A framework of analysis for assessing compliance of LAWS with IHL (API) precautionary measures

Dr. Kimberley N. Trapp[♣]

The Additional Protocols to the Geneva Conventions[1] were negotiated at a time of relative technological simplicity. Compliance with the prohibition against indiscriminate attacks,[2] supported by the obligation to take precautionary measures in planning and deciding to launch an attack (the subject of this paper),[3] was measured in terms of human effort – in gathering and assessing information about targets and their circumstances and in taking critical decisions based thereon. In the current (2016) 'Information Age', some of that human effort has been replaced by machines – in that relevant data can be gathered (by surveillance technology), assessed (in part) and disseminated remotely by computers at a rate and volume that would have been science fiction in 1977.

But technologically advanced States potentially aspire to move past this Information Age to an 'Age of Autonomous Weapons' – further cutting human effort out of the IHL compliance calculus. These technological developments – unforeseen at the time relevant treaty standards were negotiated – raise very difficult legal questions (quite apart from the obvious moral dilemma they pose): how might we assess the IHL compliance of lethal autonomous weapons systems ('LAWS')? Are the standards of IHL compliance sufficiently flexible to respond to the rate of technological development of the modern era, particularly where such development puts humans 'out of the loop' in reference to critical decision making functions? The answer to these questions is, as one might expect, rather complicated, and requires somewhat of an 'onion peel' approach.

There are in effect three layers of assessment: The outermost layer is the general international law standard applicable to the obligation to take precautionary measures under Article 57 of API. As explored in Section I below, this obligation is an obligation of conduct, not result – which is to say that compliance is measured in terms of *diligent* efforts made, not outcomes. The middle layer (explored in Section II) is informed by the general international law standard and involves a more specific assessment of compliance standards which address technological development generally, and obligations which involve the gathering and

---

♣ Senior Lecturer in Public International Law, UCL, Faculty of Laws, k.trapp@ucl.ac.uk. The framework of analysis set out in Sections I & II of this paper draws on Trapp (2013).

[1] This Chapter will focus on obligations as framed in the Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts ('API'), 8 June 1977, 1125 UNTS 3. The ICRC and States which are not party to API, however, consider the obligations that are the subject of this study to reflect customary international law. See for e.g. Henckaerts & Doswald-Beck (2005), Rules 15 and 16; Matheson (1987), 423-426.

[2] Indiscriminate attacks are prohibited under API. They are defined in part as attacks "which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated" (Art. 51(5)(b) API). This conception of 'indiscriminate' is referred to in terms of the 'proportionality' of an attack, and will be referred to as such throughout this paper.

[3] In particular, Article 57(2)(a)(i) API requires States to:

> do everything feasible to verify that the objectives to be attacked are neither civilians nor civilian objects and are not subject to special protection but are military objectives […] and that it is not prohibited by the provisions of this Protocol [including obligations regarding the proportionality of attacks] to attack them.

assessment of information in particular. Finally, the core of the relevant analysis involves specific consideration of the implications of LAWS in IHL compliance terms.

I. Precautionary measures as an obligation of due diligence

Article 57(2)(a)(i) API requires States to do ***everything feasible*** to verify (i) that a target is a military objective and (ii) that it is not otherwise prohibited by API to launch the attack (of particular relevance, that the attack would not be disproportionate), before proceeding. While 'everything feasible' sounds like a rather high standard, the obligation is nevertheless understood as an obligation of conduct, not one of result.[4] Indeed, the Commentary to API notes that 'everything feasible' is understood in terms of "everything that was practicable or practically possible",[5] making it absolutely clear that the obligation to take precautionary measures is understood in terms of efforts made. An assessment of compliance with the obligation to take precautionary measures must therefore focus on the ***process*** of verification and collateral damage assessment,[6] rather than outcomes.

