GCSP
Geneva Centre for Security Policy

JANUARY 2021 | ISSUE 16

Strategic Security Analysis

Reducing Military Risks through OSCE Instruments: The Untapped Potential in the European Arms Control Framework

Naomi Egel

www.gcsp.ch

Reducing Military Risks through OSCE Instruments: The Untapped Potential in the European Arms Control Framework

Naomi Egel

The Organization for Security and Co-operation in Europe (OSCE), jointly with the Institute for Peace Research and Security Policy at the University of Hamburg (IFSH) and in partnership with the Geneva Centre for Security Policy (GCSP), Moscow State University of International Relations (MGIMO) and the Vienna Centre for Disarmament and Non-proliferation (VCDNP), has launched an “OSCE-IFSH Essay Competition: Conventional Arms Control and Confidence- and Security-Building Measures in Europe”. The project aims at facilitating the continuity of knowledge and expertise on arms control and CSBM processes at the OSCE among students and recent graduates interested in peace and security studies. This essay has participated in the competition and has been awarded the second prize.

Key Points

- As autonomous weapons systems (AWS) increase in military importance, they pose significant risks of miscommunication, miscalculation and inadvertent conflict escalation.
- The Organization for Security and Co-operation in Europe (OSCE) has a strong history of developing confidence-building measures (CBMs) to reduce military risks stemming from other types of weapons.
- The OSCE should develop CBMs for partially autonomous weapons systems. Such CBMs should provide information about AWS features and doctrine for their use, to increase transparency and build trust between states.
- OSCE CBMs could provide a foundation for the global governance of autonomous weapons in other multilateral venues.

About the Author

Naomi Egel is a PhD candidate in political science at Cornell University. Her research examines the politics of multilateral weapons governance. She is also the Janne Nolan Nuclear Security Visiting Fellow at the Truman Center for National Policy and the Truman National Security Project.

Autonomous weapons systems (AWS) are widely regarded as a game changer in the field of international security and an increasingly important element of military operations. However, they pose heightened risks of miscommunication, miscalculation, and the inadvertent escalation of a conflict that could increase tensions and conflict between states. Although the development of AWS and their use in military operations vary widely among Organization for Security and Co-operation in Europe (OSCE) states, many states both within and outside the OSCE are incorporating increasing numbers of AWS into their armed forces, raising concerns over the unintended risks associated with these weapons systems. Even states that do not currently employ AWS have concerns over other states' use of this technology and the implications of AWS for international security and strategic stability. In the absence of formal treaties governing the use of AWS, confidence-building measures (CBMs) would provide a valuable tool for reducing military risks associated with AWS. This essay proposes that the OSCE develop CBMs for partially autonomous weapons systems, building on its successful history of developing CBMs to reduce other military risks.

Even states that do not currently employ AWS have concerns over other states' use of this technology and the implications of AWS for international security and strategic stability.

Risks posed by AWS

AWS are weapons that can perform some or all of their functions (including target detection, selection, and engagement) without requiring intervention by a human operator. This definition includes weapons for which a human operator can intervene to override an autonomous function, as well as 'semi-autonomous' weapons systems in which some functional decisions are made by human operators and some by the weapons system. Fully autonomous weapons that lack any form of human operator intervention have not yet been deployed. Examples of AWS include air defence systems, armed drones, and active protection systems for armoured vehicles. Although a degree of autonomy in weapons systems is not new (e.g. landmines could be considered partially autonomous in that they are not activated by an operator), the heightened levels of autonomy in recently developed AWS pose significant military risks.

