CONTROLLING HUMANS AND MACHINES

*Bryant Walker Smith**

INTRODUCTION

Within the debate over the "meaningful human control" of lethal weapons, this is a limited piece. It takes no position on whether an otherwise lawful killing is, or should, be unlawful when carried out by an automated agent instead of by a human agent. It also takes no position on whether such an automated agent—a so-called autonomous weapon system (AWS)—is, or should be, unlawful per se. Furthermore, the piece assumes that some concept of "meaningful human control" is, or will become, relevant to international humanitarian law's (IHL) treatment of these systems. Unlike most articles considering "meaningful human control," this article does not focus on the role that a human should play in an otherwise automated weapon system.[1] Rather, it reverses these human and machine roles to consider automated systems that limit human-initiated lethal force. After discussing the concept of control generally, this piece argues, first, that a bias toward human authority could impede eventual restrictions on that authority and, second, that the line between automated systems that initiate lethal force and automated systems that restrict that force is potentially unclear.

ON CONTROL

As others have observed, "meaningful human control" is intuitively appealing even though—or perhaps because—it is substantively ambiguous.[2] Content can be supplied to this term of art but should not be assumed.[3] This is because "control" has too many meanings in too many contexts to universally express anything useful.[4] Its object can be an actor or an action, and its effect can be to check or to

*Assistant Professor, University of South Carolina School of Law and (by courtesy) School of Engineering; Affiliate Scholar at the Center for Internet and Society, Stanford Law School; Faculty Affiliate, Rule of Law Collaborative; Chair, Emerging Technology Law Committee of the Transportation Research Board of the National Academies. I wrote this short piece as part of a conference sponsored by the Institute for International Law and Public Policy at Temple University Beasley School of Law and the International Committee of the Red Cross. I am grateful to the organizers of this conference for soliciting my piece and to the editors of the Temple International and Comparative Law Journal for publishing it. All of my relevant publications are available at http://www.newlypossible.org.

1. *See*, *e.g.*, Paul Sharre & Michael C. Horowitz, An Introduction to Autonomy in Weapon Systems 6 (Feb. 2015) (unpublished manuscript) (on file with the Center for a New American Security) (describing the relationship between a person and the machine in the section "Human-Machine Command-and-Control Relationship").

2. *See* Michael C. Horowitz & Paul Sharre, Meaningful Human Control in Weapon Systems: A Primer 6 (Mar. 2015) (unpublished manuscript) (on file with the Center for a New American Security) (explaining the definition of "meaningful human control").

3. *See id*. at 15 (explaining that "meaningful human control" must be used in context).

4. *See, e.g.*, Bryant Walker Smith, *Automated Vehicles Are Probably Legal in the United*

commandeer.[5] This can lead to confusion:

> Because they obscure rather than clarify that system, casual references to humans who are "in control," "in the loop," "out of control," or "out of the loop" and to automated systems that are "under control," "under human control," "under computer control," or "out of control" are particularly unhelpful.
>
> Phrases like these are susceptible to numerous technical, legal, and popular interpretations. Consider, for example, the "control" of an automated vehicle. An engineer might picture a real-time control loop with sensors and actuators, a lawyer might envision a broad grant of authority from human to machine analogous to a principal-agent relationship, and the public might imagine runaway cars and killer robots.[6]

In other words, control is, as a concept, essentially uncontrollable. Because of this potential for significant confusion, the standards organization formerly known as the Society of Automotive and Aerospace Engineers studiously avoided the general term "control" in defining levels of driving automation.[7] Its definitions document instead specifies the respective roles of human driver and automated driving system at each level.[8] The Defense Science Board, while rejecting levels of automation, likewise conceives of humans not as external supervisors of a given system, but rather as internal participants in a broader system.[9]

These approaches are consistent with a disciplined approach to engineering control theory, in which engineers "first describe the control system they actually intend: the goals, inputs, processes, and outputs to the extent they are determined by a human designer and the authority of the human or computer agents to the extent they are not."[10] This discipline, however, is sometimes lacking in the engineering realm and is further challenged when that realm overlaps with law.[11]

I first analyzed "control" at length in the context of automated motor

---

States, 1 TEX. A&M L. REV. 411, 436–37 (2014).

