

Group of Governmental Experts of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects

31 August 2018

English only

Geneva, 9–13 April 2018 (first week) and 27-31 August 2018 (second week)

Item 7 of the provisional agenda

Adoption of the report

Report of the 2018 Group of Governmental Experts on Lethal Autonomous Weapons Systems

I. Introduction

1. The 2017 Meeting of High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or Have Indiscriminate Effects (CCW), held in Geneva from 22 to 24 November 2017, decided, as contained in its final document (CCW/MSP/2017/8):

“The Group of Governmental Experts related to emerging technologies in the area of lethal autonomous weapons systems (LAWS) in the context of the objectives and purposes of the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons shall meet for a duration of ten days in 2018 in Geneva in accordance with Decision 1 of the Fifth Review Conference of the High Contracting Parties to the Convention (CCW/CONF.V/10), consistent with CCW/CONF.V/2.

The Rules of Procedure of the Review Conference shall apply *mutatis mutandis* to the Group. The Group shall conduct its work and adopt its report by consensus which shall be submitted to the 2018 Meeting of the High Contracting Parties to the Convention. The widest possible participation of all High Contracting Parties is to be promoted in accordance with the goals of the CCW Sponsorship Programme.

The Group will continue to be chaired by Ambassador Amandeep Singh Gill of India without prejudice to the principle of geographical rotation.”

2. The Group of Governmental Experts met from 9 to 13 April and from 27 to 31 August 2018.

II. Organization and work of the Group of Governmental Experts (April session)

3. The following High Contracting Parties to the Convention participated in the work of the Group: Albania, Algeria, Argentina, Australia, Austria, Belarus, Belgium, Bosnia and Herzegovina, Brazil, Bulgaria, Canada, Chile, China, Colombia, Costa Rica, Croatia, Cuba, Czechia, Denmark, Djibouti, Dominican Republic, El Salvador, Estonia, Finland, France, Germany, Greece, Guatemala, Holy See, Honduras, Hungary, India, Iraq, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Kuwait, Lao People’s Democratic Republic, Latvia, Lebanon, Lithuania, Luxembourg, Mexico, Montenegro, Morocco, Netherlands, New Zealand, Nicaragua, Norway, Pakistan, Panama, Peru, Philippines, Poland, Portugal, Qatar, Republic of Korea, Romania, Russian Federation, Saudi Arabia, Senegal, Serbia, Slovakia, South Africa, Spain, Sri Lanka, State of Palestine, Sweden, Switzerland, The Former

Yugoslav Republic of Macedonia, Togo, Turkey, Uganda, United Kingdom of Great Britain and Northern Ireland, United States of America, Venezuela (Bolivarian Republic of).

4. The following Signatory State to the Convention participated in the work of the Group: Egypt.

5. The following States not party to the Convention participated as observers: Ghana, Myanmar, Oman and Zimbabwe.

6. The representatives of the United Nations Institute for Disarmament Research (UNIDIR), the African Union in Geneva, the European Union and the International Committee of the Red Cross (ICRC) participated in the work of the Group in accordance with the rules of procedure.

7. The representatives of the following non-governmental organizations participated in the work of the Group in accordance with the rules of procedure: Campaign to Stop Killer Robots, Amnesty International, Article 36, Association for Aid and Relief Japan, Committee of 100 in Finland, Facing Finance, Human Rights Watch, International Committee for Robot Arms Control (ICRAC), Mines Action Canada, Norwegian Peace Foundation, PAX, Pax Christi Ireland, Project Ploughshares, Protection, Pugwash Conferences on Science and World Affairs, Seguridad Humana en Latinoamérica y el Caribe (SEHLAC), Women's International League for Peace and Freedom (WILPF), and the Centre for a New American Security (CNAS).

8. The representatives of the following entities also participated in the work of the Group in accordance with the rules of procedure: Ada-AI, Birmingham City University, Centre for Emerging Technology Intelligence, Graduate Institute Geneva, Harvard Law School, ICT4Peace Foundation, International Panel on the Regulation of Autonomous Weapons (iPRAW), King's College London, National University of Ireland (Galway), Stockholm International Peace Research Institute (SIPRI), University of Barcelona, University of China, University of Oxford, University Pablo de Olavide (Seville), University of Tasmania, University of Zurich and the Vrije Universiteit Amsterdam.

9. On Monday, 9 April 2018, the first session was opened by the Chairperson, Ambassador Amandeep Singh Gill of India. Ms. Anja Kaspersen, Director of the Geneva Branch of the UN Office for Disarmament Affairs, addressed the session on behalf of Ms. Izumi Nakamitsu, the High Representative for Disarmament Affairs.

10. At its first plenary meeting, the Group adopted its Agenda (CCW/GGE.1/2018/1), confirmed the Rules of Procedure as adopted by the Fifth Review Conference (CCW/CONF.V/4), and adopted its Programme of Work (CCW/GGE.1/2018/2). Mr. Marco Kalbusch, Senior Political Affairs Officer, UN Office for Disarmament Affairs, served as the Secretary of the Group, and was assisted by Ms. Mélanie Gerber, Associate Political Affairs Officer, UN Office for Disarmament Affairs. Ms. Isabelle Porcu-Cartier, Staff Assistant and Ms. Nadiya Dzyubynska, Staff Assistant served in the Secretariat.

11. At the same plenary meeting, the following delegations participated in a general exchange of views: Argentina, Australia, Austria, Belgium, Brazil, Bulgaria, Canada, Chile, China, Costa Rica, Cuba, Egypt, Estonia, Finland, France and Germany (joint statement), Greece, Holy See, India, Ireland, Israel, Italy, Japan, Latvia, Luxembourg, Mexico, Netherlands, Pakistan, State of Palestine, Panama, Republic of Korea, Russian Federation, South Africa on behalf of the African Group, Spain, Sri Lanka, Sweden, Switzerland, Turkey, United Kingdom of Great Britain and Northern Ireland, United States of America, European Union, Venezuela (Bolivarian Republic of) on behalf of the Non-Aligned Movement and other High Contracting Parties to the CCW, UNIDIR, ICRC, Campaign to Stop Killer Robots, International Committee for Robot Arms Control, Human Rights Watch, Mines Action Canada and PAX.

12. In accordance with its programme of work (CCW/GGE.1/2018/2), the Group commenced with a general exchange. The Group then considered the following agenda items:

(a) Characterization of the systems under consideration in order to promote a common understanding on concepts and characteristics relevant to the objectives and purposes of the Convention;

(b) 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;

(c) Review of potential military applications of related technologies in the context of the Group's work;

(d) Possible options for addressing the humanitarian and international security challenges posed by emerging technologies in the area of lethal autonomous weapons systems in the context of the objectives and purposes of the Convention without prejudging policy outcomes and taking into account past, present and future proposals.

13. The Group heard presentations from the following experts under agenda item 6 c: Expert Panel on the review of potential military applications of related technologies in the context of the Group's work: Professor Pascale Fung, Hong Kong University of Science and Technology; Professor Anthony Gillespie, University College London; Professor Mary Cummings, Duke University; Professor Dominique Lambert, Université de Namur and Professor Viacheslav Pshikhopov, Southern Federal University.

14. The Group considered the documents listed in Annex I. The Group noted with appreciation the contributions of those High Contracting Parties submitting Working Papers, presenting their national policies and positions, and the input of civil society, including industry.

