



CCW REPORT

Civil society perspectives on the CCW Group of Governmental
Experts on lethal autonomous weapon systems
9-13 April 2018



Reaching Critical Will

12 April 2018
Vol. 6, No. 4

IN THIS ISSUE

- 1 | **Lethal
[autonomous]
weapon**
- 6 | **Side event report:
Meaningful
Human Control
of Autonomous
Weapons**

The CCW Report is produced by the Reaching Critical Will programme of the Women's International League for Peace and Freedom (WILPF). WILPF is a steering group member of the Campaign to Stop Killer Robots.

The views expressed in this publication are not necessarily those of WILPF or the Campaign.

Editor: Allison Pytlak

www.reachingcriticalwill.org
info@reachingcriticalwill.org



LETHAL [AUTONOMOUS] WEAPON

Allison Pytlak | *Reaching Critical Will of Women's International League for Peace and Freedom*

In the 1987 action/comedy film *Lethal Weapon* a main character remarks, "I don't make things complicated. That's the way they get, all by themselves."

While not necessarily a fan of the film, it came to mind for two reasons. Apart from the fact that its title bears obvious relevance to the name of the Group of Governmental Experts (GGE) meeting this week, the concept of lethality in autonomous weapons has started to figure more regularly in GGE discussions. The word "lethal" is being questioned by a widening group of states in the context of whether it is an adjective or qualifier that should continue to be placed in advance of the term "autonomous weapon system". On Wednesday, Poland joined others that have made this point earlier in the week. One reason for removing this word is, as Switzerland noted, that keeping it overlooks autonomous weapons that do not necessarily inflict physical death, but lead to physical injury. The International Committee of the Red Cross (ICRC) noted that it is not lethality, but the use of force that triggers legal obligations under international humanitarian law (IHL) and international human rights law (IHRL). Civil society organisations have made similar points at this and other meetings.

A second reason why this film came to mind is because this situation is becoming complicated—all on its own. Wednesday took a hard focus on the "human element" in the use of lethal force. All governments are united in recognising the importance of human control but views differ sometimes significantly at which point in a system's life cycle it is needed, and at what threshold. The GGE chair distributed visual aids to map out trends from the discussion on Wednesday, including a diagram outlining a spectrum of autonomy, and a list of the words used most frequently by states to describe thresholds of human control, such as "meaningful" or "sufficient". This was useful in enabling discussion with

greater detail, including with examples of real and hypothetical situations that developed in a natural way, however a few states felt that this was distracting and asking things overly complicated. China urged re-focusing discussion away from these questions, and—somewhat ironically for the purposes of this editorial—back to the question of lethality.

Another dimension that surfaced often on Wednesday is legal responsibility, in that humans, not machines, are beholden to law. This reinforces the importance of human control particularly for reasons of accountability under international law, including international humanitarian law and international human rights law, as noted by Human Rights Watch.

Human control is at the heart of the issue of autonomy and necessary, as Brazil said, to meet the dictates of public conscience and the principles of humanity.

It's good to get into the weeds a bit, and have a debate that actually is a debate—provided that input will soon be channeled to move the process forward, and that these are the right weeds to be in. This is the fifth time that countries have convened at the United Nations in Geneva to address the problems these weapons would pose if they were developed and put into use. Time is passing, and technology developing exponentially. It's vital to move to the next stage, and into taking action.

Below is a non-exhaustive summary of statements delivered on this subject, with special focus on views about the nature of human control, where to situate it, and issues of accountability.

continued on next page



Lethal [autonomous] weapon, continued

Thematic debate: Further consideration of the human element in the use of lethal force; aspects of human machine interaction in the development, deployment and use of emerging technologies in the area of lethal autonomous weapons systems.

Wednesday's discussion about the "human element" demonstrates that all states present in the GGE continue to acknowledge the centrality of ensuring adherence to IHL and the necessity of human control despite differences of opinion about its extent and nature. Some states addressed the issue of legal accountability, with most noting that this rests with humans and not with machines.

Germany stated that humans must maintain control over the critical functions to the degree necessary to comply with IHL. It identified many points of a weapons' life cycle where there are human-machine interaction, such as the software engineers that create algorithms, through to engineers, those involved in political decisions about procurement, and those taking decisions about deployment.