5. *See, e.g.*, *id*.

6. Bryant Walker Smith, *Lawyers and Engineers Should Speak the Same Robot Language*, *in* ROBOT LAW 83–84 (Ryan Calo, A. Michael Froomkin & Ian Kerr eds., 2016).

7. *See* Bryant Walker Smith, *Summary of Levels of Driving Automation for On-Road Vehicles*, SAE LEVELS OF DRIVING AUTOMATION (Dec. 18, 2013, 10:33 AM), http://cyberlaw.stanford.edu/loda (showing levels of driving automation for on-road vehicles).

8. *See id*.

9. *See* DEP'T OF DEF., DEF. SCI. BD., THE ROLE OF AUTONOMY IN DOD SYSTEMS 23–24 (2012), http://www.fas.org/irp/agency/dod/dsb/autonomy.pdf (discussing why defining levels of automation is not useful).

10. *See* Smith, *Lawyers and Engineers Should Speak the Same Robot Language*, *supra* note 6, at 85 (explaining the importance of defining "control" in structural terms).

11. *See id*. (providing an example of the difference between an engineer's visions versus a lawyer's vision).

vehicles.[12] The requirement in two multilateral treaties on road traffic that "[d]rivers shall at all times be able to control their vehicles or guide their animals"[13] led to concern in Europe that these conventions would impede the broad deployment of automated motor vehicles.[14] Within the United Nations committee responsible for these treaties, States have expressed a variety of perspectives on the compatibility of this provision with automated driving and, at least implicitly, on the proper meaning of control.[15]

The successful amendment of one of these two road traffic conventions suggests that States believe that drivers are still "able to control their vehicles" if they can "override" or "switch off" any "vehicle systems which influence the way [their] vehicles are driven."[16] This clarification is striking because it codifies a driver's authority to disable vehicle systems that may be intended to enhance safety.[17] Two considerations do temper this conclusion. First, per the amendment, a vehicle system that complies with international automotive regulations does not need to permit human override.[18] Second, the ability to control is different than the exercise of control.[19]

Many legal questions implicate this "complex spectrum" of control,[20] particularly where retrospective responsibility (i.e., liability) is linked to prospective responsibility (i.e., obligation).[21] The responsibility of individual commanders for war crimes committed by others is premised on a narrow conception of "effective control."[22] Similarly, the International Court of Justice and

12. *See* Smith, *Automated Vehicles Are Probably Legal in the United States*, *supra* note 4, at 435–41 (analyzing "control" in the context of motor vehicles).

13. Geneva Convention on Road Traffic art. 8, Sept. 19, 1949, 3 U.S.T. 3008, 125 U.N.T.S. 3. In a slight change of wording, the 1968 Vienna Convention on Road Traffic states that "[e]very driver shall at all times be able to control his vehicle or to guide his animals." Convention on Road Traffic art. 8, Nov. 8, 1968, 1042 U.N.T.S. 17.

14. *See* Smith, *Automated Vehicles Are Probably Legal in the United States*, *supra* note 4, at 431–32 (explaining various viewpoints on the meaning of control and how it is interpreted in different ways).

15. *See id*. (discussing positions taken by Germany, Switzerland, and the Netherlands regarding the provision's application to driver assistance systems).

16. U.N. Econ. & Soc. Council, Econ. Comm'n for Europe, Inland Transp. Comm., Rep. of the sixty-eighth session of the Working Party on Road Traffic Safety, ECE/TRANS/WP.1/145 (Apr. 17, 2014), http://www.unece.org/fileadmin/DAM/trans/doc/2014/wp1/ECE-TRANS-WP1-145e.pdf.

17. *Id*.

18. *Id*.

19. *See Axcelis Techs., Inc. v. Applied Materials, Inc*., No. CIV.A 01- 10029DPW, 2002 WL 31761283, at *6 (D. Mass. Dec. 10, 2002) (suggesting a difference between ability to control and exercising control).

20. *See* Smith, *Automated Vehicles Are Probably Legal in the United States*, *supra* note 4, at 420–22 nn. 34–42 (explaining the various factors that complicate the term "control").

