AUTONOMOUS WEAPONS AND INTERNATIONAL HUMANITARIAN LAW

OR

KILLER ROBOTS ARE HERE. GET USED TO IT.

*Shane Harris**

"In the game of life and evolution there are three players at the table: human beings, nature, and machines. I am firmly on the side of nature. But nature, I suspect, is on the side of the machines."[1] To begin our discussion, I propose the following: (1) It is inevitable that human beings will build weapons systems capable of killing people on their own, without any human involvement or direction; and (2) It is conceivable that human beings could teach machines to recognize and distinguish when the use of lethal force complies with international humanitarian law.

Why do I think this? Consider history. In World War II, it took a fleet of 1,000 B-17 bombers, flown, navigated, and manned by a crew of 10,000 men, to destroy one Axis ground target.[2] American bombs were so imprecise that, on average, only 1 in 5 fell within 1,000 feet of where they were aimed.[3] Aerial bombing was a clumsy affair, utterly dependent on the extraordinary labor of human beings, and lots of them.[4]

Just one generation later, that was no longer true.[5] By 1972, the Vietnam War

---

*Portions of this article are adapted from Shane Harris, "Out of the Loop," *Hoover Institution*. Author, *@War: The Rise of the Military-Internet Complex* and *The Watchers: The Rise of America's Surveillance State*, Senior Intelligence and National Security Correspondent, *The Daily Beast* Fellow, International Security Program, New America.

1. GEORGE DYSON, DARWIN AMONG THE MACHINES, ix (1997).

2. *See Life and Death Aboard a B-17, 1944*, EYEWITNESS TO HIST., http://www.eyewitnesstohistory.com/b17.htm (last visited Feb. 18, 2016) (describing the use of B-17 bombers in World War II, where each mission saw as many as 1,000 B-17 bombers, each manned by a crew of ten, dispatched to neutralize each targeted objective, which might have included buildings, military installations, and the like).

3. Wolfgang W. E. Samuel, *Tactical PGMs: Implications in Perspective*, PARAMETERS J. US ARMY WAR C., 32, 35 (1979), http://strategicstudiesinstitute.army.mil/pubs/parameters/Articles/1979/1979%20samuel.pdf (last visited Feb. 20, 2016).

4. *See generally* Jonathan Yardley, Opinion, *'The Bombers and the Bombed Allied Air War Over Europe, 1940-1945' by Richard Overy*, WASH. POST (Mar. 7, 2014), https://www.washingtonpost.com/opinions/the-bombers-and-the-bombed-allied-air-war-over-europe-1940-1945-by-richard-overy/2014/03/07/7c2ba5de-9d60-11e3-a050-dc3322a94fa7_story.html.

5. *See* Nathan A. Canestaro, *Legal and Policy Constraints on the Conduct of Aerial Precision Warfare*, 37 VAND. J. TRANSNAT'L L. 431, 448–49 (2004) (discussing new technological developments in weaponry employed in the Vietnam War twenty years after the

drastically reduced manpower but increased accuracy.[6] It took only 16 F-4 fighter-bombers, each flown and navigated by only two men, to destroy an enemy bridge with only 26 tons of bombs.[7] This was also the first widespread use of laser-guided munitions, enhancing the precision of the attack.[8]

After the Vietnam War, the human necessity for aerial warfare became more attenuated and less relevant.[9] In the Gulf War, one pilot could hit two targets.[10] The effectiveness of the human-machine pairing was breathtaking. A single "smart bomb" could do the work of 1,000, planes dropping more than 9,000 bombs in World War II.[11] By the time the United States went to war in Afghanistan and Iraq, one pilot in one plane could destroy six targets.[12] The military's weapons were guided by global positioning satellites orbiting thousands of miles above the surface of the earth.[13] And increasingly, the pilots were not actually inside their planes anymore.[14]

The historical trend is sobering. As aircraft and weapons have become more precise, human beings have become less essential to military wartime operations. And that may suit the military just fine. Seven years ago, the United States Air Force released its big-picture forecast for how the service will fight wars by the year 2047.[15] Humans are to remain "in the loop" on strike missions—that is, they will still fly airplanes.[16] But by 2020, the plan is to have one pilot "control" four

---

end of WWII).