III. Organization and work of the Group of Governmental Experts (August session)

15. The following High Contracting Parties to the Convention participated in the work of the Group: Albania, Algeria, Argentina, Australia, Austria, Belarus, Belgium, Benin, Bosnia and Herzegovina, Brazil, Bulgaria, Canada, Chile, China, Colombia, Costa Rica, Croatia, Cuba, Cyprus, Czechia, Denmark, Djibouti, Dominican Republic, Ecuador, El Salvador, Estonia, Finland, France, Germany, Greece, Guatemala, Holy See, Honduras, Hungary, India, Iraq, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Kuwait, Latvia, Lebanon, Lithuania, Luxembourg, Mexico, Montenegro, Morocco, Netherlands, New Zealand, Nicaragua, Norway, Pakistan, Panama, Peru, Philippines, Poland, Portugal, Qatar, Republic of Korea, Romania, Russian Federation, Saudi Arabia, Senegal, Serbia, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, State of Palestine, Sweden, Switzerland, The Former Yugoslav Republic of Macedonia, Togo, Turkey, Uganda, United Kingdom of Great Britain and Northern Ireland, United States of America, Venezuela (Bolivarian Republic of).

16. The following Signatory State to the Convention participated in the work of the Group: Egypt.

17. The following States not party to the Convention participated as observers: Islamic Republic of Iran, Myanmar, Oman and Zimbabwe.

18. The representatives of the United Nations Institute for Disarmament Research (UNIDIR), the United Nations Mine Action Service (UNMAS), the European Union and the International Committee of the Red Cross (ICRC) participated in the work of the Group in accordance with the rules of procedure.

19. The representatives of the following non-governmental organizations participated in the work of the Group in accordance with the rules of procedure: Campaign to Stop Killer Robots, Amnesty International, Article 36, Association for Aid and Relief Japan, Committee of 100 in Finland, Center for International Security and Policy, Facing Finance, Future of Life Institute Human Rights Watch, ICT4Peace Foundation, International Committee for Robot Arms Control, Mines Action Canada, Nobel Women's Initiative, Norwegian Peace Foundation, PAX, Pax Christi Vlaanderen, Project Ploughshares, Protection, Rete Italiana

per il Disarmo, SEHLAC, Women's International League for Peace and Freedom (WILPF), Conscious Coders, International Action Network on Small Arms, Pax Christi International and Zonta International.

20. The representatives of the following entities also participated in the work of the Group in accordance with the rules of procedure: International Panel on the Regulation of Autonomous Weapons (iPRAW), University Lyon, Universidad Pablo de Olavide, University of North Carolina, University of Tampere, Quadriga University, University of Kent, Birmingham City University, University of New South Wales at the Australian Defence Force Academy, University of Tasmania, Institute for European Studies, Harvard Law School, National University of Ireland (Galway), University of Oxford, CAN Center for Autonomy and AI, Geneva Centre for Security Policy (GCSP) and RAND Corporation.

21. On Monday, 27 August 2018, the first session was opened by the Chairperson, Ambassador Amandeep Singh Gill of India. Ms. Anja Kaspersen, Director of the Geneva Branch of the of the UN Office for Disarmament Affairs, addressed the session on behalf of Ms. Izumi Nakamitsu, the High Representative for Disarmament Affairs. Ms. Mélanie Gerber, Political Affairs Officer, UN Office for Disarmament Affairs, served as the Secretary of the Group, and was assisted by Ms. Amy Dowler, Political Affairs Officer, UN Office for Disarmament Affairs. Ms. Maria José Orellana Alfaro, Ms. María-Luisa Zeballos and Ms. Nadiya Dzyubynska, Staff Assistants, served in the Secretariat.

22. In accordance with its programme of work (CCW/GGE.1/2018/2), the Group commenced with a panel discussion on agenda item 6 c. The Group considered the following agenda items:

(a) Characterization of the systems under consideration in order to promote a common understanding on concepts and characteristics relevant to the objectives and purposes of the Convention;

(b) 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 system;

(c) Review of potential military applications of related technologies in the context of the Group's work;

(d) Possible options for addressing the humanitarian and international security challenges posed by emerging technologies in the area of lethal autonomous weapons systems in the context of the objectives and purposes of the Convention without prejudging policy outcomes and taking into account past, present and future proposals.

The Group then considered and adopted the final report.

23. The Group heard presentations from the following experts under agenda item 6 c: Dr. Lydia Kostopoulos, Digital Society Institute of the European School of Management and Technology; Lieutenant Colonel Christopher Korpela, Robotics Research Center of the United States Military Academy at West Point; Professor Anthony Gillespie, University College London; Dr. Gautam Shroff, Tata Consultancy Services and Dr. Knut Dörmann, International Committee of the Red Cross (ICRC).

24. The Group considered the documents listed in Annex II. The Group noted with appreciation the contributions of those High Contracting Parties submitting Working Papers, presenting their national policies and positions, and the input of civil society, including industry.

25. A summary of the discussions held during the April and August meetings of the Group, prepared under the Chairperson's responsibility, is attached to this report as Annex III.

IV. Emerging Commonalities, Conclusions and Recommendations

A. Possible Guiding Principles

26. It was affirmed that international law, in particular the United Nations Charter and international humanitarian law (IHL) as well as relevant ethical perspectives, should guide the continued work of the Group. Noting the potential challenges posed by emerging technologies in the area of lethal autonomous weapons systems to IHL,¹ the following were affirmed, without prejudice to the result of future discussions:

(a) International humanitarian law continues to apply fully to all weapons systems, including the potential development and use of lethal autonomous weapons systems.

(b) Human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines. This should be considered across the entire life cycle of the weapons system.

(c) Accountability for developing, deploying and using any emerging weapons system in the framework of the Convention on Certain Conventional Weapons (CCW) must be ensured in accordance with applicable international law, including through the operation of such systems within a responsible chain of human command and control.

(d) In accordance with States' obligations under international law, in the study, development, acquisition, or adoption of a new weapon, means or method of warfare, determination must be made whether its employment would, in some or all circumstances, be prohibited by international law.

(e) When developing or acquiring new weapons systems based on emerging technologies in the area of lethal autonomous weapons systems, physical security, appropriate non-physical safeguards (including cyber-security against hacking or data spoofing), the risk of acquisition by terrorist groups and the risk of proliferation should be considered.

(f) Risk assessments and mitigation measures should be part of the design, development, testing and deployment cycle of emerging technologies in any weapons systems.

(g) Consideration should be given to the use of emerging technologies in the area of lethal autonomous weapons systems in upholding compliance with IHL and other applicable international legal obligations.

(h) In crafting potential policy measures, emerging technologies in the area of lethal autonomous weapons systems should not be anthropomorphized.

(i) Discussions and any potential policy measures taken within the context of the CCW should not hamper progress in or access to peaceful uses of intelligent autonomous technologies.

(j) CCW offers an appropriate framework for dealing with the issue of emerging technologies in the area of lethal autonomous weapons systems within the context of the objectives and purposes of the Convention, which seeks to strike a balance between military necessity and humanitarian considerations.

¹ Annex III, Chair's summary (pp. 22-23)

B. Characterization of the systems under consideration in order to promote a common understanding on concepts and characteristics relevant to the objectives and purposes of the Convention

27. Having examined different conceptual approaches to characterization and considered sets of specific characteristics relevant to the objectives and purposes of the Convention, and without prejudice to any future understanding on characterization, it was noted that:

(a) For some delegations, a working definition of lethal autonomous weapons systems is essential to fully address the potential risks posed. For others, absence of an agreement on a definition should not hamper discussions or progress within the CCW. Characterization, or working definitions, should neither predetermine nor prejudice policy choices; they should be universally understood by stakeholders.

(b) Purely technical characteristics such as physical performance, endurance or sophistication in targeting acquisition and engagement may alone not be sufficient to characterize lethal autonomous weapons systems, especially in view of rapid evolution in technology.

(c) Technical characteristics related to self-learning (without externally-fed training data) and self-evolution (without human design inputs) have to be further studied. Similarly, attempting to define a general threshold level of autonomy based on technical criteria alone could pose difficulty as autonomy is a spectrum, its understanding changes with shifts in the technology frontier, and different functions of a weapons system could have different degrees of autonomy.