21. *See* Smith, *Lawyers and Engineers Should Speak the Same Robot Language*, *supra* note 6, at 8 (explaining how liability can be defined in a legal, technical, and moral sense).

22. *See* United Nations Diplomatic Conference of Plenipotentiaries on the Establishment of an International Criminal Court, *Rome Statute of the International Criminal Court*, art. 28, U.N.

the International Tribunal for the Former Yugoslavia have disagreed about the conditions under which a State exercises enough control over another entity that the actions of that entity are attributable to the State.[23] In the United States, Congress deliberately declined to define "control" in the context of securities, while the Securities and Exchange Commission subsequently adopted a broad definition premised on the possession, rather than the exercise, of power.[24] Corresponding terms within various languages may also reflect subtle differences in the conception of control.[25]

The concern here is not that "meaningful human control" is indefinable. Like any term of art, this phrase can be defined as precisely as language allows. Indeed, its original proponents have stated that meaningful human control over an individual attack requires, at a minimum, adequate information, positive action, and corresponding accountability.[26]

The concern, rather, is practical: Merely defining a term will not banish its many other connotations or even its many other denotations. Even though the original proponents of "meaningful human control" focused on individual attacks, their term "is entering currency without this modifier, leading to various interpretations" involving, for example, control over "weapon systems" or "the critical functions of autonomous weapons."[27]

Once unmoored, terms can float far from their original meaning. As the examples above suggest, "control" faces an especially turbulent sea.[28] Invoking

---

Doc. A/CONF.183/9 (July 17, 1998) (explaining the responsibility of commanders and other superiors).

23. *See* Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), Judgment, 1986 I.C.J. Rep. 14, ¶115 (June 27) (explaining the effective control doctrine); Prosecutor v. Tadic, Case No. IT-94-1-T, Opinion and Judgment, ¶584 (Int'l Crim. Trib. for the Former Yugoslavia May 7, 1997) (explaining the overall control doctrine).

24. H.R. REP. NO. 73-1383, at 26 (1934) ("When reference is made to 'control,' the term is intended to include actual control as well as what has been called legally enforceable control . . . . It was thought undesirable to attempt to define the term. It would be difficult if not impossible to enumerate or to anticipate the many ways in which actual control may be exerted. A few examples of the methods used are stock ownership, lease, contract, and agency. It is well known that actual control sometimes may be exerted through ownership of much less than a majority of the stock of a corporation either by the ownership of such stock alone or through such ownership in combination with other factors."); *see also Laperriere v. Vesta Ins. Group, Inc.*, 526 F.3d 715, 722–23 (11th Cir. 2008) (per curiam) ("Congress recognized that it would be difficult, if not impossible, to enumerate or anticipate the many ways in which actual control may be exercised and expressly declined to define the term 'control,' leaving courts free to decide issues of control status on a case by case basis.").

25. *See* H.R. REP. NO. 73-1383, at 26 (stating that defining the term "control" would be difficult, if not impossible, because the term can be enumerated or anticipated in many ways).

26. *Killer Robots: UK Government Policy on Fully Autonomous Weapons*, ARTICLE 36, 3–4 (Apr. 2013), http://www.article36.org/wp-content/uploads/2013/04/Policy_Paper1.pdf.

27. UNIDIR, *The Weaponization of Increasingly Autonomous Technologies: Considering How Meaningful Human Control Might Move the Discussion Forward*, No. 2 UNIDIR Resources, 1–2 (2014), http://www.unidir.org/files/publications/pdfs/considering-how-meaningful-human-control-might-move-the-discussion-forward-en-615.pdf.

28. *See* Smith, *supra* note 4, *Automated Vehicles Are Probably Legal in the United States*, at 438 (noting that the term "control" arguably has no original meaning because it can be defined

some notion of control invites, even if unintentionally or unknowingly, myriad popular, technical, and legal views of the concept. As the next section argues, one of these views—the primacy of human authority—may be especially problematic.