6. *Id*. at 448.

7. *Id*. Destroying the same target, the Paul Doumer Bridge in Hanoi, had required 113 sorties by F-105 fighter-bombers and 380 tons of bombs just five years earlier. *Id*.

8. *Id*.

9. *See id*. at 450 (explaining the development and acquisition of precision-guided munitions that led to the minimization of human welfare costs to civilians).

10. *See, e.g*., Malcolm W. Browne, *Invention That Shaped the Gulf War: The Laser-Guided Bomb*, N.Y. TIMES (Feb. 26, 1991), http://www.nytimes.com/1991/02/26/science/invention-that-shaped-the-gulf-war-the-laser-guided-bomb.html?pagewanted=all.

11. *See* U.S. AIR FORCE, REACHING GLOBALLY, REACHING POWERFULLY: THE UNITED STATES AIR FORCE IN THE GULF WAR 21 (1991) (describing how a single F-117 stealth bomber destroyed Iraqi air force headquarters by dropping a single smart bomb down an elevator shaft). A smart bomb is a "precision-guided munition" designed to have greater accuracy through the use of GPS devices attached to the bomb. W.J. Hennigan, *Armed with GPS Kits, U.S. Crew Puts 'Smart' in Smart Bombs*, L.A. TIMES (Mar. 20, 2015), http://www.latimes.com/world/middleeast/la-fg-smart-bombs-20150320-story.html.

12. *See Six Targets One Blow: Cutting-Edge Interceptor Joins Russian Air Force*, SPUTNIK NEWS (Nov. 27, 2015), http://sputniknews.com/military/20151127/1030878163/russia-air-force-interceptor-jet-combat.html (discussing the capability of certain aircraft to destroy six targets at a time).

13. Canestaro, *supra* note 5, at 450.

14. *Id*. at 452.

15. *See generally* U.S. AIR FORCE, UNMANNED AIRCRAFT SYSTEMS FLIGHT PLAN 2009-2047 (2009), https://fas.org/irp/program/collect/uas_2009.pdf.

16. *See* Joe Pappalardo, *The Future for UAVs in the U.S. Air Force*, POPULAR MECHANICS (Feb. 25, 2010), http://www.popularmechanics.com/military/a5383/4347306/ (noting that humans remain in control either from the sky or the ground during strike missions).

aircraft.[17] That is, he or she will not sit in the cockpit, or even a mocked-up cockpit in a trailer thousands of miles away.[18] The pilot will communicate with the fleet via computer terminal, and perhaps even an application on a smartphone.[19]

The efficiencies that the United States Air Force could wring out of this pilot-as-controller setup are, like the advances of an earlier area, breathtaking.[20] The pilot will issue a flight plan, and the aircraft themselves will complete many key aspects of the mission unassisted.[21] They will take off, fly to the target, and avoid detection by adversaries.[22] The United States Air Force's ultimate goal is for one human to control a fleet of drones that can attack multiple targets with as close to perfect precision as possible.[23]

At this point, one would have to stretch the English language as much as imagination allows to say that human pilots are still "in the loop." You might say they are "on the loop," or clinging to it. But this is warfare being conducted by machines largely unassisted by humans. And that is by design. One person could not keep track of all the aircraft needed to conduct multiple airstrikes.[24] This progression towards something that looks like "autonomy" in flight begs the question: When, if ever, would the military allow drones to decide when to attack?

***

There are two ways to answer that question. The first is to examine how close the military is to developing or acquiring the technology to allow drones to attack on their own—that is, without a human giving them the order to fire. To some extent, we are already there—and have been for some time.

---

17. *See id*. (stating that by 2020 humans will be able to control multiple unmanned aerial vehicles (UAV) at one time).

18. *See id*. (asserting that a pilot miles from danger can assess a situation from the screen of his cockpit).