The UK insists that lethal autonomous weapons (LAWS), if they existed, must be held to the same degree of review as other weapons. Accountability is written into the Article 36 review process, and once deployed, accountability is established through the military chain of command. The UK stresses the importance of some minimum degree of human oversight.

Poland noted that lethality as a characteristic does not add much to the conversation. It prefers the use of "fully autonomous weapon systems". It feels that humans should play a decisive role in the deployment and use of force. Poland spoke about the ethical dimension of this issue as being particularly relevant and that as it is deeply-rooted in our human nature to control the environment, it would be against human nature to create something outside the controllable environment. Like the UK, it feels states should continue to improve the Article 36 review process and to establish limits.

Egypt took note of the statements linking human control to the engagement of critical functions, but it is not clear what engagement means. Egypt believes that selection of targets must be subject to human control, without this the weapon is indiscriminate. It feels that autonomous weapons should have a "kill switch" and human control maintained at all stages of the weapon life cycle. Direct responsibility still lies with the human operator.

Brazil stated LAWS should comply with the Martens Clause. It feels that broad terms such as "on the loop" are not helpful because they do not explain

the degree of meaningful human control. Meaningful human control has to have the capacity to deploy and also to limit a machine's action. Another aspect to be considered must be the machine's impact on the human operator's decision-making process.

Egypt stated that LAWS is an umbrella term that could include many potential and very different elements, ranging from fully autonomous weapons functioning as biological agents with free will; or weapons systems that once deployed by a human operator are not subject to human control; or weapons that once switched on could still be overridden by humans, or that remain fully controlled by humans, like drones. It said that focusing on critical functions may be distracting for the fact that we lack a definition of critical functions.

The International Committee for Robot Arms Control stressed that it will create confusion to broaden the discussion into weapons with AI or emerging AI. It feels that the ICRC definition, as concerned with autonomy in critical functions, is sufficient for the GGE's definitional purposes. ICRAC supports those states that have stated that the focus should be on human control and human-machine interaction. Earlier on Tuesday it published a set of "Guidelines for the human control of weapons systems".

Costa Rica, with a view to advancing the discussion on the topic, recommended avoiding an unending discussion on the characterisation of technical components. States should look at the issue of human control.

Sierra Leone, noting that since autonomy can be gradually increased or decreased, determining when a system becomes fully autonomous does not serve the interests of the GGE. It urged concentrating on level of harm, and observed that what is emerging from the discussion is that states should focus on autonomy in critical functions as basis for further work.

Panama preferred more general definitions and would like to focus on human control of critical functions, such as engagement with targets, so as to ensure inclusion of ethical and legal considerations.

Canada suggested mapping out the functions of LAWS, and the degree of autonomy allowable, as an aid to reaching a definition. The technological dynamics of "systems of systems" that must be considered.

Throughout, some delegations (France, UK, Russian Federation, Ireland, among others) referred to their existing definitions, which are contained in the compilation document.

Lethal [autonomous] weapon, continued

New Zealand observed that meaningful human control in one situation might be insufficient in another. It recommends further exploration of the concepts of “human in the loop” and “human on the loop”, in order to organise our collective thinking. It also raised questions about how will a human know when human intervention is necessary, and how does the role of bias in programming affect these weapons?

The US explained that its Department of Defence policy mandates that autonomous and semi-autonomous weapons be designed to allow for appropriate human judgement, and that weapons are also rigorously tested to allow for appropriate human judgement.

Japan emphasised that meaningful human control is crucially relevant to weapons systems because the deployment of systems that select and target without it is not acceptable from a moral point of view. The human operator must have effective control over the weapon to ensure moral responsibility and legal accountability.

China recommends that other states explain how they use the term human control using technical indicators and not generic standards, noting that terms are helpful for explaining what does not constitute LAWs, but cannot replace a concrete definition. For China, the issue of lethality is a fundamental issue of humanitarian concern; removing it will detach the discussion from the Group’s mandate.

Netherlands believes that fully autonomous weapon systems should not be developed, as they could alter their mission without human oversight. At the same time, it does not see autonomous weapons as taking over the majority of military operations. It agrees with other states that the entire targeting cycle of the weapon (explained as design, deployment, and assessment of collateral damage) requires meaningful human control.