RESTRICTIONS ON HUMAN AUTHORITY

Other scholars have recognized that machines can limit as well as initiate killing.[29] Smart locks for guns and launch codes for nuclear missiles, for example, restrict the access of some—indeed, most—individuals to particular weapons.[30] Landmines that neutralize themselves and cluster munitions that explode only under particular conditions may at least be preferable to those that do not. Geofencing might be used to prevent the targeting of hospitals or friendly units.[31] I cannot conceive of all the ways in which technologies might someday check the

---

or understood in so many different ways).

29. *See* Kenneth Anderson & Matthew C. Waxman, *Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can*, *in* JEAN PERKINS TASK FORCE ON NATIONAL SECURITY AND LAW ESSAY SERIES, 15 (The Hoover Institution, Apr. 10, 2013), https://www.law.upenn.edu/live/files/3963-anderson-k-waxman-mlaw-and-ethics-for-autonomouspd ("We should not rule out in advance possibilities of positive technological outcomes—including the development of technologies of war that might reduce risks to civilians by making targeting more precise and firing decisions more controlled (especially compared to human-soldier failings that are so often exacerbated by fear, panic, vengeance, or other emotions—not to mention the limits of human senses and cognition). It may well be, for instance, that weapons systems with greater and greater levels of automation can—in some battlefield contexts, and perhaps more and more over time—reduce misidentification of military targets, better detect or calculate possible collateral damage, or allow for using smaller quanta of force compared to human decision-making."); Noel Sharkey, *The Human Control of Weapons: A Humanitarian Perspective*, 4 (2013), https://www.law.upenn.edu/live/files/3948-sharkey—-human-control-of-weapons-pf-draftpdf ("It is vital that we take the opportunity to use the evolution of technology to ensure that the partnership between human and machine increases rather than diminishes the ability of humans to ensure the legitimacy of the targets of attack. Rather than making more and more hi-tech weapons with the aim of more effective killing and destruction of targets, would it not be better to create hi-tech weapons with the aim of having greater humanitarian impact?"); Peter Asaro, *On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making*, 94 INT'L REV. OF THE RED CROSS 687, 701–02 (2012) ("If technologies did exist that could distinguish civilians from combatants better than any human, or better than the average combatant, then those technologies should be deployed in a manner to assist the combatant in applying the principle of distinction, rather than used to eliminate human judgement. Similarly, if a technology were capable of determining a course of action which could achieve a military objective with minimal collateral damage, and minimize any disproportionate harms, then that technology could be employed by a human combatant charged with the duty of making an informed choice to initiate the use of lethal force in that situation.").

30. *See* Margot Hirsch, *These Are Weapons That Could Prevent Gun Violence*, THE WASH. POST (Dec. 11, 2015), https://www.washingtonpost.com/news/in-theory/wp/2015/12/11/these-are-the-guns-that-could-prevent-gun-violence/ (discussing "smart guns" as a means of curbing certain types of gun violence); David Szondy, *Nuclear Weapons Write Their Own Security Codes*, GIZMAG, (Nov. 24, 2014), http://www.gizmag.com/llnl-nuclear-code/34864/ (discussing a system involving random number generation as a safeguard against nuclear weapons tampering).

31. *See* Kevin Poulsen, *Why the US Government Is Terrified of Hobbyist Drones*, WIRED (Feb. 5, 2015), http://www.wired.com/2015/02/white-house-drone/ (defining GPS Geofencing as the creation of a virtual no-fly zone that prevents drone flight).

misjudgments or misadventures of combatants.

These applications, whatever they are, could diminish the "control" that some or even all humans have over the weapons entrusted to them. Again, this conclusion depends on the meaning given to the term: Control in the sense of commandeering implies a total authority that control in the sense of checking does not. "Meaningful human control" appears to fall somewhere in between these two extremes: A human combatant must do more to initiate a lethal attack than merely fail to stop it, and yet that decision to attack need not be absolute. In other words, an informed human decision to kill is a necessary—but not necessarily a sufficient—condition for killing. In theory, both a human and an automated system could hold a veto over a lethal attack.

