

CCW Meeting of Experts on « LAWS »

Panel on possible challenges to IHL due to increasing degrees of autonomy

[Introduction]

- To begin with, as mentioned in the ICRC's general statement on Monday, and as with many of you, **our thinking and views** about "Lethal Autonomous Weapon Systems" (or Autonomous Weapon Systems, as I will call them) **continue to evolve** including on the legal issues raised by these weapons.
 - In the little time I have, I will outline some open questions raised by AWS under IHL rules governing the conduct of hostilities (what Bill Boothby has referred to as "weapons law" and "targeting rules"), including some questions that may require clarification.
- Indeed, although there is **no dispute** that, like any new means and methods of warfare, AWS must be used and be capable of being used in accordance with the rules of IHL, **difficulties in interpreting and applying existing IHL rules** to these new technologies of warfare may arise in view of their unique functional characteristics, the intended and expected circumstances of their use and their foreseeable effects.
 - Ultimately these challenges may raise questions whether existing law is sufficiently clear and whether there is a need to clarify it or develop new rules to deal with these challenges.

[Frame of discussion]

- As we have oft^{en} repeated, AWS refer to **weapon systems with autonomy in the critical functions of selecting and attacking targets**, regardless of the level of sophistication, i.e. a weapon in which there is no "meaningful" human control over these critical functions. ["when a machine selects a specific target and the human is unaware of which particular targets are to be engaged, then the machine is in control" – Scharre].
- There may be many different kinds of AWS, depending on specific functions of the weapon and the purpose for which it has been developed.
- In this respect, (as UNIDIR has pointed out in its 1st report on the Weaponization of Increasingly Autonomous Technologies), **autonomy should be seen as a characteristic of a technology**, attached to a function or functions of a weapon system, and not the weapon system itself.
 - In this sense, "autonomous weapons" can be seen as an **umbrella term** covering a range of different kinds of weapon systems, each of which differs from the other based on a number of factors or variables.
 - This point is relevant to the general question of whether or not AWS are unlawful: I suggest that the answer to this narrow question is that "it depends". It depends

on the specific kind of AWS and its ability to, **predictably** and **reliably**, respect the relevant IHL rules in all of the circumstances in which it is expected to be used.

- As we heard yesterday from a number of speakers, including my colleague Dr. Neil Davison, the factors relevant to the ability of a specific AWS to respect IHL may include:
 - The type of task the particular weapon system is being used for (offensive v defensive)
 - The context in which it is being used (air, ground, sea; simple v cluttered environment)
 - The type of target (material or personnel)
 - The type of force it is using (kinetic v non-kinetic)
 - The freedom of the weapon to move in space (fixed or mobile, narrow or wide geographical area)
 - The time frame of the action of the weapon (does it attack only at a specific point of time or over a longer period of time)
- These factors are **crucial to determining the foreseeable effects** of the particular weapon – thereby ensuring that it can be used in conformity with the rules of IHL.
 - The **kind and degree of human control that would be required for compliance with IHL** would vary as a matter of law according to these factors and the specific characteristics of the particular AWS (including the degree of autonomy programmed in its critical functions).

[Challenges to respect for IHL posed by predictability and reliability of the weapon]

- As weapon systems become more autonomous, are increasingly able to determine their own actions, and are deployed for lengthier periods and in more complex environments, the results/outcomes of their use will necessarily become **less predictable**.
 - “FAWS by def’n unpredictable => if you could create a perfect algorithm which you have near-100% certainty that it will perform as well, or better than, the IHL-compliant soldier => arguably this might meet the predictability standard.
- The degree of **predictability** (which requires knowing what precisely the weapon will do; the results/outcomes of use), which also depend on the factors I have listed, are crucial considerations both at the legal review stage **and**, if the weapon passes the legal review, at operational stage by military commanders **when planning and deciding to launch attacks** using AWS.
 - The key question at the review stage is how to determine whether the AWS will operate in accordance with IHL if its performance is unpredictable?