AWS are increasingly important in strategic planning and military operations. Although some policymakers argue that AWS can help to reduce risks to both combatants and civilians (e.g. the 2018 US Department of Defense Artificial Intelligence Strategy), the increasing level of autonomy that is being built into weapons systems poses considerable military risks, most notably those of miscommunication, misunderstanding and the inadvertent escalation of a conflict. These risks posed by AWS undermine the stability of the balance of military power among nations. For example, one major purported advantage of AWS is the speed with which they can react and respond to changes in their environment. However, the increased speed at which AWS operate also hampers humans' ability to correct mistakes made by AWS, which increases the risk of the inadvertent escalation. More broadly, by reducing human involvement in decision-making processes, AWS increase the risk of signals and actions being misinterpreted by other parties, both when autonomy is a feature of a particular weapon and when it is involved in the decision-making cycle. Although AWS do pose military risks on their own (e.g. they could malfunction), the greater risks they pose involve how other actors could (mis)interpret and react to an actor's use of these weapons systems.

Different types of autonomy have varied implications for military risk and the stability of the international system. For example, AWS that rely on unsupervised machine learning to improve their functioning are inherently unpredictable, since the user does not specify the nature or directions of the improvements made and does not know what elements the programme controlling an AWS uses to sort and categorise information. This increases the risk of the inadvertent escalation of a conflict if a system 'learns' to respond in ways that appear escalatory or unclear to others. In contrast, AWS that are pre-programmed and do not 'learn' on their own are more predictable and thus pose a lower risk of miscommunication and inadvertent escalation. However, the algorithms used in AWS are generally designed to react to a set of information with specific parameters. But even if these algorithms can accurately classify information and produce reliable results in training scenarios, the 'fog of war' endemic to military engagement makes it extremely difficult to design AWS for real-world military engagement. Even taking civilian autonomous systems only slightly out of the context in which they have been programmed to operate has produced unpredictable – and in some cases disastrous – results (e.g. self-driving cars crashing into other cars because of what appears to humans to be a minor change in the context in which they operate). Although miscalculation, miscommunication and inadvertent escalation are certainly not new military risks, the increasing use of AWS in military operations heightens these risks in new ways.

AWS that rely on unsupervised machine learning to improve their functioning are inherently unpredictable.

The OSCE's contribution to risk reduction

Currently, global governance of AWS is under discussion within the framework of the Convention on Certain Conventional Weapons (CCW). This process, however, has been slow moving and fraught with disagreement over what AWS are, particularly regarding the distinction between fully and partially autonomous weapons systems. Rather than relying solely on the CCW to address the risks posed by AWS, regional organisations could develop mechanisms to reduce risks and build confidence that states' increasing use of partially autonomous weapons would not be destabilising and escalatory. Such efforts could also help to energise the process under way in the CCW and facilitate a global agreement there. The OSCE has made such contributions before: OSCE information exchanges on small arms and light weapons, beginning with the 1993 OSCE Principles Governing Conventional Arms Transfers, helped pave the way for the 2013 Arms Trade Treaty.

CBMs are voluntary measures designed to communicate 'credible evidence of the absence of feared threats by reducing uncertainties and by constraining opportunities for exerting pressure through military activity'. Given the OSCE's membership and history, OSCE CBMs would make a particularly significant contribution to reducing the risks posed by AWS. The OSCE's membership is both broad (the organisation has 57 participating States, encompassing not only all European states, but Central Asian and North American states as well) and includes seven of the top ten arms producers in the period 2015-2019. OSCE CBMs, thus, have a legitimacy based on both the breadth and number of states that subscribe to them and the involvement of key arms producers.

Among international organisations (both regional and global), the OSCE is distinctive for its history of establishing strong norms for using CBMs as tools for risk reduction. AWS CBMs would build on the OSCE's robust record of reducing military risk through CBMs. The Vienna Document - the cornerstone of OSCE CBMs - has provided a strong framework for building

confidence among OSCE participating states, reducing military risks and increasing security in the OSCE region. Other OSCE instruments like the Conventional Forces in Europe Treaty and the Open Skies Treaty provide further information about states' activities and are thus able to build trust among OSCE members (even though both treaties are currently under strain). However, these agreements are designed to address the risks posed by conventional weapons rather than risks from AWS.