"Meaningful human control," however, at most captures only one side of this potential "partnership between human and machine."[32] While resisting the myth of technological infallibility, the concept entrenches the myth of individual infallibility. Driving, which is a leading cause of death,[33] illustrates the tendency to privilege the autonomy of a particular actor over the safety of others. Motor vehicles are generally not equipped with speed limiters, seatbelt ignition locks, or alcohol ignition locks, each of which could substantially improve road safety even as their rare failure results in even rarer tragedy.[34]

This preference for human discretion has also influenced military decisions.[35] According to Jeffrey Lewis of the Monterey Institute for International Studies, during the early Cold War "the United States Air Force, particularly Strategic Air Command, generally resisted the introduction of technical safeguards out of concerns that such measures might make it more difficult to use the weapons in the event of a conflict."[36] Strategic Air Command's putative decision to originally set a launch code of "00000000" for its missiles would, if true, provide the cardinal

---

32. Sharkey, *The Human Control of Weapons: A Humanitarian Perspective*, *supra* note 29, at 4.

33. *See, e.g.*, National Center for Injury Prevention and Control, *10 Leading Causes of Death by Age Group, United States—2013*, CENTERS FOR DISEASE CONTROL AND PREVENTION (2013), http://www.cdc.gov/injury/wisqars/pdf/leading_causes_of_death_by_age_group_2013-a.pdf; *see also* National Center for Injury Prevention and Control, *10 Leading Causes of Injury Deaths by Age Group Highlighting Unintentional Injury Deaths, United States—2013*, CENTERS FOR DISEASE CONTROL AND PREVENTION (2013), http://www.cdc.gov/injury/images/lc-charts/leading_causes_of_injury_deaths_highlighting_unintentional_injury_2013-a.gif (highlighting the frequency of driving-related deaths in the United States).

34. *Cf.* R. v. Michaud, 2015 ONCA 585 (Can.), http://www.ontariocourts.ca/decisions/2015/2015ONCA0585.htm (accepting the plaintiff's argument that an Ontario law requiring a speed limiter on his truck "leaves him in physical danger in some situations" but nonetheless upholding the requirement as reasonable).

35. *See* Noel Sharkey, *The Automation and Proliferation of Military Drones and the Protection of Civilians*, 3 L. INNOVATION & TECH. 229, 237 (2011) (explaining the concerns regarding the lack of compliance of autonomous weapons under international law).

36. Dan Lamothe, *Air Force Swears: Our Nuke Launch Code Was Never '00000000'*, FOREIGN POLICY (Jan. 21, 2014, 1:34 AM), http://foreignpolicy.com/2014/01/21/air-force-swears-our-nuke-launch-code-was-never-00000000/.

example of this preference.[37]

At some point, automated systems may perform better than humans in making real-time distinctions between civilians and combatants or between friendly and enemy combatants.[38] Such a development could bolster the argument for AWSs. But it could also lead to the argument, perhaps grounded in the IHL principle of distinction, that individual combatants should not have unfettered discretion over the use of their weapons. As a practical matter, this would be a difficult argument: A restrictive system that fails to identify a threat (and thus generates a false negative) could, in an extreme case, leave a combatant without the means to defend herself or others from imminent danger.[39]

Broad endorsement of the importance of human "control" could make this argument for restrictive technologies even more difficult. If humans are in the best position to make real-time decisions about the use of lethal force, then a system that restricts that discretion might be as suspect as one that displaces that discretion. A principle based in part on a fear of flawed robots may distract from the demonstrated possibility of flawed humans.

A more holistic approach to the use of lethal force should, on the assumption that both human and machine are flawed, consider the roles of each vis-à-vis each other.[40] Such an approach would explicitly treat humans as elements of a system rather than as users of that system. It would appropriately focus more attention on actors other than the immediate initiator of a lethal attack, including the designers of the human-machine system itself.

These issues are not new: Militaries are already complex systems of autonomous agents—generally human—who act and interact in ways far more nuanced than chains of command (or the lack thereof) may suggest.[41] This focus may in turn demand a more robust conception of command responsibility and the extension of that responsibility to cases of machine action. Although command responsibility is beyond the scope of this article, its potentially vexing distinction between act and omission is not.