(d) Lethality as made explicit in the mandate of the GGE does not prejudice the application of and respect for all rules relevant to the conduct of hostilities.

(e) Autonomy in the military targeting and engagement cycle has to be studied further keeping in view that autonomy can exist throughout or during parts of the targeting cycle and could start to be applied increasingly in other contexts such as close combat.

(f) In the context of the CCW, a focus on characteristics related to the human element in the use of force and its interface with machines is necessary in addressing accountability and responsibility.

C. 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

28. In the context of the objectives and purposes of the CCW, it was noted that the nature and quality of the human-machine interface is important to address concerns related to the development, deployment and use of emerging technologies in the area of lethal autonomous weapons systems. In line with the Chair's 'sunrise slide', the following touch points in the human-machine interface were considered: 0) political direction in the pre-development phase; 1) research and development; 2) testing, evaluation and certification; 3) deployment, training, command and control; 4) use and abort; 5) post-use assessment. It was noted that:

(a) Accountability threads together these various human-machine touch points in the context of the CCW. Humans must at all times remain accountable in accordance with applicable international law for decisions on the use of force.

(b) Where feasible and appropriate, inter-disciplinary perspectives must be integrated in research and development, including through independent ethics reviews bearing in mind national security considerations and restrictions on commercial proprietary information.

(c) Weapons systems under development, or modification which significantly changes the use of existing weapons systems, must be reviewed as applicable to ensure compliance with IHL.

(d) Where feasible and appropriate, verifiability and certification procedures covering all likely or intended use scenarios must be developed, the experience of applying such procedures should be shared bearing in mind national security considerations or commercial restrictions on proprietary information.

(e) Accountability for the use of force in armed conflict must be ensured in accordance with applicable international law, including through the operation of any emerging weapons systems within a responsible chain of command and control.

(f) Human responsibility for the use of force must be retained. To the extent possible or feasible, this could extend to intervention in the operation of a weapon if necessary to ensure compliance with IHL.

(g) Necessary investments in human resources and training should be made in order to comply with IHL and retain human accountability and responsibility throughout the development and deployment cycle of emerging technologies.

(h) Keeping in mind the foregoing, and recognizing the authority and responsibility of States in this area, it would be useful to continue discussions on reaching shared understandings on the extent and quality of the human-machine interaction in the various phases of the weapons system's life cycle as well as clarifying the accountability threads throughout these phases.

D. Review of potential military applications of related technologies in the context of the Group's work

29. The valuable contribution of experts from the tech community, industry, academia and civil society to building awareness and understanding of the potential military applications of emerging technologies in the area of lethal autonomous weapons systems in the context of the Group's work was recognized. These inputs have been channeled mainly through experts participating in national delegations, panels put together at the invitation of the Chair, side events and open calls for contributions on the CCW website. They have ensured that the Group's policy consideration advances in step with developments in the technology field and a minimum degree of transparency regarding potential military applications is built up.

30. The value of initiatives by industry, the science and technology community, academia and other organizations to develop a common scientific and policy vernacular across the globe was also recognized.

31. Moving forward, ways and means to preserve this momentum and cross-fertilization of knowledge through dialogue in the context of the CCW need to be found. Enhanced participation of cross-disciplinary experts, with due regard for gender balance, in delegations attending CCW meetings should be encouraged to ensure that the Convention's consideration of the issue stays in step with the advance of technology.

E. Possible options for addressing the humanitarian and international security challenges posed by emerging technologies in the area of lethal autonomous weapons systems in the context of the objectives and purposes of the Convention

32. In the context of the CCW, delegations raised a diversity of views on potential risks and challenges posed by emerging technologies in the area of lethal autonomous weapons systems including in relation to harm to civilians and combatants in armed conflict in contravention of IHL obligations, exacerbation of regional and international security dilemmas through arms races and the lowering of the threshold for the use of force. Proliferation, acquisition and use by terrorists, vulnerability of such systems to hacking and interference, and the possible undermining of confidence in the civilian uses of related technologies were also raised.

33. Delegations presented different options to address these potential risks and challenges in the context of the objectives and purposes of the CCW. Their pros and cons were discussed under four categories, including a legally-binding instrument, a political declaration, and clarity on the implementation of existing obligations under international law, in particular IHL.

- Under the first category, a proposal for a legally-binding instrument stipulating prohibitions and regulations on lethal autonomous weapons systems was made. A mandate to negotiate a legally-binding instrument to ensure human control over the critical functions in lethal autonomous weapons systems was proposed.
- Under the second category, a proposal for a political declaration that would outline important principles such as the necessity of human control in the use force and the importance of human accountability, and with elements of transparency and technology review, was made.
- Under the third category, proposals were made to further discuss the human-machine interface and the application of existing international legal obligations. The need to identify practical measures, best practices and information sharing for improving compliance with international law, including legal weapons reviews required by Article 36 of the Additional Protocol I to the Geneva Conventions, was also underlined.
- As IHL is fully applicable to potential lethal autonomous weapons systems a view was also expressed that no further legal measures were needed.

34. It was felt that the options were not necessarily mutually exclusive, and the work carried out so far in the GGE on principles, characterization, human-machine interface and review of potential military applications of emerging technologies in the area of lethal autonomous weapons systems and related emerging commonalities offered useful building blocks for future work. Existing understandings need to be consolidated, open questions clarified and further common ground built on the basis of consensus.

35. The Group emphasized that the CCW offers an appropriate framework for dealing with the issue of emerging technologies in the area of lethal autonomous weapons systems. Within the context of the broader policy work internationally necessitated by the combinatorial effects of emerging technologies in the area of lethal autonomous weapons systems, the Convention's modular and evolutionary character, the balance it seeks to strike between humanitarian considerations and military necessity as well as the opportunity it offers to engage multiple stakeholders make it an ideal platform for focused and participative discussions for reaching common understandings on the subject.

Recommendation

36. Some delegations made proposals for strengthening the mandate of the GGE going forward. These proposals are listed in Annex III. Some delegations called for retaining the existing mandate; others underlined that the current mandate offered sufficient flexibility and scope for stepping up work while continuing to explore options for an outcome. A suggestion was made for rationalizing the number of days for the GGE (as reflected in Annex III).

37. In the light of the above discussions, the Group recommends that:

The Group of Governmental Experts related to emerging technologies in the area of lethal autonomous weapons systems (LAWS) in the context of the objectives and purposes of the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons shall meet for a duration of ten days in 2019 in Geneva in accordance with Decision 1 of the Fifth Review Conference of the High Contracting Parties to the Convention (CCW/CONF.V/10), consistent with CCW/CONF.V/2.

The Rules of Procedure of the Review Conference shall apply *mutatis mutandis* to the Group. The Group shall conduct its work and adopt its report by consensus which shall be submitted to the 2019 Meeting of the High Contracting Parties to the Convention. The widest possible participation of all High Contracting Parties is to be promoted in accordance with the goals of the CCW Sponsorship Programme.