Obligations of conduct, unlike obligations of result, are subject to a due diligence standard – and diligence, as a matter of international law, involves an 'available means' analysis. As a result, international jurisprudence and doctrine have highlighted the importance of accounting for available resources in assessing compliance with obligations of conduct[7] (particularly obligations to develop a capacity to keep informed, discussed further in section II below). The implication of conditioning the obligation to take precautionary measures on feasibility, practicability and diligence is that an assessment of compliance will turn (to a certain extent) on the technological means available to belligerents. While this was an accepted consequence of the way in which the obligations were framed,[8] it does mean that parties to an armed conflict which *could* do more (account taken of their state of technological advancement and available resources) cannot get away with implementing the lowest common denominator of precautions simply because their adversaries are not in the same technologically privileged position as they are. This result does not undermine the reciprocal nature of IHL obligations (which apply equally even though compliance is

[4] The distinction between obligations of conduct and obligations of result is derived from the Civil Law tradition and turns on an analysis of whether the primary rule requires absolutely that state conduct produce a certain result (obligation of result), or whether it requires only that a state make certain efforts to produce a desired, but uncertain, result (obligation of conduct). See e.g. *Case Concerning the Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v. Serbia and Montenegro)* (Judgment) [2007] ICJ Rep 43, para. 430 applying this distinction.
[5] ICRC, *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949*, <www.icrc.org/ihl.nsf/WebList?ReadForm&id=470&t=com>, para. 2198.
[6] Committee Established to Review the NATO Bombing Campaign against the Federal Republic of Yugoslavia ['ICTY Committee of Experts'], 'Final Report to the Prosecutor' (2000) 39 ILM 1257 ['NATO Bombing Report'], para. 29.
[7] See Trapp (2011), §3.1 for further detail.
[8] Commentary to API (n. 5), para. 2199.

assessed relative to particular capacity as a matter of law), but may well be the cause of some resentment and dissatisfaction in asymmetrical conflicts.[9]

II. Elements of 'Everything Feasible' in the Information Age

'Due diligence' is the general international law standard against which assessment of compliance with the API obligation to take precautionary measures is measured. In the specific circumstances of the obligation to do 'everything feasible' to verify, due diligence takes on a very particular form and is informed by technological development. While the API obligation to take precautionary measures was articulated at a time when 'feasibility' and expectations of civilian casualties would have been heavily conditioned by available technology, the obligation to take precautionary measures is nevertheless framed in flexible enough language to account for exponential technological advancements.

At the time of writing, these advancements have principally been in reference to information gathering, assessment and dissemination capabilities. In the Information Age, at least militarily developed States have access to a vast amount of information about the particular circumstances of targeted military objectives, can be kept appraised of changed circumstances in real time as a result of their persistent surveillance capabilities (resulting from information gathered by unmanned aerial vehicles ('UAVs') and satellite imagery), and are developing the networking capabilities to disseminate critical information quickly to relevant (human) actors. Compliance with the API obligation to do 'everything feasible' is therefore the product of a functional partnership between machine and human operators – with the ultimate exercise of judgment and discretion as to the proportionality of an attack, and indeed the decision to attack, the responsibility of human actors. And due diligence – measuring the efforts made to verify that an objective is military and to account for 'expected' incidental losses in launching an attack – will turn on an assessment of two separate obligations:

First, there is an obligation to use available means to develop relevant information gathering, analysis and dissemination capabilities. The Commentary to API notes the "*duty* of Parties to the conflict to have the means available to respect the rules of the Protocol."[10] The second obligation relevant to compliance with the API obligation to take precautionary measures is an obligation to put available technologies and gathered information to good and diligent use.[11] As discussed further below, both the obligation to develop relevant information

[9] See Schmitt (2005), 2-3.
[10] Emphasis added, Commentary to API (n. 5), para. 1871.
[11] Obligations which require an assessment of (and appropriately tailored action based on) information have long been considered to be composite obligations – consisting of the two equally important obligations to develop capacity (to gather/assess/disseminate information) and diligently utilise developed capacity and its fruits. See Trapp (2011), §§3.11 and 3.12.

technologies and the obligation to put those technologies to good and diligent use are obligations of conduct and subject to a diligence assessment. Were it otherwise, states with limited resources would be held to the technologically advanced standards of developed States in terms of their information gathering, analysis and dissemination capabilities, and this would effectively guarantee a breach of the obligation to take precautionary measures – no matter how diligently such states put their more limited information gathering and dissemination capabilities to use.