Still, the OSCE's recent adoption of CBMs for Information and Communication Technologies (adopted in 2013 and updated in 2016) demonstrates that the OSCE and its approach to CBMs are well suited to developing risk-reduction CBMs applicable to advanced technologies. The OSCE's cyber CBMs have also provided a template for other states or regional organisations to develop their own cyber CBMs, building confidence and reducing risk both within and beyond the OSCE's membership; CBMs for AWS could play a similar role. Although the original Vienna Document was negotiated at the end of the Cold War, its updates and revisions, together with the development of other OSCE CBMs, reflect how CBMs are not merely an outdated Cold War legacy, but continue to play an important role in building confidence and reducing risks. Moreover, the OSCE's success in negotiating cyber CBMs in the context of tensions between the United States and the Russian Federation indicates that such tensions are not an insurmountable barrier to reaching agreement on CBMs of various kinds within the OSCE. Although negotiating new agreements in international organisations is rarely easy, the OSCE's track record shows that it is capable of doing so. This stands in contrast to many other international organisations in which geopolitical tensions have prevented progress on agreements to govern new risks. The OSCE's resilience in the face of such geopolitical challenges demonstrates its importance and efficacy as a vehicle for addressing security threats and military risks.

CBMs provide information about other states' capabilities and intentions in order to reduce risks arising from miscalculation and miscommunication and to build trust between parties.

Given the difficulty of negotiating legally binding treaties governing weapons in the current geopolitical environment, informal and voluntary CBMs provide a valuable and practical way to reduce military risks arising from AWS. CBMs provide information about other states' capabilities and intentions in order to reduce risks arising from miscalculation and miscommunication and to build trust between parties. Although CBMs are, by design, voluntary measures, OSCE CBMs have provided a valuable framework for reducing tensions and lowering risks associated with other weapons and military activities. Additionally, in 2019 the OSCE Parliamentary Assembly called on OSCE members to support international negotiations to ban lethal autonomous weapons (also referred to as fully autonomous weapons). CBMs for partially autonomous weapons would support efforts in this area while also reducing the risks posed by these kinds of weapons systems.

CBMs for AWS

Despite the risks posed by AWS, little information is available regarding which types of partially autonomous weapons systems states possess and how they employ these systems. Rather than compete with the CCW's efforts to govern fully autonomous weapons, the OSCE could develop CBMs for partially autonomous weapons (in which some functions are performed autonomously, but humans can intervene in certain aspects of the weapons' operations). By including a spectrum of partially autonomous weapons systems, OSCE CBMs could also complement efforts under way in the CCW framework. The CCW's discussions on AWS have focused on fully autonomous weapons (i.e. weapons that can select and fire at targets without human intervention). Yet the focus on fully autonomous weapons ignores the many different ways in which autonomy is incorporated into weapons systems. Few AWS can at present be classified as fully autonomous, but many of them still pose risks for miscommunication and unintended escalation. Exchanging information through OSCE-designed AWS CBMs would bypass this binary and offer a template for reducing risks posed by a fuller spectrum of AWS. Drawing on precedents in the Vienna Document and the OSCE's cyber CBMs, CBMs for partially autonomous weapons systems could include the elements discussed below.

Exchanging information that shows how human operators remain fully in control of AWS would help reduce uncertainty regarding this crucial element of such weapons systems.

Firstly, they could include a *register* of the various AWS that each state possesses (e.g. different kinds of armed drones, sentry-type systems, etc.). A 2017 Stockholm International Peace Research Institute study identified 381 different types of AWS through open-source research. Although this is not a comprehensive account of all types of AWS, the diversity and multiplicity of even this sample indicate the utility of a register of AWS, which would provide greater clarity regarding states' AWS capabilities. At present, states reveal very little information about the overall landscape of their AWS capabilities, which further increases the risk of miscommunication and miscalculation that could lead to inadvertent escalation. Providing more information about states' AWS capabilities is a crucial first step towards reducing risks stemming from AWS. Although the Vienna Document calls for OSCE states to notify one another in their annual reporting when they deploy new types or versions of major weapons systems - including the number of weapons systems deployed - this requirement does not distinguish between AWS and manually operated weapons systems. A separate register of AWS would provide the foundation for other CBMs and forms of risk reduction.