[Legal Reviews of AWS]

- **A new weapon system with autonomy in its critical functions**, like any new weapon system, must undergo a **rigorous legal review**, involving the requisite multidisciplinary expertise, to determine whether it can be used in accordance with IHL rules. (In other words, the obligation to carry out a legal review of new weapons creates an onus on the State developing or acquiring the new weapon to show that it can be used in accordance with IHL.)
 - The ICRC encourages States that have not yet done so to establish weapons review mechanisms and stands ready to advise States in this regard. In this respect, States may wish to refer to the ICRC's *Guide to the Legal Review of New Weapons, Means and Methods of Warfare* (copies at the back of the room).
- (As Bill Boothby has mentioned), reviewers must assess the weapon's lawfulness in relation to the normal or expected circumstances of its use. This requires **foreseeing** how the weapon will perform in the circumstances and environment in which it is intended to be deployed (the factors I have just mentioned), based on the weapon's design and how it actually functions.
 - This will enable the review to set clear parameters on their operational use and deployment. If it is found that it could be used lawfully only in limited circumstances, **these limits on use** imposed must then be incorporated in the instructions and rules of engagement applying to the weapon, to ensure it is not misused.
 - I should stress that the permitted circumstances of use may in some cases be so limited and complex, and therefore unrealistic to apply in real-world scenarios, that it may be more appropriate to prohibit the weapon's use altogether.
 - **A significant challenge in reviewing the legality of an AWS** will be how to test it to be sure that it will do what the human operator wants it to do, in other words **how to test the weapon's predictability**.
 - Predictability about the way an AWS will interact with its environment must be sufficiently high to allow an accurate legal review. As I said before, the greater the autonomy in the tasks of the weapon, the lesser the ability to determine whether the system will function in accordance with IHL.
 - If at the legal review stage, it is not possible to **accurately predict**, through testing or otherwise, whether or not the weapon will respect IHL, **then surely it cannot pass the legal review**.
 - A key challenge in this respect is **how to properly test the AWS** to ensure it can be used in accordance with IHL. As we have heard, there are no standards for testing.

[Operational decisions to use AWS]

- **Predictability** of the weapon system – again, meaning its foreseeable effects -- is also relevant at the operational stage, assuming the weapon has passed the legal review:

- Indeed, not all legal questions will have been resolved at the legal review phase. The rules of distinction, proportionality and precautions in attack are applied contextually, on a case-by-case basis, by commanders and soldiers the field. It is the commander that plans and decides upon the attack, using the means of warfare at his/her disposal and taking into consideration the conflict environment.
- In this respect, **a commander must fully understand the capabilities of the weapons at his/her disposal**, in order to be able **to foresee the weapon's effects** on a target, and its potential incidental (or collateral) effects, and therefore allow him/her to comply with what is required by IHL rules on the CoH. This is what some militaries refer to as “weaponeering” – i.e. matching the choice of the weapon to fit the target in the particular environment. This requires **predictability** regarding how the weapon will perform. This may be a **challenge** for AWS used in certain environments.
 - If a commander **knowingly** deploys a weapon system the effects of which are unpredictable => accountable for any violations that may occur as a result.
- In relation to weapons with autonomy in the critical functions of selecting and attacking targets, there are a number of stages at which IHL rules on targeting may or must be considered applied. Stage 1: the programming of an AWS (to define its task, its target, and its freedom of action in time and space); Stage 2: the deployment of an AWS, which involves a commander or operator's decision to activate and deploy a particular weapon system to achieve a particular aim in a particular context; and Stage 3: the actual targeting phase, i.e. where the AWS selects and attacks specific targets.
 - It is clear that the programming (1) and deployment (2) stages by definition involve human intent and control.
 - It is in stage 3, the targeting phase -- where an AW would select (i.e. search for or detect, identify, track) and attack (i.e. intercept, uses force against, neutralise, damage or destroy) the target – that the question arises whether IHL requires human involvement in these critical functions.
 - No doubt that the IHL rules on the conduct of hostilities are addressed to human beings: while the primary subjects of IHL are States, IHL rules of distinction, proportionality and precautions in attack are addressed (implicitly or explicitly) to those who plan and¹ decide upon an attack.
 - These rules create obligations for human combatants and fighters which are responsible for respecting them, and will be held accountable for violations.
- If a commander deploys an AWS (stage 2), which will itself independently select targets and/or independently determine the moment when to attack the selected target, in other words, where the actions of the weapon are **unpredictable**, is this permissible under IHL?

¹ Art.57(2) says “or”, but Commentary interprets this as “and”, based on teleological interpretation. Boothby agrees.