A. Obligation to Develop Relevant Capabilities

Concretely, measuring diligent compliance with the obligation to develop relevant capabilities will turn on a factual analysis of the socio-political circumstances of the relevant State, its general technological capabilities, and its particular development of technology relevant to battlefield identification and assessment of potential targets. For instance, States with limited resources, including relatively small intelligence budgets, or States which are not perpetually engaged in (or threatened with) armed conflicts, are unlikely to have developed persistent surveillance capabilities. To the extent that such States find themselves in a situation of armed conflict, they would certainly need to act diligently in adapting their limited existing capabilities to battlefield purposes in order to meet their API obligations. On the other hand, States which already have extensive persistent surveillance capabilities and are engaged in long-term armed conflicts would not be acting diligently if some development of those capabilities were not aimed specifically at (for instance) effective target identification.

The obligation to diligently develop relevant information capacity itself has three elements:

(i) Gathering Information

States with relevant technical resources, in particular States which are perpetually at war, would be expected to develop 'just-in-time' capabilities which enable them to gather information as to the nature and circumstances of a target on an as needed basis.[12] Such technological developments, coupled with more traditional reliance on human intelligence and reconnaissance, amount to a diligent effort to develop information gathering capabilities relevant to meeting the API obligation to take precautionary measures.

(ii) Timely analysis

Any assessment of a State's efforts to develop a capacity to analyse gathered information needs to be realistic – accounting for the volume of raw data collected and the simple present

[12] See for e.g. Best (2011), 8

day impossibility of analysing, interpreting and integrating *all* such data in a timely manner (given much of the analysis is still carried out by human operators). A realistic assessment of a State's efforts to develop relevant (and time effective) analysis capabilities also needs to be highly sensitive to competing priorities. This is because international law is silent as to the way in which a state allocates its resources, and indeed must be so given the ever increasing extent of international regulation and concomitant demands on limited financial, technical and human resources. Diligence therefore needs to be measured on the basis of a State's efforts to *improve* on timely analysis, appreciating the impossibility of entirely overcoming the limitations of human ingenuity and the current limitations of artificial 'intelligence'.

(iii) Dissemination and Usability

While timely production of intelligence may always be a difficulty given the vast amount of raw data collected, the timely dissemination of such intelligence once analysed and its usability is a key area of military capability development. Again, precautionary measures require diligence in this regard – continued efforts to respond to the circumstances of armed conflict with a view to maximal IHL compliance.

(iv) Conclusion

Diligence requires that a State be alive to the challenges of the armed conflicts in which it is engaged (for instance, the challenges of asymmetrical warfare) – particularly as regards the necessity of properly identifying targets (whether immovable, movable or human), and that its information technology development strategy is on its face be responsive to those challenges. An important part of diligent development is a good faith effort to rectify identified inadequacies, which in turn depends on a State's not turning a blind eye to technical difficulties encountered by its military in the field. This might be considered the 'lessons learned' feature of diligent development[13] – requiring technological development which responds precisely, but subject to competing budgetary priorities, to the need for accurate, timely and actionable intelligence.

B. Obligation to put developed capabilities to good and diligent use

For the purposes of assessing diligent compliance with the API obligation to take precautionary measures in the Information Age, it is assumed that critical decisions are taken by human operators. And such human operators (military commanders for instance) are held

[13] The 'lessons learned' feature of an obligation to do 'everything feasible' is implicit in the NATO Bombing Report's conclusion that states can rely on a proven track record of distinguishing between military objectives and civilians and civilian objects. See NATO Bombing Report (n. 6), para. 29. *À contrario*, where information gathering, assessment and dissemination methods have resulted in several mistakenly identified military objectives, diligence requires a State to re-evaluate its processes and address any inadequacies.

to a standard of reasonableness in their critical decision making, in particular decisions as to whether a target is indeed a military objective and any proportionality calculus necessitated by the circumstances of the military objective.[14] Any assessment of individual compliance with the obligation to take *feasible* precautionary measures (including in reference to a military commanders' *expectation* of civilian losses) will necessarily draw on the level of technological advancement of the State on which he or she depends. This is because compliance is assessed on the basis of information *reasonably* available *in the circumstances*,[15] and such circumstances will to a large extent be driven by a State's development of information gathering, analysis and dissemination capabilities (coupled with certain temporal factors in respect of the particular attack).[16]

III. LAWs and API Compliance

The final layer of IHL compliance examined in this paper is how presently applicable standards of compliance (explored in Sections II & III above) adapt to further technological development – whereby information is assessed and actioned within a single weapons system without human initiation or further intervention. This paper will address the case of LAWS which operate at the highest levels of autonomy and with humans 'out of the loop' (in that the weapons system performs critical functions – target acquisition, tracking, selection, and attack – without human initiation or intervention). The measure of interaction between human and machine which is a feature of compliance with IHL in the Information Age would thereby be limited to before the fact programming and parameter setting.