19. *See* Courtney Howard, *UAV Command, Control & Communications*, MIL. & AEROSPACE ELECTRONICS (July 11, 2013), http://www.militaryaerospace.com/articles/print/volume-24/issue-7/special-report/uav-command-control-communications.html (discussing how DreamHammer's Ballista software allows unmanned drones to be run from computers, tablets, or smartphones).

20. *See* Pappalardo, *supra* note 16 (describing the many advantages that this technology creates).

21. *See id*. (detailing the Air Force Flight Plan stating that UAVs will have collision-avoidance systems as well as be able to refuel each other).

22. *See id*. (describing an example mission where a pilot miles away from the target assesses the situation, confirms the target, and authorizes the UAV to fire).

23. *See* P.W. SINGER, WIRED FOR WAR: THE ROBOTICS REVOLUTION AND CONFLICT IN THE 21ST CENTURY 126 (2009) ("For example, the army's FCS plan is to have two humans sit at identical consoles and jointly supervise a team of ten land robots."); *see also* Sumit Passary, *US Air Force Hires Civilian Drone Operators to Control Surveillance Drones*, TECH TIMES (Nov. 30, 2015), http://www.techtimes.com/articles/111893/20151130/us-air-force-hires-civilian-drone-operators-to-control-surveillance-drones.htm ("[The] Pentagon's goal is to recruit more than 1,200 drone operators but it is well below the recruitment requirement.").

24. *See* SINGER, *supra* note 23, at 126 (citing a NATO study that concluded that having one operator control two rather than one, UAV at a time reduces their performance levels by an average of fifty percent).

In 1988, the U.S.S. *Vincennes,* a guided missile cruiser on patrol in the Persian Gulf, shot down an Iranian passenger jet after the ship's Aegis targeting system mistook it for a military fighter.[25] The crew of the *Vincennes* could tell from the plane's course, speed, and radio signal that it was a civilian aircraft.[26] But Aegis, which had been programmed to identify large Soviet bombers, believed otherwise.[27] All 290 passengers and crew aboard the Iranian jet died, including 66 children.[28]

"Even though the hard data was telling the crew that the plane wasn't a fighter jet, they trusted what the computer was telling them more," P.W. Singer, Director of the 21st Century Defense Initiative at the Brookings Institution, writes in his book *Wired for War*.[29] "Aegis was on semiautomatic mode, but not one of the eighteen sailors and officers on the command crew was willing to challenge the computer's wisdom. They authorized it to fire."[30]

Aegis was not "autonomous." Human beings were still "in the loop" to a degree, and could overrule the computer's conclusions on whether to fire.[31] But Singer notes that the Navy had such faith in Aegis' abilities to identify a true enemy that the *Vincennes* was allowed to fire of its own volition, without the crew seeking permission from more senior officers in the fleet.[32] That is not exactly putting the computer in charge, but it delegates an extraordinary amount of decision-making normally conducted by humans—trained in the law of war—to a machine.

Singer writes that, historically, "[t]here have been all sorts of new technologies that people insisted in absolutist terms would 'never ever' be allowed to run on their own without a human in the loop.[33] Then, as the human roles were redefined, they were gradually accepted, and eventually were not even thought about."[34] For instance, even the 1988 Iranian airliner accident discussed above, where a targeting system miscalculated and caused the death of 290 passengers and crew, did not deter military research into more automated weapons systems.[35]

---

25. *See id*. at 124–25.

26. *See id*. at 125 (explaining that the crew of the U.S.S. *Vincennes* trusted the Aegis system more than the hard data before them).

27. *See id*. (describing how the Aegis system erroneously registered the passenger plan as an Iranian F-14 fighter, a plane half the size, thereby deeming it an enemy target).

28. *Id*.

29. *Id*.

30. SINGER, *supra* note 23, at 125.

31. *See id*. at 124 (describing the operation of the Aegis computer system and that the human sailors could override the Aegis computer in any of its modes).

32. *See id*. at 124–25 ("Even though the hard data was telling the crew that the plane wasn't a fighter jet, they trusted what the computer was telling them more. . . . They authorized [Robo-cruiser] to fire . . . the only ship in the area authorized to fire without having to seek permission from more senior officers in the fleet.").