- From a strictly legal standpoint (leaving aside the questionable military utility) deploying a weapon system the effects of which are wholly or partly unpredictable creates a **significant risk** that IHL will not be respected.
- Here, the person ordering the deployment of the weapon “will be expected to make a reasonable decision about the appropriate amount of risk” that is acceptable in using an AWS (Turner). Can the AWS perform in accordance with IHL in the specific context and given the specific circumstances in which it will be deployed? To make such a judgement, the commander must have a thorough understanding of the AWS’ particular capabilities, and of its limits. He/she must understand how the AWS will react in the specific environment in which it will be deployed. Again, this requires sufficient predictability of how the weapon will perform.
- In sum, there would be a need for States to examine more closely, in their deliberations on the legal issues raised by AWS, how to assess predictability of AWS (and for that matter reliability, which I have not gotten into for want of time).

[Questions for consideration of States at multilateral level]

- Some argue that the faithful implementation of IHL’s general rules – including through the legal reviews of new weapons and proper training of the users of these weapons -- should be sufficient to address any humanitarian (including ethical) concerns regarding autonomy in the critical functions of weapon systems.
 - If some **specific usages** of a specific weapon system are problematic under IHL, there may be a need for technical fixes, or limits on use, or outright prohibition of the specific weapon system.
 - However, history has shown that when using certain weapons, the **specific characteristics of the weapon combined with the inconsistent application of the general rules of IHL to that weapon** (be it at the legal review stage or at the operational stage) may reveal a need to clarify the law and ultimately **develop weapon-specific rules**.
- What is clear is that there are too many questions regarding the interpretation of IHL rules as they apply to AWS, including issues of accountability for violations (which I do not have time to address here) to leave to national legal reviews.
 - Leaving it up to each State to determine the lawfulness of specific AWS they are developing or acquiring risks inconsistent application of IHL, with e.g. some States applying limits, and others not.
 - This risk, as I see it (at least in the future), relates notably to a possible interpretation leading to absence of human judgment – and reliance instead on the AWS’ algorithm – to carry out the highly context-based assessments required by the IHL rules of proportionality and precautions in attack.
 - Indeed the many questions about AWS raised under IHL (not to mention the ethical and other questions), underscore the need for continued discussions in multilateral

forums such as the CCW, with a view to considering various policy and other options available to States to address the legal (and other) challenges raised by AWS.

[Conclusion]

- To sum up, based on what we have heard in the last two days, today the **main obstacles to developing IHL-compliant AWS appear to be technological**.
 - In this respect, based on current and foreseeable technology, there are **serious doubts** about the ability of AWS to comply with IHL rules of distinction, proportionality and precautions in attack **in all but the narrowest of scenarios and the simplest of environments**. (These narrow scenarios and predictable environments are those in which current AWS operate today, most if not all of them under human supervision.)
 - But there may very well be a day when advances in robotic technology will lead to the development of AWS that are capable of independently selecting their targets and determining their own actions, outside of tightly constrained spatial and temporal limits, while fully complying with IHL rules on targeting – including systems capable of autonomously performing the highly contextual evaluations required by the rules of proportionality and precautions in attack. This remains a hypothetical, futuristic scenario, but one that should be contemplating when discussing the legal and ethical implications of AWS.
 - The challenge (as UNIDIR has pointed out) is whether it is **technically possible** (and desirable) **to replace human judgment with data readable by a computerized system** (“data-based proxies”) for making the assessments required by the IHL rules of distinction, proportionality and precautions in attack (UNIDIR, no.2, p.8).
 - Undoubtedly, the absence of human judgment in these assessments represent “a paradigm shift and a major qualitative change in the conduct of hostilities” (2011 Challenges report).
 - The question boils down to whether human decision-making in programming and deploying the weapon is sufficient to fulfil human responsibilities as subjects of IHL, or whether a point is reached where humans are so far removed (in time and space from the selection and attack of targets) that the human decision-making process regarding the use of force is substituted with machine decision-making.
 - In this respect, there remains the crucial question of whether an AWS, notably those with the ability to **select and attack human targets without human intervention**, would be acceptable under the principles of humanity and the dictates of public conscience, as expressed in the Martens Clause, an overarching issue that this meeting will address in more depth later this week.