---

37. *See id*. (citing Bruce G. Blair, *Keeping Presidents in the Nuclear Dark*, CTR. FOR DEF. INFO.:BRUCE BLAIR'S NUCLEAR COLUMN (Feb. 11, 2004), https://web.archive.org/web/20120511191600/http:/www.cdi.org/blair/permissive-action-links.cfm (discussing the concern that too many safeguards would interfere with a necessary launch than the concern of accidental launches)).

38. *See* Sharkey, *The Automation and Proliferation of Military Drones and the Protection of Civilians*, *supra* note 35, at 237 (explaining the advancements in automated systems needed to distinguish civilians and combatants, such as through more accurate and discriminative sensing and visions systems).

39. In contrast, false positives—errantly identifying a threat—pose the greatest challenge for AWSs. *See* John P. Sullins, *An Ethical Analysis of the Case for Robotic Weapons Arms Control*, NATO CCD COE (2013), https://ccdcoe.org/cycon/2013/proceedings/d2r1s9_sullins.pdf (discussing the possibility of a false positive error of indiscriminant use of force by a robot).

40. Sharkey, *The Human Control of Weapons: A Humanitarian Perspective*, *supra* note 29, at 6.

41. *Id*. at 6–7 (explaining automatic and controlled processes in human psychology).

LETHAL ACTION VERSUS INACTION

The preceding discussion assumed a bright line between automated systems that initiate lethal attacks and those that restrict such attacks. Under this assumption, the chief objection to "meaningful human control" was that it risked, however subtly, precluding the latter in order to prohibit the former. This section introduces a different complication: An automated system that selectively restricts an attack initiated under the "meaningful control" of a human, in some situations, is functionally equivalent to an automated system that simply initiates the attack directly.

Cluster munitions provide a contemporary example. A lethal attack is at least arguably initiated when the cluster munition is dropped or launched under "meaningful human control" as originally defined.[42] At some point after this initiation, a basic munition would split into submunitions, each of which might explode anytime, anywhere, and on anything or anyone.[43] In contrast, an advanced munition would split into advanced submunitions, each of which would be "designed to detect and engage a single target object," "equipped with an electronic self-destruction mechanism," and "equipped with an electronic self-deactivating feature."[44]

If they function properly, these advanced submunitions would automatically restrict the attack in both duration and target.[45] This restriction would be absolute: Even if the pilot wanted the submunitions to deploy more indiscriminately, she would be unable to direct them to do so. The restriction would not thwart "meaningful human control" as long as that term is interpreted to permit automated systems that limit rather than initiate lethal attacks.[46] And it would seem to be desirable; indeed, such designs are intended to "avoid indiscriminate area effects and the risks posed by unexploded submunitions."[47]

Further refinements of these weapons may be even more desirable. For example, a submunition might also be designed to automatically deactivate if it

---

42. *See* Convention on Cluster Munitions art. 2.2, May 30, 2008, 39 U.N.T.S. 2688 (prohibiting parties from using, developing, or transferring certain cluster munitions). The Convention has been ratified by nearly 100 states —but not by the United States, China, or Russia; for a list of all signatories, see Convention on Cluster Munitions, Convention Status, http://www.clusterconvention.org/the-convention/convention-status (last visited Feb. 23, 2016).

43. *Id*. at art. 2.2 ("'Cluster munition' means a conventional munition that is designed to disperse or release explosive submunitions each weighing less than 20 kilograms, and includes those explosive submunitions.").

44. *Id*. (providing characteristics of more advanced submunitions).

45. *See* Tyler Rogoway, *How Dumb Cluster Bombs Are Becoming Heinously Smart*, FOXTROT ALPHA (Dec. 29, 2014), http://foxtrotalpha.jalopnik.com/how-dumb-cluster-bombs-got-heinously-smart-1673486769 ("The self destruct and self inert capabilities of emerging submunitions . . . alleviates, at least to [a] small degree, concerns about the unintended effects of using cluster munitions . . . .").

46. *See* Horowitz & Sharre, *supra* note 2, at 14.