33. *Id*. at 125.

34. *Id*.

35. *See, e.g*., *id*. at 59 (referencing the Senate Armed Services Committee meeting in 2000, where it was mandated that by 2010, one-third of all aircrafts that go beyond enemy lines should be unmanned and by 2015 one-third of all ground combat vehicles should be driverless).

The Defense Advanced Research Projects Agency (DARPA), the Pentagon research arm that first developed stealth aircraft technology,[36] has joined up with the United States Air Force to study ways to give drones autonomous control over their weapons.[37] The Persistent Close Air Support (PCAS) Program is ostensibly aimed at speeding up the process by which tactical air controllers can call in strikes, either to piloted or unmanned aircraft.[38] It takes about an hour now, and researchers want to whittle that down to six minutes.[39] To do that, the program will build equipment that lets unmanned aircraft respond autonomously to a request for weapons fire from the controllers.[40] It will be up to the drones to figure out how best to attack the target.[41]

DARPA is also running a program to take automation in commercial aviation to new heights.[42] Recently, the agency has awarded three contracts for its Aircrew Labor In-Cockpit Automation System (ALIAS).[43]As DARPA describes the

---

36. *See DARPA's Stealth Revolution*, DARPA, http://www.darpa.mil/about-us/timeline/darpas-stealth-revolution (last visited Mar. 3, 2016) ("In the mid-1970s, DARPA oversaw the development of HAVE Blue, the first practical combat stealth aircraft, which made its first test flight by the end of 1977. This led to the procurement by the Air Force of the F-117A stealth fighter, which became operational in October 1983.").

37. *See* John Keller, *Northrup Grumman Joins DARPA/Air Force Persistent Close Air Support Program to Safeguard Ground Forces*, MIL. & AEROSPACE ELECTRONICS (May 1, 2011), http://www.militaryaerospace.com/articles/print/volume-22/issue-5/news/news/northrop-grumman-joins-darpa-air-force-persistent-close-air-support-program-to-safeguard-ground-forces.html (indicating the program would allow human attack controllers to enable strike aircrafts to respond autonomously to weapon delivery requests).

38. *See id*. ("Close air support manned and unmanned aircraft may be able to speed their services to ground troops faster than ever before with technology from the PCAS research program.").

39. Daniel Patt, *Persistent Close Air Support (PCAS)*, DARPA, http://www.darpa.mil/program/persistent-close-air-support (last visited Mar. 3, 2016) (noting the PCAS program seeks to reduce the time between the call for a strike and a target being hit from sixty to six minutes). Dr. Daniel Patt of DARPA further explains the Close Air Support system currently in use:

> To maintain a decisive tactical advantage in 21st-century combat, warfighters need the ability to safely, rapidly and collaboratively deploy ordnance against elusive mobile targets. Unfortunately, air-ground fire coordination—referred to as Close Air Support or CAS—has changed little since its emergence in World War I. Pilots and dismounted ground agents can focus on only one target at a time and must ensure they hit it using just voice directions and, if they're lucky, a common paper map. It can take up to an hour to confer, get in position and strike—time in which targets can attack first or move out of reach.

Id.

40. *See* Keller, *supra* note 37 (discussing how the PCAS research program may be able to speed up services to ground troops).

41. *See id*. (indicating aircraft would be able to respond autonomously to weapons delivery requests).

42. *See* Graham Warwick, *Darpa's Alias Aims to Automate Existing Aircraft*, AVIATION WEEK (Apr. 21, 2014), http://aviationweek.com/defense/darpas-alias-aims-automate-existing-aircraft (indicating DARPA seeks to further automate manned aircraft to reduce pilot workload, increase flight capabilities, and increase safety).

43. *ALIAS Prepares for Accelerated Take-Off*, DARPA (Mar. 10, 2015), http://www.darpa.

program:

> ALIAS envisions a tailorable, drop-in, removable kit that would enable high levels of automation in existing aircraft and facilitate reduced need for onboard crew. The program intends to leverage the considerable advances that have been made in aircraft automation systems over the past 50 years, as well as the advances that have been made in remotely piloted aircraft technologies, to help shift and refocus pilot workloads, augment mission performance and improve aircraft safety.[44]

That sure sounds like a plan to replace pilots with machines, or at least to fundamentally change what it means for a human to "fly" an aircraft.