Secondly, OSCE CBMs could involve an *exchange of information regarding the nature and extent of autonomy and human control* in these weapons systems. For example, AWS systems may be autonomous in their subsystems dealing with such matters as navigation, target selection and/or decision to engage a target. Human operators may have the ability to override any automated decisions or any part of the autonomous operation (i.e. 'human on the loop'), or autonomous functions may have specific decision points that require human approval (i.e. 'human in the loop'). Exchanging information that shows how human operators remain fully in control of AWS would help reduce uncertainty regarding this crucial element of such weapons systems. Moreover, while several governments that currently employ and/or are developing AWS claim that humans would always remain in ultimate control, the practical meaning of this claim is ambiguous, and 'human control' can take various forms. Greater transparency regarding the elements of autonomy and human control in AWS would help build confidence that these weapons systems are being designed in ways that would prevent inadvertent conflict escalation by machines.

Thirdly, CBMs could also include an *exchange of information regarding the doctrine governing AWS use*, which would be distinct from the software controlling AWS. For example, there may be geographical or operational doctrinal limits on use, such as using autonomous systems to select structures for targeting, but not to select individuals. Sharing information about doctrine for AWS use would go far in building confidence and reducing military risk by clarifying the intent behind states' development and use of AWS. Exchanging information related to the nature of autonomy in AWS and doctrine for their use would also provide a foundation for sharing best practices, which would further build confidence and reduce the risk of misinterpreting or misunderstanding other states' intentions regarding AWS. The Vienna Document calls on states to exchange information about their respective defence policies, including military doctrine. However, it does not specify the elements of military doctrine or the level of detail that should be exchanged, and thus does not explicitly cover the development and use of AWS.

Exchanging information related to the nature of autonomy in AWS and doctrine for their use would also provide a foundation for sharing best practices...

Fourthly, OSCE CBMs could involve *site visits and observations*, building on the Vienna Document's provisions for site visits and observations of military activities as part of its CBMs. Site visits and observations as part of AWS CBMs could involve, for example, observations of training activities involving AWS or demonstrations of how human operators intervene in and maintain control over AWS operations. Such provisions would help to build confidence that OSCE states' reporting on other CBMs accurately represents their development and use of AWS. Although site visits and military observations are often contentious in other multilateral forums, the OSCE has normalised voluntary site visits to military bases as part of the CBMs in the Vienna Document and other CBM instruments. Given this precedent, site visits and observations could reasonably be an element of AWS CBMs.

Information about the types of AWS that states employ, the aspects of autonomy and human control in AWS, the intended uses of AWS, and demonstrations of AWS use would together provide greater information regarding the scope of AWS deployment and use in military operations. They would also strengthen confidence that AWS are being designed and deployed in ways that aim to avoid inadvertent conflict escalation. Although these CBMs would not fully eliminate the military risks posed by AWS (including the risks posed by AWS malfunctions), they could help to reduce misunderstandings and thereby moderate the responses by other states. Additionally, such CBMs could be applied to future AWS. Much of the concern over the risks posed by AWS is centred on cutting-edge and future systems of this kind. Although the exact nature of future AWS cannot be predicted, CBMs that are designed to reduce the risks posed by various elements of autonomy would be useful for reducing the risks posed by both existing and future AWS.

Given the ever-increasing integration of autonomy into weapons systems, these AWS CBMs could take the form of an addendum to the Vienna Document, which covers all conventional weapons systems. The CBMs proposed here build on precedents in the Vienna Document, and thus amending the Vienna Document to address risks arising from AWS would be a logical progression that would be acceptable to states. Alternatively, AWS CBMs could be agreed as a stand-alone set of CBMs, in the style of the OSCE's CBMs for Information and Communication Technologies. Regarding the institutional format used to discuss, negotiate and develop AWS CBMs, the OSCE Structured Dialogue (established in 2016) could provide a dedicated venue for negotiating such CBMs. An alternate approach would be for the OSCE to set up a new venue for such negotiations.