ADVANCE COPY

Annex I

List of documents

CCW/GGE.1/2018/1	Provisional agenda. Submitted by the Chairperson
CCW/GGE.1/2018/2	Provisional Programme of Work. Submitted by the Chairperson
CCW/GGE.1/2018/WP.1	General Principles on Lethal Autonomous Weapons Systems. Submitted by the Bolivarian Republic of Venezuela on behalf of the Non-Aligned Movement (NAM) and Other States Parties to the Convention on Certain Conventional Weapons
CCW/GGE.1/2018/WP.2	Strengthening of the Review Mechanisms of a New Weapon, Means or Methods of Warfare. Submitted by Argentina
CCW/GGE.1/2018/WP.3	Working Paper on Lethal Autonomous Weapons Systems. Submitted by Poland
CCW/GGE.1/2018/WP.4	Humanitarian Benefits of Emerging Technologies in the Area of Lethal Autonomous Weapons Systems. Submitted by the United States of America
CCW/GGE.1/2018/WP.5	Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control? Submitted by the International Committee of the Red Cross (ICRC)
CCW/GGE.1/2018/WP.6	Russia's Approaches to the Elaboration of a Working Definition and Basic Functions of Lethal Autonomous Weapons Systems in the Context of the Purposes and Objectives of the Convention. Submitted by the Russian Federation
CCW/GGE.1/2018/WP.7	Position Paper. Submitted by China

Annex II

List of documents

CCW/GGE.1/2018/1	Provisional agenda. Submitted by the Chairperson
CCW/GGE.1/2018/2	Provisional Programme of Work. Submitted by the Chairperson
CCW/GGE.1/2018/INF.1	List of Participants
CCW/GGE.2/2018/WP.1	Human Machine Touchpoints: The United Kingdom's perspective on human control over weapon development and targeting cycles. Submitted by the United Kingdom of Great Britain and Northern Ireland
CCW/GGE.2/2018/WP.2	Categorizing lethal autonomous weapons systems - A technical and legal perspective to understanding LAWS. Submitted by Estonia and Finland
CCW/GGE.2/2018/WP.3	Human-machine interaction in the development, deployment and use of emerging technologies in the area of lethal autonomous weapons systems. Submitted by France
CCW/GGE.2/2018/WP.4	Human-Machine Interaction in the Development, Deployment, and Use of Emerging Technologies in the Area of Lethal Autonomous Weapons Systems. Submitted by the United States of America
CCW/GGE.2/2018/WP.5	Areas of Convergence on LAWS. Submitted by Brazil
CCW/GGE.2/2018/WP.6	The Australian Article 36 Review Process. Submitted by Australia
CCW/GGE.2/2018/WP.7	Proposal for a Mandate to Negotiate a Legally-binding Instrument that Addresses the Legal, Humanitarian and Ethical Concerns Posed by Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (LAWS). Submitted by Austria, Brazil and Chile

Annex III

Chair's summary of the discussion of the 2018 Group of Governmental Experts on emerging technologies in the area of lethal autonomous weapons systems

Agenda item 6 (a)

Characterization of the systems under consideration in order to promote a common understanding on concepts and characteristics relevant to the objectives and purposes of the Convention

Delegations exchanged views on different approaches to the characterization of lethal autonomous weapons systems (LAWS) based on emerging technologies related to intelligent autonomous systems in order to promote a common understanding on concepts and characteristics relevant to the objectives and purposes of the Convention on Certain Conventional Weapons (CCW). Four broad approaches to characterization were highlighted and served as a reference point during the discussions without prejudice to other possible approaches.

Separative approach

An approach whereby characteristics and concepts not relevant to the objectives and purposes of the CCW are set aside (“via negativa”), while gathering the characteristics and concepts that are definitely relevant to the objectives and purposes of the CCW (“via positiva”).

Cumulative approach

An approach whereby categories of characteristics are added to a master list and then concepts and characteristics therein are evaluated against certain technical, legal-humanitarian or political-security criteria to assess their relevance to the objectives and purposes of the CCW. Such categories could include physical performance, targeting performance and other technical characteristics. They could also include characteristics that are related to the human-machine interface, the human-machine relationship or secondary characteristics such as reliability, predictability and, subordination to command and control.

Accountability approach

An approach, which considers a set of characteristics related to the functions and type of decisions handed over to machines, and which avoids using levels of autonomy and other technical characteristics or categories related to the loss of human control. This approach would depend rather on the context and scenario in which the systems under consideration would be used and would involve a combination of technical and human-interface evaluations centred on accountability of States and humans.

Purpose oriented and effect-based approach

This approach focuses on desirable and undesirable consequences of possible lethal weapons systems based on emerging autonomous intelligent systems and technologies.

Delegations listed comprehensive sets of characteristics that were of interest to the deliberations while reaffirming the applicability of International Humanitarian Law (IHL) to all new weapons developed for use in armed conflict, including those with autonomous functions, and the responsibility of States to ensure such compliance. Some delegations noted that the key obstacle to further work was the absence of working samples as well as a common understanding on a working definition. It was felt, however, that while a definition would be eventually essential the absence of an agreed definition should not prevent the Group from moving forward with the discussions. Others pointed to the lack of common understanding on various issues related to emerging technologies in the area of lethal autonomous weapons systems and encouraged more effort to be invested in education and

deepening the collective understanding. There was a desire to ensure that understandings on characterization stood the test of time and were not overtaken by technological developments. Physical or technical attributes alone would not be sufficient to characterize lethal autonomous weapons systems and the Group sought to focus on considerations related to the human element in the use of force and build understandings on the human-machine interface throughout the lifecycle of weapons systems. In examining approaches to characterization, the Group sought to treat technology characteristics as capabilities and not as human features or stand-alone physical objects.

Some delegations suggested that existing weapons systems were not the subject of the discussions, whereas others expressed the view that systems under consideration were already in existence. Likewise, some delegations felt that lethality was an essential characteristic spelt out in the mandate while others felt that the term “lethal” as a characteristic needed to be further examined in the light of the fundamental notion of use of force, which triggers legal obligations under international law irrespective of lethality. Several delegations expressed the view that a focus on lethality would fail to address injuries to persons or damage to objects that are protected by IHL. The interface of lethal systems with non-lethal decision support or other systems was highlighted as well.

Delegations also discussed important questions related to the ability of the machine for self-learning and self-evolution, which could potentially enable the machine to redefine targets. Some delegations considered distinguishing autonomy from semi-autonomy or automation to be helpful, while others promoted the consideration of autonomy as a broad spectrum, noting that autonomy was not an on/off phenomenon and there is lack of a clear line beyond which human control is lost or autonomy becomes problematic. The possibility to interrupt or abort the operation of a system was also discussed as a characteristic. In discussing technical characteristics, delegations underlined the need to apply a human-centric focus in the discussions to focus on the human element in the design and (ultimate) decision-making chain when choosing targets, authorizing or using (lethal) force.

Delegations generally viewed human control as core concept for advancing discussions on emerging technologies in the area of lethal autonomous weapons systems and narrowing the focus of the Group’s future work. Many delegations were of the view that agreement on each and every characteristic was not essential and the Group could proceed in a step-by-step manner when characterizing the systems under consideration. This was also true for one approach that suggested a distinction between anti-personnel and anti-material systems. Some delegations noted that, in view of the connectivity of the characterization discussion with the human-machine interface discussion, consideration of specific systems and use scenarios could be helpful in building a shared understanding of the problems that might arise for human control and accountability as well as of characteristics related to the human element in the use of force.

Delegations stressed that autonomy can exist throughout or during different parts of the targeting cycle. Accordingly, some semi-autonomous machines can have highly autonomous critical functions while highly autonomous machines can have no or limited autonomy in critical functions. Moreover, different functions of a weapons system may have different levels of autonomy and there may be no single general level of autonomy across the system. Thus, purely technical criteria may not be sufficient in framing a characterization of existing weapons or future weapons and could only serve as a reference point given the fast evolution of emerging technologies. Delegations underlined that a focus on the level, nature and primacy of human control rather than purely technical criteria was also logical given the purpose and scope of the CCW.