***

So, if the technology to build autonomous weapons systems (AWS) is, essentially already with us and only going to improve, what would the legal constraints be on the use of such weapons? Or, more to the point, could the military build an AWS capable of determining whether an attack is lawful, just as the military trains soldiers, sailors, airmen, and Marines to do now?

As we saw earlier, pilots are going to become controllers overseeing fleets of drones that fly largely on their own. In theory, you could have one human controller orchestrating all drones, known as a fire controller, but a roomful of lawyers or other decision-makers monitoring one or a few drones and deciding when to let each drone attack its target. At that point, the fire controller would send a command to the drone, which decides the optimal moment to strike. The drone camera could take into account the number of civilians nearby, the blast radius of the missile—all factors that human beings consider now, with the aid of technology to help them decide when a strike is legal.

We can imagine, then, programming a drone to tell a human being when conditions for a lawful strike have been seemingly met, and then presenting the fire controller with the option to fire. Lawyers are already preparing for the emergence of such autonomous, or semi-autonomous, killing machines.[45]

"Advances in artificial intelligence (AI) and other technologies will soon make possible the development of fully autonomous weapons, which would revolutionize the way wars are fought," Human Rights Watch observed in a lengthy online "Question & Answer."[46]

---

mil/news-events/2015-03-10.

44. *Id*.

45. *See Human Rights Watch Q&A*, HUM. RTS. WATCH (Oct. 21, 2013), https://www.hrw.org/news/2013/10/21/qa-fully-autonomous-weapons (discussing how the U.S. Defense Department has prepared policies for "killer robots").

46. *Id*.

> These weapons, unlike the current generation of armed drones, would be able to select and engage targets without human intervention. Military officials in the United States and other technologically advanced countries generally say that they prefer to see humans retain some level of supervision over decisions to use lethal force, and the U.S. Defense Department has issued a policy directive embracing that principle for the time being.[47]

But what happens when time passes? Human Rights Watch asked whether AWS could ever observe the requirement under international humanitarian law to protect civilians in armed conflict.[48] They were not optimistic.[49] "Distinguishing between combatants, who may be targeted, and civilians, who may not be, is a core requirement of international humanitarian law. There is no certainty that fully autonomous weapons would have the capacity to make such distinctions reliably," Human Rights Watch says.[50] To which, I suspect, many technologists in and out of the military would reply, let us try.

***

So I return to the two ideas posited at the outset: (1) It is inevitable that human beings will build weapons systems capable of killing people on their own, without any human involvement or direction; and (2) It is conceivable that human beings could teach machines to recognize and distinguish when the use of lethal force complies with international humanitarian law. For the sake of argument, say I am right. The question I then pose is: Should we build these autonomous killing machines and use them in war?

I will provoke the discussion by answering, yes. Human beings have proven reliably fallible when it comes to distinguishing between civilians and combatants. We do a pretty good job, but we make a lot of mistakes. Machines could do a better job.

We like to tell ourselves that human beings can make judgments, especially snap judgments, in a way that computers cannot. If science were to prove this to be true, Google and every major automobile manufacturer would not be betting their futures on driverless cars and selling them as a solution to dramatically cut the number of human-caused traffic accidents.[51] We use computers to provide myriad life-saving functions. We resist giving computers the power to take life because we think they cannot do it more safely or humanely than we can. Maybe we are wrong.

---

47. *Id*.

48. *Id*.

49. *See id*. (discussing the implications if one State is allowed to use AWS).

50. *Id*.

51. *See generally* Paul Ingrassia & Joseph White, *Automakers, Not Silicon Valley, Lead in Driverless Car Patents: Study*, REUTERS (Jan. 5, 2016), http://www.reuters.com/article/us-tech-ces-autos-idUSKBN0UJ1UD20160105.