Far-reaching benefits of AWS CBMs

In addition to reducing the military risks posed by AWS, the process of developing AWS CBMs in the OSCE would also help to strengthen policymakers' understanding of the role of autonomy in military operations. This would facilitate further international cooperation on AWS governance. Within the context of the CCW - at present the primary global venue considering AWS governance - debates over AWS are often stymied by diplomats' unfamiliarity with these systems. Greater transparency and information regarding states' AWS capabilities (including the varying elements of autonomy in these systems) and use would ameliorate this impediment to reaching an agreement. Developing AWS CBMs in the OSCE would thus help to provide a foundation for the global governance of AWS. Moreover, while OSCE CBMs would not include China (a leader in the development and use of AWS), OSCE CBMs could lay the groundwork for a global agreement that would include China. An OSCE-developed template of CBMs could also be applied outside the OSCE region, either in a global agreement or through agreements in other regional organisations.

CBMs for partially autonomous weapons could also strengthen confidence in a future agreement banning fully autonomous weapons.

Even if the CCW eventually bans fully autonomous weapons, OSCE CBMs would still be very useful for reducing the risks posed by partially autonomous weapons systems. Autonomy is and will continue to be important in military operations, regardless of whether fully autonomous systems are banned. CBMs for partially autonomous weapons could also strengthen confidence in a future agreement banning fully autonomous weapons. By providing information about the extent of human control over AWS and how AWS would be used, such CBMs could be used in assessments of states' compliance with such an agreement.

OSCE AWS CBMs could also provide the basis for a future OSCE code of conduct for AWS involving doctrinal restrictions on use. The OSCE has precedents here, too, with its Principles Governing Conventional Arms Transfers, its Code of Conduct on Politico-Military Aspects of Security, and its Principles Governing Non-Proliferation. Like CBMs, codes of conduct (including principles governing behaviour) are voluntary rather than legally binding agreements. However, whereas CBMs reduce risk and build trust by providing information about other states' capabilities and intentions, codes of conduct go further by endorsing certain behaviours and proscribing others. An OSCE AWS code of conduct could further reduce military risks by proscribing certain uses of AWS or forms of autonomy (e.g. fully autonomous weapons that lack human supervision) and/or endorsing certain requirements for human control over AWS. By building confidence and trust among states, CBMs could facilitate further cooperation and agreements.

Conclusion

CBMs for AWS would make a significant contribution to reducing the military risks posed by AWS and - equally important - are a realistic goal. Although agreement within the OSCE cannot be assured, its strong history of developing CBMs and its ability to make progress towards reducing military risk even in a climate of geopolitical tensions makes the OSCE uniquely well positioned to develop AWS CBMs. The growing proliferation of AWS and their increasing importance in military operations means that risk reduction is an urgent priority. Given the numerous challenges to arms control and cooperative security measures in other international organisations, OSCE CBMs offer one of the best opportunities for making progress in reducing the military risks of AWS. Moreover, they would provide a foundation for further governance of AWS by other international organisations. The CBMs proposed here would provide greater transparency regarding states' capabilities and intentions regarding AWS, and in doing so, would reduce the risks of miscommunication, miscalculation and inadvertent conflict escalation that these systems pose.

The growing proliferation of AWS and their increasing importance in military operations means that risk reduction is an urgent priority.

Bibliography

Altmann, J. and F. Sauer, 'Autonomous Weapon Systems and Strategic Stability', *Survival*, Vol.59(5), 2017, pp.117-142.

Borghard, E. and S. Lonergan, 'Confidence Building Measures for the Cyber Domain', *Strategic Studies Quarterly*, Vol.12(3), Fall 2018, pp.10-49.

Boulanin, V. and M. Verbruggen, *Mapping the Development of Autonomous Weapons System*, Stockholm, Stockholm International Peace Research Institute, 2017.

Boulanin, V. et al., *Limits on Autonomy in Weapons Systems*, Stockholm, Stockholm International Peace Research Institute, 2020.

Garcia, D., 'Future Arms, Technologies, and International Law: Preventive Security Governance', *European Journal of International Security*, Vol.1(1), 2016, pp.94-111.

