

Convention on Certain Conventional Weapons (CCW)
Group of Governmental Experts on Lethal Autonomous Weapons Systems
25-29 March 2019, Geneva

Statement of the International Committee of the Red Cross (ICRC)

under agenda item 5(e) Possible options for addressing the humanitarian and international security challenges posed by emerging technologies in the area of lethal autonomous weapon systems in the context of the objectives and purposes of the Convention without prejudicing policy outcomes and taking into account past, present and future proposals

Thank you Mr. Chairperson,

As we have discussed this week, autonomous weapon systems raise **concerns about loss of human control over the use of force**, which could have serious humanitarian consequences, in terms of adverse consequences for both civilians and combatants in armed conflict.

As the ICRC expressed yesterday, these unique characteristics of autonomous weapon systems – namely the loss of the ability of combatants to exercise the context-specific judgments required of international humanitarian law (IHL) rules, and the loss of human agency and diffusion of moral responsibility in decisions to use force -- **raise challenges for IHL compliance and for ethical acceptability** and, therefore, raise the question of whether new internationally agreed policies, standards or rules are needed.

The **common ground** on which States can build, and have been building in their interventions this week, is their agreement in the 2018 GGE that **human responsibility – or control** – must be retained over weapon systems and the use of force.

As the ICRC has called for in its Working Paper on the “Element of Human Control” submitted to the CCW Meeting of High Contracting Parties last November (CCW/MSP/2018/WP.3), this **“human-centred” approach** must guide the development of limits on autonomy in weapon systems, and in particular the practical elements of human control over the critical functions of weapon systems needed for legal compliance and ethical acceptability.

The Provisional Programme of Work lists a number of options that have been proposed for addressing the humanitarian and international security challenges posed by autonomous weapon systems. These options, which are “not necessarily mutually exclusive”, include: **a legally binding instrument; a political declaration; guidelines**, principles or codes of conduct; and improving implementation of existing legal requirements, including **national legal reviews of new weapons**.

As the ICRC stated in its Working Paper, all of these approaches share the same need to develop **common understandings of the type and degree of human control necessary in practice** to ensure compliance with IHL, and ethical acceptability. This effort must be driven by the necessity to **preserve human judgement and responsibility in targeting decisions**, where **human supervision, predictability** and **context** are important factors.

Working on the parameters of a **positive obligation for human control** will also enable States to **identify “autonomous weapon systems of concern”** that fall outside of effective human control, and would therefore be unlawful and/or ethically unacceptable.

Insofar as the sufficiency of existing law – and in particular of IHL – is concerned, it is clear, as the ICRC submitted again yesterday, that **existing IHL rules** – in particular distinction, proportionality and precautions required in attack – **already provide some limits** to autonomy in weapon systems, insofar as they establish responsibilities for combatants, who must retain the ability to make the context-specific assessments required of IHL rules during the operation of the weapon system. We described the IHL limits in more detail in our statement.

However, it is also clear that existing IHL rules – while already limiting autonomy in weapons to a certain degree – **do not provide all the answers**. Key questions include:

- What is the type and degree of human control over weapon systems, including the level of human supervision and spatial and temporal constraints on their operation, required to ensure compliance with IHL rules?
- What are the minimum levels of predictability (or maximum levels of unpredictability) that are tolerable in the functioning of weapon systems, to ensure compliance with IHL rules?

Debates in the GGE this week and in the last two years have shown that there are **serious questions** whether existing IHL is sufficiently clear, or whether there is a need to **clarify IHL or to develop new rules** – i.e. new law. This again points to the urgent need to, at minimum, reach common understandings on the practical elements of human control.

The ICRC has welcomed the attention that this discussion has brought to improving and developing legal review processes. Effective **legal reviews are critical** to ensuring that a State’s armed forces comply with IHL in light of rapid technological developments. **Greater transparency** in how States interpret and apply their obligation to carry out legal reviews of new weapons to new technologies can help to identify good practices and thus assist States seeking to comply with their legal review obligations, particularly when reviewing autonomous weapon systems, whose unique characteristics can make legal reviews more difficult. But the bottom line is that while robust legal reviews remain essential, they are **not a substitute** for States working towards internationally agreed limits on autonomy in weapon systems.

Moreover, the limits dictated by **ethical considerations** may go beyond those found in existing IHL rules. In particular, ethical concerns have been expressed about the **loss of human agency** in decisions to use force, **diffusion of moral responsibility** and **loss of human dignity** are most acutely felt with autonomous weapon systems that present risks for human life, and **especially with the notion of anti-personnel systems**, that is, those designed to target humans directly.

As a result, **public conscience may demand limits or prohibitions on particular types of autonomous weapons** – such as anti-personnel systems – and/or their **use in certain environments** – such as where civilians and civilian objects are present.

Perhaps these **ethical boundaries are already evident with existing autonomous weapons**, which generally only target objects and not humans, and are used in very narrow circumstances for defensive

purposes in environments where there are few civilians. And even in these narrow circumstances, human supervision and ability to intervene and deactivate is retained.

Regarding a proposed **ban on so-called “fully autonomous weapon systems”** -- such as weapons that are unsupervised, unpredictable and unconstrained in time and space (that is, beyond any human control) -- the ICRC notes that many States are of the view that such weapons would almost certainly be unlawful by their very nature (a view that it shares), and many have expressly stated they have no intention of developing and using such weapons. Yet the ongoing militarization and weaponization of autonomous technologies demands a standard of human control that is relevant to current and emerging developments as well as to future technological and operational developments.

Most States appear to agree that the **yardstick** against which the legal and ethical acceptability of all existing and future weapon systems with autonomy in their critical functions must be assessed is **meaningful (effective or significant) human control**, or put another way, ensuring appropriate levels of human judgment in decisions to use force. **However, the question remains:** what would be the practical elements (or criteria) of this positive obligation of human control?

In terms of **articulating a standard of human control**, and therefore of limits on autonomy in weapon systems and the use of force, could a requirement for human control (responsibility) take **one of these two possible approaches?**

- A requirement for direct or remote human control over all weapons i.e. that would generally exclude autonomous weapon systems, with specific exceptions? **Or**
- A requirement for meaningful / effective / substantive human control or judgment over all weapons with autonomy in their critical functions, with specific prohibitions?

In terms of **the type and degree of human control** that would be needed to comply with IHL and satisfy ethical concerns:

- Must there be **constant human supervision**, with the ability to intervene and de-activate, for the duration of the weapon's operation? And what would be the exceptions, if any, to this requirement?
- Must there be a requirement that the human operator be able **to predict, with a high degree of certainty**, that the weapon will attack a specific target at a specific point in time, and its effects? And what would be the exceptions, if any, to such requirement?
- What are the standards of **reliability** of weapons with autonomy in their critical functions, also as regards testing?
- Assuming that autonomous weapons that are designed to select and attack materiel targets (objects) are acceptable, **what operational constraints** apply, in particular to the environment in which they operate (e.g. populated or unpopulated area), the duration (time-limit) of their operation? What would be the spatial and temporal limits that would be applied to mobile (as opposed to stationary) systems?

The ICRC believes that these key questions can help guide efforts to identify a standard of human control that is clear, robust and practical, and that withstands the test of time.

Thank you.