Pakistan believes a lack of meaningful human control will fundamentally alter the nature of warfare and that the goal for the GGE should be to ensure that it is retained, although it noted that while this concept has gained some traction, it does not on its own provide a definition or regulatory framework.

Russia agrees with preserving human control over LAWs, and emphasised issues of responsibility and accountability. Legal accountability rests with states and individuals; in the latter case it is those who do the programming, or authorise or operate the weapon in question.

Switzerland feels that this GGE is a chance to clarify and advance various concepts. It believes that it is important to study the degree of control that must be exercised at various stages of a machine’s life cycle, suggesting that through guidelines and training of armed forces, the degree of predictability can be increased but

that those making the decision to deploy an autonomous weapon must consider all circumstances at the time of deployment. Distinction and proportionality requires individual value judgments that machines cannot perform at present.

Ireland raised the point that since autonomous weapons are generally created through computer programming, this raises the question of bias, including gender bias. It feels that the question of ensuring effective human control requires that a human operator make the decision to employ force against a target and monitor the use of force during the subsequent attack.

Kazakhstan feels human control should be “direct and permanent” in all stages of the development, deployment, and use of an autonomous weapon. It acknowledged that the proliferation of terminology in the GGE raises concerns about clarity.

Chile stressed the importance of also retaining control over accountability and responsibility, particularly as technology advances. It also asked several relevant questions to consider in the context of human control, such as: can a machine respond legally for its actions? Can discernment be programmed? Can a machine comply with IHL?

South Africa agrees with simplifying the discussion to focus on the regulatory framework. It is of the view that human control in the selection and use phases must be maintained, with a focus on accountability and responsibility in accordance with international law. South Africa agrees with the concerns of the ICRC about intended circumstances of use.

Sierra Leone views this as an issue of accountability at all stages but highlighted that it’s difficult to hold those involved in design, such as engineers, responsible for what a machine does. It is more clear cut in the case of those who deploy. Sierra Leone agrees with others who say that the critical functions must have human oversight, and that a ban on the development and deployment is needed. Defensive weapons should not fall under LAWs.

France outlined three aspects of human-machine interaction: military interest, operational efficacy, and predictability and reliability on the battlefield. It emphasised that to be deployable a system must make it possible for humans to maintain responsibility of command and subordination.

Austria underlined that the most important take away of this session is the agreement that humans should maintain meaningful control over LAWs. It described control in targeting and selecting as essential and would like to look further at established

continued on next page



Lethal [autonomous] weapon, continued

lines around this. Like others, Austria noted that machines are not beholden to law, and that fundamental concepts of responsibility and accountability cannot be outsourced to machines. It referred to the Polish statement, particularly around ethics.

Costa Rica believes that the ambiguity present in terms like “sufficient” “minimum” or “appropriate” human control is purposeful. It has also observed that many states agree that LAWS should not execute critical functions—which is something yet to be defined. The future definition of LAWS should include the human element in a multi-dimensional way, that accounts for human capacity for compassion, intelligence, and ethics, among other qualities. Human dignity should rest at the centre of this discussion, Costa Rica said.

Estonia is of the view that meaningful human control does not reflect a new norm of international humanitarian law, but rather that a requirement of human control exists already in IHL. It reinforced that as individuals are the actors accountable to law, and weapons are the tools that humans use in conflict, it is each person’s responsibility to use these instruments consistently with the law. It feels that human control can occur at different touchpoints, ranging from targeting decisions to procurement, or programming. Estonia also expressed concern about the cyber security of autonomous weapon systems.

Israel prefers human analysis to be present in two stages: development and acquisition, and operational planning and deployment. During these phases, humans will determine how LAWS operate. Those involved should also account for operational scenarios, and ensure compliance with IHL. It envisions that LAWS could only be authorised for use in certain scenarios.

Panama said it is vital that meaningful human control be exerted at all stages and mostly in those of critical functions. It is also concerned with the reliability and functioning of LAWS and if they would be able to operate in compliance with international law or distinguish between dangerous and inoffensive individuals or contexts. It points out that those in favour of LAWS argue that they will always be subject to human control, but that is not compatible with military logic. In comparison to conventional weapons, the human factor may be removed because it’s the slowest link in the chain and LAWS will never be able to value the intent of human action or gauge necessity/proportionality. LAWS don’t feel compassion or empathy.