Holst, J. and K. Melander, 'European Security and CBMs', *Survival*, Vol.19(4), 1977.

Holtom, P., 'The OSCE and the Arms Trade Treaty: Complementarity and Lessons Learned', *OSCE Yearbook,* Vol.21, 2015, pp.327-342.

International Committee of the Red Cross (ICRC), *Autonomous Weapons Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons*, Versoix, Switzerland, ICRC, 2016.

Israeli Permanent Mission in Geneva, 'Statement by Ms Maya Yaron, Minister-Counsellor, Deputy Permanent Representative to the Conference on Disarmament', meeting of the Convention on Certain Conventional Weapons Group of Governmental Experts on Lethal Autonomous Weapons Systems, Geneva, 9-13 April 2018.

Sauer, F., 'Stopping 'Killer Robots': Why Now Is the Time to Ban Autonomous Weapons Systems', *Arms Control Today* 46(8), October 2016, pp.8-13.

Scharre, P., *Army of None: Autonomous Weapons and the Future of War*, New York, W.W. Norton, 2019.

Scharre, P., 'The Militarization of Artificial Intelligence', *Texas National Security Review*, 2 June 2020.

United States Department of Defense (USDoD), Directive Number 3000.09: Autonomy in Weapon Systems, Washington, DC, USDoD, 2017.

United States Department of Defense, *Summary of the 2018 Department of Defense Artificial Intelligence Strategy*, Washington, DC, USDoD, 2018.

United Kingdom Permanent Mission in Geneva, 'Statement for the General Exchange of Views', meeting of the Convention on Certain Conventional Weapons Group of Governmental Experts on Lethal Autonomous Weapons Systems, Geneva, 9-13 April 2018.

Wezeman, P. et al., *Trends in International Arms Transfers, 2019*, Stockholm, Stockholm International Peace Research Institute, 2020.

Endnotes

1. V. Boulanin et al., *Limits on Autonomy in Weapons Systems*, Stockholm, Stockholm International Peace Research Institute, 2020, https://www.sipri.org/publications/2020/other-publications/limits-autonomy-weapon-systems-identifying-practical-elements-human-control-0.

2. P. Scharre, 'The Militarization of Artificial Intelligence', *Texas National Security Review*, 2 June 2020, https://www.cnas.org/articles-multimedia?author=paul-scharre.

3. J. Holst and K. Melander, 'European Security and CBMs', *Survival*, Vol.19(4), p.147.

4. T. Minárik, 'OSCE Expands Its List of Confidence-Building Measures for Cyberspace: Common Ground on Critical Infrastructure Protection', NATO Cooperative Cyber Defence Centre of Excellence, 10 March 2016, https://ccdcoe.org/incyder-articles/osce-expands-its-list-of-confidence-building-measures-for-cyberspace-common-ground-on-critical-infrastructure-protection/.

5. E. Borghard and S. Lonergan, 'Confidence Building Measures for the Cyber Domain', *Strategic Studies Quarterly*, Vol.12(3), Fall 2018, pp.10-49, https://www.airuniversity.af.edu/Portals/10/SSQ/documents/Volume-12_Issue-3/Borghard-Lonergan.pdf.

6. V. Boulanin and M. Verbruggen, *Mapping the Development of Autonomous Weapons System*, Stockholm, Stockholm International Peace Research Institute, 2017, https://www.sipri.org/publications/2017/other-publications/mapping-development-autonomy-weapon-systems.

GCSP
Geneva Centre for Security Policy

Where knowledge meets experience

The GCSP Strategic Security Analysis series are short papers that address a current security issue. They provide background information about the theme, identify the main issues and challenges, and propose policy recommendations.

Geneva Centre for Security Policy - GCSP
Maison de la paix
Chemin Eugène-Rigot 2D
P.O. Box 1295
CH-1211 Geneva 1
Tel: + 41 22 730 96 00
Fax: + 41 22 730 96 49
e-mail: info@gcsp.ch
www.gcsp.ch

ISBN: 978-2-88947-302-1