- Delegations suggested a variety of attributes and interpretations that could be used in characterizing emerging technologies in the area of lethal autonomous weapons systems. These included:
- A system operating with neither human control after activation nor subordination to the chain of command

- A system capable of understanding higher level intent and direction with the ability to take appropriate action by choosing its course of action without depending on human oversight and control, although these may still be present
- A system capable of carrying out tasks governed by IHL in partial or full replacement of a human in the use of force, notably in the targeting cycle
- A system that once launched or deployed assumes a complex adaptive self-learning mode
- An adaptive system capable of navigating through a complex environment by redefining scenarios and approaches
- A rules-based system able to switch to autonomous mode
- A system that can select and attack targets without human intervention, in other words a system that self-initiates an attack
- Fully-autonomous systems, that is, unmanned technical means, other than ammunition, that are designed to carry out combat and support tasks without any participation of an operator
- A weapon system which can act autonomously in delivering (lethal) effects to a target and may also act autonomously in detection and target selection prior to engagement of the target. The level of autonomy can vary from basic levels of automation through a spectrum of an increasing number of autonomous functions and decreasing human control up to and including fully autonomous systems which operate across a range of functions without direct human control.

Delegations also discussed if the following systems should fall under the mandate of the GGE:

- A system with manual override, self-destruct, or self-deactivate function
- Systems that are technologically advanced but have no autonomy
- Systems with some degree of autonomy like close-in weapons systems, which autonomously engage incoming targets based on clearly defined parameters
- A rules-based system that is subject to a clear chain of command and control
- A learning system that offers options

Some delegations emphasised that autonomy in non-critical functions could be adequately addressed within existing legal frameworks. Others considered it important to cover not only critical functions but also other situations for use of force such as close combat when discussing the characteristics of lethal autonomous weapons systems. Delegations discussed if the method and nature of human control over critical functions were the key consideration in judging the weapons systems' compliance with IHL. Some delegations emphasized that consideration of the intended or expected circumstances of the use of the weapon was an important factor for compliance with IHL.

Delegations emphasised the importance of reaffirming relevant States' obligations for a legal review of current and new weapons systems. The importance of examining and sharing best practices was also underlined, which included but was not limited to: sharing national policies and approaches for assessing and dealing with autonomous technologies; considerations related to the human element in the use of force; intelligibility and explainability; and adoption of new standards, methods and protocols for testing and validation. In the context of the discussion on a purpose-oriented approach, some delegations stated that policy should drive definitions and related characteristics, not the other way around.

Agenda item 6 (b)

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

The Chair presented as a reference for delegations a set of three slides: first, on four broad areas of touch points in the human-machine interface (the ‘sunrise’ slide); second, a list of terms used by different delegations; and third, some qualitative, purpose and practice related considerations with regard to the human-machine interface. The solid line indicates that while national frameworks could cover all areas of the human-machine interface, the extent of international or industry level regulation could be limited.

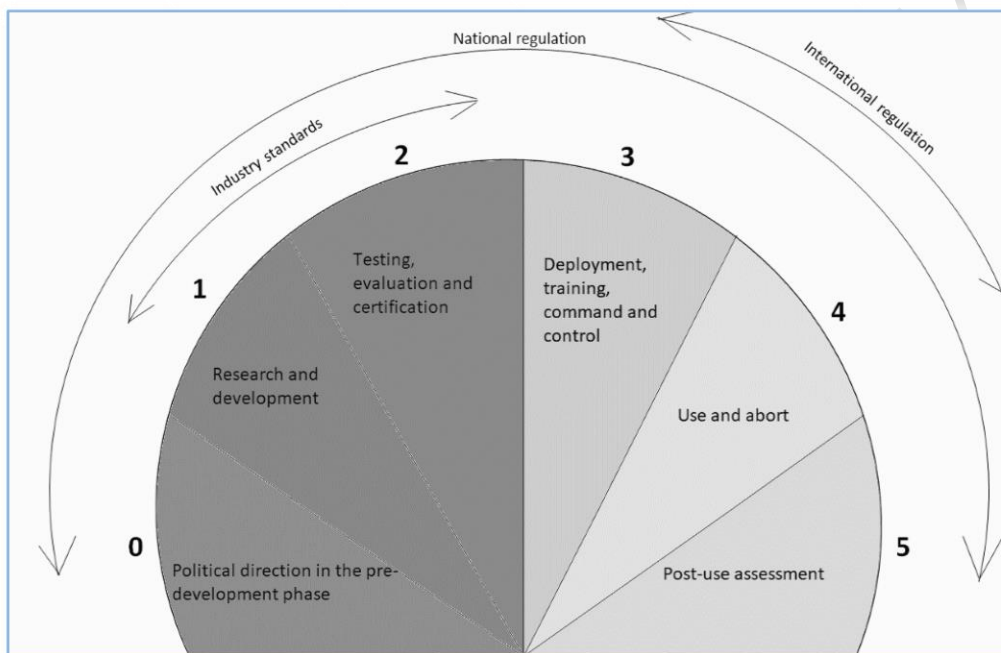


Image 1: Human-machine touchpoints in the context of emerging technologies in the area of lethal autonomous weapons systems, as updated at the August meeting.

Linking back to the discussion on characterization, delegations underlined the need to apply a human-centric focus in discussing technical characteristics of lethal autonomous weapons systems. The nature and quality of the human-machine interface was viewed as critical for addressing concerns related to the development, deployment and use of emerging technologies in the area of lethal autonomous weapons systems. Many delegations supported the notion that maintaining human control over weapons systems, particularly their critical functions, was necessary for compliance with IHL while some underlined difficulties with the notion of human control. Different terminologies and conceptualizations were brought forward by delegations with regard to human control. One was the importance of maintaining human control over the critical functions of autonomous weapons systems. Another was the human element in the different phases of the lifecycle of a weapons system and the level and quality of human control that can be applied at each stage.

Regardless of the approach, delegations reaffirmed the essential importance of human control, supervision, oversight or judgement in the use of force. It was pointed out that while terms such as human control were not expressly foreseen by IHL, their use could be derived from the requirement for compliance with IHL in the application of force.

Using the ‘sunrise’ slide of the various touch points in the human-machine interface, delegations stated that human control, supervision, oversight or judgement could be exerted in varied forms and at different times in the following phases: first, in the research and development phase; second, at the stage of testing and evaluation, verification and validation;

third, the deployment, command and control stage; and fourth, the use and abort stage. Some delegations mentioned additional stages, such as a stage preceding the first one, consisting of national policies and regulations with regard to weapons development or the actual decision to procure weapons, or a stage after the fourth related to evaluation and assessment of use. The addition of training to the third stage was also discussed.

Some delegations noted that there may be no single touch point or notion that can fully describe the role of humans throughout the life cycle of a weapons system, including among others the development, testing, deployment, use and post-use assessment phases instead of focusing on a 'silver bullet' such as the notion of 'human in the loop'. The distribution of human control across the various different touch points does not mean a dilution of responsibility because each touch point is threaded with the notion of accountability. Many delegations thought it would be fruitful for the Group to focus on the qualitative and quantitative dimensions of human control required to ensure compliance with international law and in particular IHL for each of these touch points. Delegations also called for elaboration of the human-machine interface in order to balance the weaknesses of one with the relative strengths of the other as well as to maintain the human as the essential element in the human-machine team, with overall responsibility for coordination and decision-making.

Some delegations noted that human control is already exerted in the research and development phase, which includes weapons reviews. In the research and development phase, teams of software developers will jointly create algorithms, technical experts will design and/or "train" the software and engineers will be responsible for the hardware and the integration of the software. Hardware and software design must allow an operator to actually exercise control during the operation of the system, through specific instruments in the human-machine interface and relevant procedures programmed into the system's processes to enable human input and intervention. Developers could seek to establish in the design phase itself as appropriate and feasible that any autonomous weapons system is receptive to human intervention, judgement and control.