The framework of analysis set forth in Section II above assumes a functional partnership between machine and human operators – with the ultimate exercise of judgment and discretion as to the proportionality of an attack, and indeed the decision to launch an attack, the responsibility of human actors. This feature of Information Age compliance with API obligations is precisely what is missing from LAWS, where analysis of available information, proportionality judgment based thereon and kill decisions are folded into a single weapons system without a 'human in the loop'. The implications of this 'consolidation' of relevant tasks in assessing API compliance are set forth below.

[14] ICTY, *Prosecutor v Galic*, IT-98-29-T, Trial Chamber Judgment and Opinion (5 December 2003): "it is necessary to examine whether a reasonably well-informed person *in the circumstances* of the actual perpetrator, making reasonable use of the information *available* to him or her, could have expected excessive civilian casualties to result from the attack" (emphasis added).
[15] See State practice reviewed in Henckaerts & Doswald-Beck (2005), 363-5.
[16] See Trapp (2013) for an analysis of temporal factors affecting an assessment of compliance with precautionary measures.

A. Elements of API compliance fold into each other

In cases of a functional partnership between machines and human operators during an armed conflict, there is a logic to splitting API compliance into two separate elements, the one in reference to capacity and technological development (which will facilitate compliance with API obligations), the other in reference to the use to which that capacity and technology is put by human operators (in compliance with API obligations). In the case of LAWS, however, the capacity / technology (and its development) is not separate from its end use – even if not all the information gathering is done within the one weapons system, the analysis of the gathered data and critical decision making is all the one weapons system or at least part of the same weapons system network. The two separate elements of a diligence assessment in respect of compliance with API obligations in the Information Age therefore collapse one into the other in an Age of Autonomous Weapons. Assessing compliance with an obligation to do 'everything feasible' becomes a significantly more focused judgment – as everything hangs on the development and testing of the technology.

This has clear implications for Article 36 API – which imposes an obligation on State Parties to 'determine whether [a weapon's] employment would, in some or all circumstances, be prohibited' by the Protocol. In the Information Age, these assessments tend to focus on the precision of the weapon. In an Age of Autonomous Weapons, Article 36 assessments need to focus on the entire range of API obligations – and indeed it may be impossible to determine, *a priori*, whether LAWS can comply with the obligation to take precautionary measures. This is not least because artificial intelligence now in development would potentially allow for weapons systems to 'learn' from experience. Any weapons system subject to the rigours of Article 36 in laboratory conditions will therefore be different – an earlier development of – the weapons system later deployed in combat situations.

B. Decrease in margin of appreciation

The second implication of LAWS in assessing compliance with the API obligation to take precautionary measures is in reference to the margin of appreciation States enjoy in the Information Age. In particular, States enjoy a margin of appreciation in meeting the capacity development obligations inherent in due diligence obligations. As discussed above, due diligence obligations are conditioned by an 'available means' analysis, and the margin of appreciation is an important aspect of any such resource based analysis. This is in part because States are faced with competing priorities, and how they manage their resources will depend to a large extent on the nature and number of threat they face, and the acuteness of any such threats. International law therefore has very little to say about how states should prioritize resource allocation.

But, in respect of LAWS (assuming the higher end of the autonomy spectrum and the lower end of the human in the loop spectrum, with weapons systems selecting and engaging targets without any initiating or further intervention by a human operator) – the margin of appreciation that even technologically advanced States enjoy has to decrease. This is because everything hangs on the technology in an 'everything feasible' compliance calculus – and that of course increases the standard against which capability development is measured (as 'diligence' is not shared between technological development and human end use).