Argentina noted it would be impossible to create systems that are completely immune from cyber attacks. It stressed the factor of unpredictability as

a real risk and an argument for human control. It would not like to rule out the possibility of developing such technologies for humanitarian purposes such as demining, but not for offensive military action.

Bulgaria stated that human-machine interaction should be built up at every stage, starting from research and development. The targeting cycle must include a person to make decisions about taking life, and lethal force must be authorised by an official in the chain of command. It suggests focusing broadly on human control in order to move forward at the GGE, but explained that human control should be exercised in terms of IHL and IHRL.

Sweden reiterated its position from earlier in the week that LAWS do not yet exist. Also focusing on process, Sweden noted the multiplicity of terms being used in relation to human control and that these are all widely supported, so a line needs to be drawn. Ensuring predictability and control is highly important, and Sweden found the New Zealand statement about asking for more discussion about “in and on the loop” useful. Sweden also noted that there are some weapons that have been lawfully in existence for decades, like anti-ship missiles, which might suddenly become illegal based on how LAWS are defined and the political outcomes of this process. It elaborated on points made Tuesday, regarding different scenarios such as maritime ones where the context is quite different than conflict on land or in an urban setting.

The ICRC spoke about legality, ethics, and human supervision. It advocates for a human-centred approach with a minimum level of human control over weapons that have autonomy in their critical functions, and ethical acceptability with their use under the Martens Clause. It further noted that operational constraints (e.g. stopping incoming projectiles or searching for targets over a wide area, or attacks in a populated area versus over the sea) are all factors to consider. It referred states to its paper from the November 2017 meeting and Working Paper No.5 published for this meeting.

IPRAW recommends that LAWS have a two-step design: situational understanding and the ability to stop an attack. Further details are outlined in its new report.

Article 36 noted that some states have argued human control should manifest at various stages of the weapon’s life cycle, in particularly during the targeting process. It argued that we must not allow movement toward having machines undermine the legal, including IHL, requirements demanded of humans.



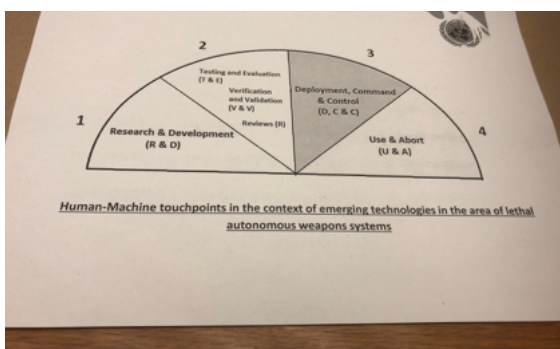
Lethal [autonomous] weapon, continued

Human Rights Watch stated that meaningful human control is the best standard to use, and helps to promote compliance with international law and ethical obligations. Like others, HRW spoke to legal accountability particularly around unlawful harm.

Protection said that human control over the use of LAWS is necessary for weapons whether they are lethal or not. States need to discuss the types and natural limits of human control in the selection and engagement of targets.

The ICRC observed that we cannot talk about the meaningful human control of LAWS because then they would no longer be LAWS. For human control to be meaningful, states should examine the ways in which humans interact with machines, drawing from the psychology of human learning. There should be a focus on what the human operator must do in the targeting process under IHL and IHRL and in post-targeting, and the design of autonomous weapons systems must render them incapable of operating without sufficient human control, among other recommendations.

In the second half of the day, the chair introduced visual resources to illustrate the nature of discussion on this thematic area to date. The first features a half pie chart with four categories of human-machine touch points: 1) research and development; 2) testing, verification, reviews; 3) deployment, command, control; and, 4) use and abort.



Some terms			
(Maintaining)	(Substantive)	Human	(Participation)
(Ensuring)	(Meaningful)		(Involvement)
(Exerting)	(Appropriate)		(Responsibility)
	(Sufficient)		(Supervision)
	(Minimum level of)		(Validation)
			(Control)
			(Judgement)

A second resource categorised vocabulary, which has been used frequently over the last three days. For example, one column contains words such as "maintaining" "ensuring" "exerting" "preserving" while another has "substantive" "meaningful" "appropriate".

A third resource lists "related dimensions", or concerns, that require consideration such as about the purpose of our deliberations, and feasibility.