Delegations highlighted the necessity to demonstrate that weapons systems have the capacity to be used in accordance with IHL, in particular the rules on the conduct of hostilities, distinction, proportionality and precautions in attack. Training of armed forces to ensure the ability to use a given system in compliance with international law, specifically IHL, was also important. The importance of legal weapons reviews under Article 36 of Protocol I to the Geneva Conventions was reiterated by delegations.

In the context of deployment and command and control, delegations underlined the need for a system to maintain the double principle of command and subordination. Subordination means that the framing, re-definition and adjustment of a weapons system's mission needs to be done by humans. Concrete decisions related to the 'when and where' of the use of force need to be taken by humans. It was added that communication links, even if intermittent, must be maintained – through a variety of means – between the chain of command and the weapons system, in order to maintain sufficient control and allow humans to take ultimate decisions regarding the use of force.

In context of the deployment and use of a weapons system in an armed conflict, delegations noted that military personnel activate the weapons systems and monitor their functioning. This would require that the operator know the characteristics of the weapons system, is assured that they are appropriate to the environment in which it would be deployed and has sufficient and reliable information on them in order to make conscious decisions and ensure legal compliance. It was also noted that control over use encompasses the procedural requirements to maintain control over the systems during planning, tasking and operation in a two-step approach: the ability to understand the situation and its context, for example when battle space situations change or during malfunctions, and the option to appropriately intervene if necessary by overriding the system or manipulating the machine, either at all steps of the targeting cycle, or at least during the target selection and engagement phases. It was also expressed that machines cannot simply be programmed to comply with IHL, and that therefore positive measures are necessary to prevent indiscriminate action and injury by lethal autonomous weapons systems caused by a breakaway from human control. To develop

such measures, concepts such as ‘meaningful human control’ and ‘human judgment’ need to be further elaborated and clarified.

Delegations mentioned that an autonomous system must not acquire the capacity to repurpose itself automatically and repeatedly, and the human operator must maintain the capacity to intervene as is the case in current civilian applications. Further, human control over a machine must be adapted to the specific sets of tasks and the environment in which a system is operating and must allow the human to make meaningful decisions that comply with IHL and other requirements or to disengage the system if required. Several clarifications on how to describe the extent of human interaction with lethal autonomous weapons systems were suggested, such as substantive, meaningful, appropriate or sufficient human control. Delegations also suggested a minimum level of or minimum indispensable extent of human control.

Delegations stated that post-use, it is necessary to have the possibility to establish accountability. Delegations stressed the importance of ensuring that humans remained at all times accountable for the use of force and related decisions and actions. Ensuring accountability might require actions of a weapons system to be recordable, auditable or explainable. Reliability, predictability and trust were mentioned as other qualitative aspects of control.

The discussions benefited from presentations by delegations of four real-world examples:

- Delegations discussed an existing defensive weapon system with the capability of countering rocket, artillery and mortar threats. The case study served as an example of a weapons system whose operations involve a mix of human decision-making and automation. High levels of automation and some autonomous functions allow the system to operate at the necessary speed, while human operator oversight and inputs assures the safety of the operation and allows for intervention if necessary. The presentation provided information on why and how the weapon system was developed, how it functioned, and whether it has worked as intended. The presentation was made with the intention to increase the understanding in the GGE of how autonomy in various functions can enhance the ability of weapons to serve their purposes while also respecting IHL.
- Delegations also discussed an example of a weapons system under development whose operations could lead to greater compliance with IHL. This system, an underwater autonomous vessel equipped with a sonar, ship registry data, and torpedoes, would be able to recognize and differentiate between civilian and military vessels based on the input from the sonar system and comparison of the input with the onboard ship registry. In case a civilian vessel is detected, the torpedoes would not be launched or would be diverted.
- Delegations received an overview of an anti-tank artillery weapons system for which human control is limited to the selection of the target area. Rather than using high explosives, the system’s warhead contains two metallic slug munitions that must hit their target very precisely to be effective. Each munition is equipped with heat and radar sensors which can scan a 200m diameter area. If a target is detected, the warhead is activated; otherwise it self-destructs. This system, which provides no opportunity for a human to control or abort following the release of the sub-munitions, has been operational for fifteen years and passed legal weapon reviews in two States. The presenter argued that the autonomy-enabled precision of this system allowed it to have a smaller footprint than a traditional explosive warhead.
- Delegations also heard about a system that employs autonomy in naval mine countermeasures. The system utilizes autonomy to reduce or eliminate the most time-consuming steps in traditional mine counter-measures processes. Its primary component is an unmanned surface vessel (USV) that carries mine-hunting unmanned underwater vehicles (UUVs) and neutralizing systems. The system uses automated algorithms enabling the in-situ planning, scheduling, and deconfliction of UUV missions. The UUVs are able to detect and identify targets autonomously and send data back to the USV, which launches neutralizers on that basis. The neutralizer

proceeds autonomously to the mine location, validates the target and autonomously renders the mine inert. The system communicates with an operator on a mothership who is able to abort the mission.

During discussions on agenda item 6 (b) the following terms were listed non-exhaustively by the Chair for further discussion.

(Maintaining)	(Substantive)	Human	(Participation)
(Ensuring)	(Meaningful)		(Involvement)
(Exerting)	(Appropriate)		(Responsibility)
(Preserving)	(Sufficient)		(Supervision)
	(Minimum level of)		(Validation)
	(Minimum		(Control)
	indispensa		(Judgment)
	ble extent		(Decisi
	of)		on)

It was also stated that the GGE could usefully focus at its future meetings on practice-related considerations, including the feasibility of some measures of human control, supervision or judgment, keeping in mind qualitative requirements such as reliability, predictability and explainability or auditability.

Agenda Item 6 (c)

Review of potential military applications of related technologies in the context of Group’s work

Delegation consideration of this agenda item proceeded through two interactive expert panel discussions, held on 12 April and 27 August 2018.

The April panel was composed of five independent experts, each providing their views on the possible military applications of emerging technologies in the area of lethal autonomous weapons systems.

Professor Mary Cummings (Professor, Duke University, Fellow of AIAA and Co-Chair of WEF’s Council on Artificial Intelligence and Robotics) spoke of the three core functions in weapons release in the military – the decider, the executor and the validator. In the case of drones, there was a human decider, a robotic executor and a human validator. In the case of lethal autonomous weapons systems, the validator would be a computer. In the next five to fifteen years, technology in the battlefield would move to cooperative control with multiple vehicles coordinating together, sharing information and using computer vision for target identification and validation. She expressed concern about the lack of knowledge among engineers on the subject as well as the challenges concerning certification of emerging technologies in the area of lethal autonomous weapons systems for both civilian and military uses. It is the civilian world that currently has dominance over artificial intelligence (AI), which is the reason why the military establishments are seeking help from multinational companies to develop algorithms. Due to the innate neuro-muscular lag of humans to perceive and act upon a situation, she said that lethal autonomous weapons systems would be far more discriminatory provided existing computer perception issues were sorted out. On human-machine teaming, she said that the ideal system in the battlefield would be the one that could strongly leverage the human-machine team.

Professor Dominique Lambert (University of Namur in Belgium, member of Royal Belgian Academy and International Academy for Philosophy of Science) noted that from the viewpoint of ethics, the question of responsibility is crucial. Ultimately, it is the human being who would have to take responsibility for his or her actions. A machine or artificial system cannot be held accountable or penalized for its actions. He also spoke about the limits to the algorithmic approach in problem solving. He said that even if we were able to introduce an ethical component into algorithms, it may be able to generate legal databases and provide information that could assist in legal and ethical decisions, but it would not be able to replace the role of a judge who actually makes a ruling. There is a need to draw a distinction between

assistance in decision making and taking a decision itself. One cannot ignore the interpretation aspect which comes into play in different scenarios. Ethical decisions cannot be replaced by formal processes.