47. Convention on Cluster Munitions, *supra* note 42, at art. 2.2; *Cf., e.g.*, U.S. Department of State, *Cluster Munitions*, http://www.state.gov/t/pm/wra/c25930.htm (last visited Feb. 4, 2016) ("[C]luster munitions can often result in much less collateral damage than unitary weapons, such as a larger bomb or larger artillery shell would cause, if used for the same mission.").

detects a person of small stature or if it lands near a structure coded on a map as a hospital, or it might be designed to explode only if a vehicle with certain characteristics passes by it. Each of these restrictions would presumably represent a laudable improvement over a basic cluster munition. And, as restrictions, they would seem to remain consistent with a principle of "meaningful human control."

This combination of human initiation plus automated limitation of an attack, however, may be difficult to distinguish in practical terms from the automated initiation of distinct attacks. In the former case, a human combatant might deploy a weapon against a broad target on the assumption that the weapon itself would selectively exclude particular elements of that target from lethal effect.[48] In the latter case, a human combatant might activate a weapon that automatically determines what targets to include rather than to exclude.[49] Indeed, Israel's Harpy—"an anti-radar weapon that flies a search pattern over a wide area searching for [and then dive-bombing into] enemy radars"[50]—could arguably constitute either an automated attack or the automated restriction of a human attack.

These categories could overlap in another way as well: Human combatants might come to rely on automated systems that limit attacks just as they may come to rely on automated systems that recommend targets. A human who expects that an automated system will restrict an attack to lawful targets may be more willing to initiate such an attack.[51] Even if the automated system succeeds, this reliance would implicitly delegate more authority to the automated system.

Similar issues regarding human factors arise in the "mushy middle" of driving automation, where driving tasks are shared sequentially or simultaneously by the human driver and the automated driving system.[52] One response to the problem of undue reliance by the human has been to design the vehicle to monitor whether the driver actually remains alert.[53] In this way, both human and machine supervise the other.[54] Similarly, advanced weapon systems might incorporate some form of dual supervision in which machine "control" is as important as human "control."

---

48. *See* Jack M. Beard, *Autonomous Weapons and Human Responsibilities*, 45 GEO. J. INT'L. L. 617, 652–53 (2014) (discussing automated weapons that are programmed to strike particular targets while excluding other targets).

49. *Cf*. U.S. Department of Defense Directive 3000.09 (Nov. 21, 2012) http://www.dtic.mil /whs/directives/corres/pdf/300009p.pdf (defining "autonomous weapon system" as one that can selectively engage targets without human input beyond initial activation).

50. Beard, *supra* note 48, at 633.

51. *Id*. at 653 (discussing how a reasonable person would operate an automated weapon).

52. Bryant Walker Smith, *A legal Perspective on Three Misconceptions in Vehicle Automation*, *in* ROAD VEHICLE AUTOMATION 85, 86 (Gereon Meyer & Sven Beiker eds. 2014).

53. Bryant Walker Smith, *Tesla and Liability*, THE CTR. FOR INTERNET AND SOC'Y: BLOG (May 20, 2015, 9:42 AM), https://cyberlaw.stanford.edu/blog/2015/05/tesla-and-liability.

54. *See* Smith, *Lawyers and Engineers Should Speak the Same Robot Language*, *supra* note 6, at 96 (discussing shared authority between the automated vehicle and the human driver).

CONCLUSION

Control can be obviously confusing and—more alarmingly—subtly confusing. It means so much that, without careful and persistent qualification, it ends up meaning very little. For this reason:

> [C]ontrol is more useful as a structure than as a standalone term. Those who would deploy it should first describe the control system they actually intend: the goals, inputs, processes, and outputs to the extent they are determined by a human designer and the authority of the human or computer agents to the extent they are not.[55]

"Meaningful human control" hints at, but ultimately risks missing, much of this system complexity. A human combatant is not merely the user of an automated weapon system that she checks or commandeers. Rather, human and machine both empower and limit the other. Even this duality obfuscates the wide range of elements in a system designed to effectively deploy lawful force. A military, after all, is much more than a soldier and a gun.

The fundamental functional question is whether such a system—a modern military committed to international humanitarian law—can remain robust when human or machine elements fail. Whether any given element is human or machine is a secondary question. So too is what particular relationship exists between any two elements. Asking who controls what (or what controls whom) within the system does not necessarily reveal whether the system is itself under control.

55. *Id*. at 85.