That the margin of appreciation decreases, perhaps even significantly, is made clear in the relevant literature. For instance, in the ICRC Report produced after the 2014 round of expert meetings, one commentator suggested that if it is not possible to 'guarantee' that the weapons system would comply with IHL in all circumstances, then it would be unlawful.[17] Such statements perhaps go somewhat farther than legally warranted (in that capacity development, as part of the precautionary measure obligation, is subject to a due diligence standard, while *guaranteeing compliance* is what would be required of an obligation of result). The standard of compliance adapted to technological developments between 1977 and the Information Age, and 'everything feasible' will have to respond to an increasingly *exclusive* technological environment. And it is clear that when everything hangs on the technology, 'diligence' will require significantly more of States than is presently required.

C. Human in / out the Loop

The final implication of LAWS in assessing API compliance is in reference to the 'human in the loop' issue (assuming again the lower end of that spectrum for the purposes of discussion). For the foreseeable future, given the stage of Artificial Intelligence ('AI') and Artificial Learning Intelligence development, 'everything feasible' should be measured against a 'human in the loop' standard – precisely because it is always feasible for them to be so. Again this increases the standard against which technology is measured when determining compliance with precautionary measures. Until machines can exercise judgment, and engage in a balancing act that even those who regularly do so might not be able to explain (much less programme AI to do), the 'human in the loop' standard against which autonomous weapons compliance will be judged is simply too high for such weapons systems to pass Article 36 API review.

[17] ICRC (2014), p 22.

IV. Conclusion

While the obligation to take precautionary measures and to do 'everything feasible' is of universal application, assessment of compliance must have regard for the very particular circumstances of each Party to an armed conflict. It is therefore true that great resources mean great responsibility. And States which have the resources available to even contemplate the development of LAWS have great responsibilities indeed.

The flip side of great responsibility is that great resources engender great opportunities. In particular, the capacities developed in the Information Age (including to gather and disseminate information remotely) give States the freedom to counter the dangers to which their national forces are exposed in meeting API obligations to take precautionary measures. It is only right that States which have the technological means available to put their armed forces out of harm's way in conducting war should also be held to the standard of *feasibility* suggested by those technological means in protecting civilian populations from its ravages. In this regard, LAWS goes one significant step further than current technologies – as it the ultimate force protection weapon. And IHL simply does not allow for a State to prioritise protecting its own forces above protecting the innocent. Given the increased standards of IHL compliance in reference LAWS discussed above, it is difficult (nay…. near impossible) to see a path to IHL precautionary measure compliance for LAWS.

BIBLIOGRAPHY

J Beard, 'Law and War in the Virtual Era' (2009) 103 AM. J. INT'L L. 409

R Best, *Intelligence, Surveillance, and Reconnaissance (ISR) Acquisition: Issues for Congress* (Congressional Research Service, December 2011).

J-M Henckaerts and L Doswald-Beck, *Customary International Humanitarian Law*, Vol I: Rules (CUP 2005)

I Henderson, *The Contemporary Law of Targeting: Military Objectives, Proportionality and Precautions in Attack under Additional Protocol I* (Martinus Nijhoff 2009), 159

ICRC, Expert Meeting, 'Report on Autonomous Weapon Systems: Technical, Military, Legal and Humanitarian Aspects', Geneva, Switzerland, 26 - 28 March 2014 (available at https://www.icrc.org/en/download/file/1707/4221-002-autonomous-weapons-systems-full-report.pdf)

MJ Matheson, 'Remarks' in *Sixth Annual American Red Cross-Washington College of Law Conference on International Humanitarian Law: A Workshop on Customary International L aw and the 1977Protocols Additional to the Geneva Conventions* (1987) 2 AM. U. J. INT'L L. & POL'Y 419

GS McNeal 'Targetted Killing and Accountability,' (2014) 102 GEORG. L.R. 681

APV Rogers, *Law on the Battlefield* (2nd ed, Juris Publishing 2004)

KN Trapp, 'Great resources mean great responsibility: a framework of analysis for assessing compliance with API obligations in the information age', in Dan Saxon (ed.), *International Humanitarian Law and the Changing Technology of War*, Martinus Nijhoff (2013), 153.

KN Trapp, *State Responsibility for International Terrorism* (CUP, 2011)

M Schmitt, 'Precision Attack and International Humanitarian Law (2005) 87 INT'L REV OF THE RED CROSS 445

M Schmitt, 'War Technology, and International Humanitarian Law', Occasional Paper Series, Program on Humanitarian Policy and Conflict Research, Harvard University (July 2005), <www.hpcrresearch.org/sites/default/files/publications/OccasionalPaper4.pdf>