Professor Viacheslev Pshikhopov (Director of R&D Institute of Robotics and Control Systems of Southern Federal University and Head of Robotics and Intelligent Systems Laboratory of Russia) discussed the anti-ship P-700 Granit from the 1980s to underline that decision making through intelligent technology where you have group control is an effective approach. At the same time, he acknowledged the difficulty in having an equivalent number of operators when given a group of weapons systems. He pointed out that in terms of motion control, there are limits for human intervention. He also spoke of the disadvantages of the existing control methods with a human operator, pointing out that human error causes ten per cent of total losses in unmanned systems. He highlighted the benefits of using emerging technologies in the area of lethal autonomous weapons systems by introducing an example from the United States showing that intelligent technology allows identification of mobile objects with lower error level than when humans carry out the same task. He posed a question – what would be a better choice from the point of IHL – to let the human operator take the wrong decision and let people die or to use smart lethal autonomous weapons systems with much less casualties? In response to a question about apportioning blame, he said that war was always inhumane and if it were to happen then diplomats and policy makers should also take the blame, and not just engineers, researchers or the military service.

Professor Pascale Fung (Director of the Centre for AI Research, Hong Kong University of Science and Technology, IEEE fellow, ISCA fellow and expert of the WEF Global Future Council on AI and Robotics) said that existing and emerging AI-based technologies continue to be used by militaries for some operations and several are being put to civilian uses including for speech recognition, autonomous flight, and facial and object recognition. She pointed out that rules-based systems are more susceptible to hacking and that machine learning allows the systems to become fully autonomous by optimizing functions such as minimal harm and risk. She highlighted the current trend away from the rules-based systems to machine learning based systems in the civilian sector. For the military, a hybrid approach could be a valid option. She expressed concern over the black box nature of some decision-making algorithms as well as their scalability. Machine learning could be used to reinforce proportionality, necessity and discrimination. She pointed to the need of cross-cutting education and science, technology, engineering and mathematics training for ethicists, doctors, lawyers, philosophers and humanists, as well as experience sharing across different cultures.

Professor Anthony Gillespie (Visiting Professor at University College London, Fellow of the Royal Academy of Engineering) referred to technologies, capabilities, competence and authority. He spoke of targeting functions taking the example of Observe, Orient, Decide and Act (OODA) as a sequence of multiple post-task decisions. He pointed out that there must be success-based criteria for each task, probably based on Rules of Engagement, to ensure operation within bounds set by humans. In his view, it should be possible to define the technical competence or authority level for machine decisions. He spoke about the qualitative aspects of the human-machine interface, trust issues between human and machine, separation of decision-making and reliability. He pointed to the importance of training operators so that they are aware of the limits of behaviour of the automated decision-making system, which is under their command. It is imperative for the human to trust the operative capability and reliability of the machine. He added that IHL sets a much higher threshold than international human rights law and commercial considerations. The difference between military and civilian systems is that the former will always be a part of a command and control chain.

Delegations held a lively moderated exchange with the panel on a variety of aspects of the issue. They considered which related technologies were most likely to be applied in the military domain, and whether some domains were more likely to see the early deployment of such technologies. The question of whether there would be situations in which the complete loss of human control, supervision or judgement would be acceptable or even desirable was raised. The concept of human-machine teaming was interrogated, specifically

whether it brought something different to the table with respect to the exercise of human control, supervision or judgement.

Delegations raised questions concerning the unique difficulties posed by emerging technologies in the area of lethal autonomous weapons systems for legal weapons reviews, including testing and whether a self-learning system would need to be reviewed every time it changed itself. The concept of machine decision-making was further explored, including whether this represented true machine agency or just algorithmic response to inputs, and whether complexity was relevant in this regard. There was discussion of ethics in the context of programming and design, as well as how meaningful human control could be applied over autonomous swarms. Delegations raised concerns about the black box nature of current AI algorithms and questioned whether advances in explainability would address this. Delegations sought examples of other sectors that provided good examples of optimizing human-machine collaboration in decision-making.

The August panel comprised five experts, each of whom tackled the issue of potential military applications of related technologies from different angles, and in light of the discussions of the Group in April.

Dr. Lydia Kostopoulos (Researcher, Digital Society Institute for European School of Management and Technology in Berlin) presented a matrix covering characteristics of autonomous weapons systems, and human involvement at different stages of the weapon's life cycle, based on the previous discussions of the Group. She also presented a related matrix covering issues of trust and confidence in autonomous weapons systems related to different stages of its life cycle and different forms of human involvement. She noted that a common vocabulary was important to progress discussions on human accountability and responsibility.

Dr. Gautam Shroff (Vice President and Chief Scientist, Tata Consultancy Services) discussed the importance of considering the level and degree of the human element in the use of force to ensure that the unintended or unpredictable application of force is prevented. The glaring errors made by otherwise highly accurate deep learning systems, and their high degree of confidence in patent errors, underline the importance of human judgement in the application of force. He also discussed the possibility of AI systems helping to prevent war crimes. He noted the apparent desire within the Group to prevent the unintended or unpredictable application of force and argued this would require considering the level and granularity of human involvement.

Lieutenant Colonel Christopher Korpela (Associate Professor and Director, Robotics Research Center, United States Military Academy) discussed human roles in the targeting cycle, noting that the human commander is held accountable for the use of force regardless of the weapons system used. Lt. Col. Korpela said that weapons must effectuate human intention in the use of force and that decisions on the use of force are not delegated to machines. He noted that if autonomy could provide humanitarian benefits and compliance with IHL, it would be more ethical to use it than not to do so.

Professor Anthony Gillespie (Fellow, Royal Academy of Engineering) presented his work looking at ensuring that the human who has delegated authority to a machine is held accountable. He suggested there was need for further consideration of where in the decision chain leading to the use of force it was acceptable for there to be no further human involvement. He underlined the importance of a clear distinction between human and machine actions for accountability and questioned whether rules of engagement could be put into a form intelligible to machines.

Mr. Knut Dörmann (Deputy Head, Legal Division, ICRC) welcomed the renewed interest in weapons reviews, noted the importance of States sharing national processes leading to the approval of weapons with autonomous functions, and provided an overview of four areas of challenges for weapons review of autonomous weapons systems: (i) deciding when a system warranted a review (not just new systems, but changes to existing systems); (ii) determining criteria for assessment (IHL, treaty law, operating environment, Martens clause, ethical considerations); (iii) assessment process (much is left to the discretion of States, there was a need to ensure sufficient expertise); (iv) and how to deal with uncertainty (which raises novel compliance issues and renders testing difficult).

In the interactive discussion that followed, delegations made several comments and posed a number of questions. The relationship between requiring human intervention and prohibiting lethal autonomous weapons systems was raised. The concept of human-machine pairing was raised, as well as the question of whether human involvement was more of a concern from an operational or an ethical perspective. The ability of machines to react to sudden changes of circumstances was raised, as was the question of holding developers accountable. More information on the use of autonomy in extreme environments such as outer space and the deep oceans was sought. It was also asked whether an autonomous system would be capable of determining whether it was operating in the context of a war or not.

With regards to legal weapons reviews, delegations were interested to know whether existing processes met the standards discussed and whether they were sufficient to ensure compliance with IHL. An important emerging commonality in the discussions on weapons reviews was the legal requirement for States party to the Additional Protocol I to the Geneva Conventions to conduct such reviews. A related query was how and by whom international standards for weapons review could be developed and how to monitor their implementation if such reviews were applied more broadly. Concerns regarding verification of weapons reviews, which could be tantamount to interference in the national affairs of a State, were also expressed. One delegation sought views on whether a distinction between weapons systems focused on individuals and those focused on other weapon systems, or between weapons operating in different environments, would be useful in the context of weapons reviews. It was noted that it was difficult to assess the quality of weapon reviews given only a limited number are publicly available. It is also important that they are conducted not only by lawyers but also incorporate adequate inter-disciplinary expertise. Given the lack of guidance in Article 36 about how to conduct reviews, it was agreed that sharing of national practices and experiences by States would be of value.