These diagrams facilitated another free and interactive discussion throughout the afternoon that was also aided by the presentation from the US about its Counter Rocket, Artillery, and Mortar (C-RAM) programme. Themes that emerged are the challenges of trying to agree a "one size fits all" definition or understanding of autonomous weapon systems given different contexts in which these could be deployed or used, including in situations with no risk of civilian harm. Some felt that the different "slices" of the pie chart are not necessarily distinct but have overlap or interact with one another. Another state advised against linking because it might dilute accountability. It was suggested that a zero level of autonomy be added to the half pie chart. Caution was expressed against going into excessive levels of detail and to keep sight of the overarching goals of the GGE.

The US presentation was welcomed for its practicality, although some delegates noted that the nature of the system itself may be beyond what this GGE is concerned with. It generated several questions around reliability as well as specific questions about the nature of the technology and system, the potential of aborting missions, and the role of commanders. The Republic of Korea took the opportunity to describe the de-mining robots it is developing for use in the demilitarised zone. These may end up being a lot like a LAWS as it also detects, identifies, and engages targets and to that end, feels we shouldn't stigmatise technology itself. Toward the end of this session, some states began to question the utility of too detailed discussion and voiced concern that it may become overly complicated or stall progress. •

SIDE EVENT REPORT: MEANINGFUL HUMAN CONTROL OF WEAPON SYSTEMS*Lenka Filipova | Women's International League for Peace and Freedom*

On Wednesday 11 April the Campaign to Stop Killer Robots organised a side event focusing on meaningful human control of autonomous weapons systems. The event followed up on the plenary discussions focusing on further consideration of the human element in the use of lethal force in autonomous weapons systems and aspects of human-machine interaction. The very well-attended side-event was moderated by Ms. Rasha Abdul-Rahim of Amnesty International.

The first presenter was Mr. Paul Scharre from the Center for a New American Security. He asked two overarching questions: "If we had all the technology we could imagine, what role would we want humans to play in war?" and "What decisions in war require uniquely human judgment and why is that?" He spoke about the ethics of autonomous weapons and the lead provided by international humanitarian law (IHL) in answering these questions. One important point is that the IHL treats humans differently because we are the legal agents, and not machines. Unlike machines, humans are obliged to comply with the law and this responsibility cannot be delegated to machines—IHL applies to people.

Ms. Bonnie Docherty, is a senior researcher in the Arms Division at Human Rights Watch and an associate director of Armed Conflict and Civilian Protection at the International Human Rights Clinic at Harvard University. Ms. Docherty provided a brief analysis of what has been said during this week of CCW; she examined how the requirement for meaningful human control can address many of the legal and ethical questions around autonomous weapons systems, as well as outline the legal precedent. She explained that, during this week are hearing different approaches to addressing concerns regarding LAWS.

On one side there are states that now include Austria and the African Group calling for a pre-emptive ban, while on the other side a French-German proposal for a non-binding political declaration was also submitted and is receiving some support. Ms. Docherty pointed out that nevertheless, states have agreed on at least one essential point: that it is necessary to maintain human control over the use of force. She highlighted that at least 40 states have brought up the concept of human control, including the International Committee of the Red Cross (ICRC), and numerous civil society organisations. She stressed that the concept of meaningful human control is not new to disarmament and other areas of international law. Such a requirement to maintain human control over the use of weapons would eliminate many problems, including the potential affront to human dignity, and promote the protection of human rights, particularly the right to life.

Professor Noel Sharkey, from Sheffield University and the chair of the International Committee for Robot Arms Control, introduced different levels of human control of weapons, and described the psychology of human reasoning and decision-making. Dr. Sharkey supported his presentation through various exercises with the audience, introducing the so-called "deliberative process" and "automatic reasoning" in order to help GGE participants better understand cognition and what human control brings to decision-making. While automatic reasoning does not require attentional and memory resources, as introduced on series of colour words, the deliberative process takes time, consuming both attentional and memory resources.

The panel was followed by a fruitful discussion with the audience, and several challenging questions including about if whether it is possible to regulate these weapons without the necessity of regulating non-weapons. It also explored normal

accident theory, and loitering weapons. Another question examined the analogy between LAWS and dogs, what is the difference between autonomous weapons systems and trained dogs, which in its essence are also semi-autonomous weapons without human control. Why we do not have regulations on such animals? •