Agenda item 6 (d)

Possible options for addressing the humanitarian and international security challenges posed by emerging technologies in the area of lethal autonomous weapons systems in the context of the objectives and purposes of the Convention without prejudging policy outcomes and taking into account past, present and future proposals

Delegations reiterated the need to address various international security-related and humanitarian challenges arising from emerging technologies in the area of lethal autonomous weapons systems in the context of the objectives and purposes of the CCW. Among the security related challenges, the possibility of a new arms race, the danger of proliferation to non-State actors including terrorists, lowering of the threshold for the use of force, and the risk of cyber-attacks, hacking and spoofing of network-centric weapon systems were mentioned. In addition, the potential detrimental effects on global and regional security and stability were highlighted.

With regard to humanitarian challenges, delegations noted that emerging technologies in the area of lethal autonomous weapons systems raised a number of concerns including compliance with IHL, possible gaps in legal and political responsibility and accountability frameworks, and ethical questions raised by increased machine autonomy in the use of force. Several delegations also stressed the need to further assess possible humanitarian and military benefits of emerging technologies in the area of lethal autonomous weapons systems. Many delegations stressed that any possible regulation should not impede the development or application of related technologies in the civilian sector for beneficial uses.

Several delegations noted the value in taking a gender perspective when considering policy responses to the various concerns raised by emerging technologies in the area of lethal autonomous weapons systems.

Some delegations advocated the establishment of transparency and confidence-building measures (CBMs) as a means to prevent risks and unintended consequences such as a new arms race or proliferation to terrorist groups. In this connection, experience sharing in national policies and practices guiding the development, testing and use of intelligent

autonomous systems technologies broadly could be useful, bearing in mind national security considerations and commercial restrictions on proprietary information. It was further suggested that relevant scientific and commercial communities should be included in any transparency and CBM efforts, including to prevent the potential leak of newly developed technologies from the civilian sector to illegitimate non-State actors.

Support was expressed for the harmonization of legal weapons review processes and the elaboration of internationally agreed norms and standards. It was suggested that a technical comparative analysis of comprehensive weapons review mechanisms on development and procurement of new weapons could be undertaken, which would help to identify shortcomings and potential solutions. This could also result in a compendium of good national practice on weapons reviews and reviews of new means and methods of warfare and contribute towards establishing transparency and information exchange measures and cooperative technical analysis arrangements among States. The idea raised under agenda item 6 (c), that success and failure criteria could be programmed into autonomous weapons systems, and that these should be rigorously verified in the testing phase, was mentioned in this regard. The human control element was suggested as a necessary design requirement for all upcoming and future weapons developments. The discussion also touched on the feasibility of Article 36 reviews being applied in every stage of the life cycle of a weapon.

There was a call for greater technical expertise, which could be achieved by establishing technical bodies or expert groups to follow the development of emerging technologies. The possibility of establishing a standing group of technical experts under the CCW to monitor developments in science and technology and serve an early warning function was introduced in this connection. Such a mechanism could also provide a platform for States to continue to share how technology development and deployment is proceeding in the context of the objectives and purposes of the CCW, while allowing other stakeholders to provide inputs.

Some delegations expressed the view that the better application of existing international law, in particular IHL, was sufficient for maintaining the necessary human involvement in the use of weapons and force. Delegations also reaffirmed that the regime of international responsibility and accountability for the use of force in armed conflict fully applies when emerging technologies in the area of lethal autonomous weapons systems are employed and that individuals could be held accountable under applicable provisions at all stages of weapons development and deployment.

Other delegations viewed that new legally-binding provisions were necessary for addressing the humanitarian and international security challenges posed by emerging technologies in the area of lethal autonomous weapons systems. Such provisions could encompass measures mentioned above, including a comprehensive ban, prohibitions and restrictions of the type already seen within the CCW framework, or a positive requirement for maintaining human control over the critical functions of a weapons system.

Some delegations supported the 2017 proposal to establish a political declaration, which would affirm, *inter alia* and without prejudice to policy outcomes, that humans should be responsible for a) making final decisions with regard to the use of force and b) maintaining control over autonomous weapons systems. This could be followed by the development of codes of conduct and further confidence building measures. A technology review functionality was proposed as part of such an approach. This could serve a lighthouse function on relevant developments in emerging technologies in the area of lethal autonomous weapons systems. Some delegations, citing the possible inadequacy of non-binding approaches in the context of lethal autonomous weapons systems, viewed a political declaration as an interim step prior to the conclusion of a legally-binding instrument, perhaps in the form of a new protocol to the CCW. Several delegations called for a moratorium on the development of lethal autonomous weapons systems.

Some delegations underlined the need to further develop understanding of the subject within the context of the GGE's mandate before considering which option(s) to pursue.

There was broad support for continuation of the GGE in 2019. In addition to the agenda items, delegations therefore discussed various options for the mandate of the Group in 2019, including:

- A continuation of the existing discussion mandate;
- An amended discussion mandate, reflecting the progress made by the Group to date, which could focus the work of the Group on compiling elements for an outcome document;
- A negotiating mandate.

The following amendment was proposed by Germany to the existing mandate:

“The Group of Governmental Experts related to emerging technologies in the area of lethal autonomous weapons systems (LAWS) in the context of the objectives and purposes of the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons shall meet for a duration of ten days in 2019 in Geneva in accordance with Decision 1 of the Fifth Review Conference of the High Contracting Parties to the Convention (CCW/CONF.V/10), consistent with CCW/CONF.V/2, *with a focus on exploring and agreeing options for an Outcome Document.*”

With regards to the proposal for a negotiating mandate, Austria, Brazil and Chile proposed the following language:

“The 2018 Meeting of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects decides to establish an open-ended Group of Governmental Experts to negotiate a legally-binding instrument to ensure meaningful human control over critical functions in lethal autonomous weapons systems.”

A view was expressed that IHL is fully applicable to emerging technologies in the area of lethal autonomous weapons system and its modernization, or adaptation, to such weapons systems is not needed. Another view was that IHL was applicable but was insufficient. It was also stated that the meeting of the Group in 2019 should be rationalised to five days. There was some discussion on the relative merits of holding two one-week meetings or one two-week meeting; the former allowing for intersessional work and increasing the chance that experts with many competing demands could attend and the latter being potentially more cost effective (for participants from capitals).

Delegations commended the discussions which had taken place during the two weeks and the progress made in terms of reaching a greater understanding as well as a shared vernacular on the subject matter in general and the notions of the human element and of accountability in the use of force in particular. The relevant references in Secretary-General’s new agenda for disarmament were highlighted by some delegations in this regard. The interactive discussion on characterization, human-machine interface and military applications of relevant technologies had helped underscore areas of convergence, such as the applicability of IHL, while clarifying areas that required further work. Delegations underscored the value added of the Possible Guiding Principles that had emerged out of the discussions in 2018 as an ‘early harvest’. They could be updated as work progressed and could also help policy communities working on issues related to intelligent autonomous systems outside of the CCW framework. Delegations welcomed the contributions and involvement of the civil society, industry and youth in the process and expressed appreciation for the contribution to the discussion made by the ICRC. Delegations agreed that any future policy measures should not impede progress in or access to peaceful uses of intelligent autonomous technologies and that these technologies should be seen as attributes and not be anthropomorphized. The suitability of the CCW for a continuation of the discussions on emerging technologies in the area of lethal autonomous weapons systems was reaffirmed by delegations.