

AUTONOMOUS WEAPONS & HUMAN CONTROL

CENTER FOR A NEW AMERICAN SECURITY
ETHICAL AUTONOMY PROJECT
APRIL 2016



Center for a
New American
Security

Many nations are building increasingly autonomous next-generation robotic & uninhabited systems on the ground, in the air, and at sea. *How much autonomy will they have? How much should they have?*

Human judgment and control in lethal attacks

Increasing autonomy in weapons raises the question of how much human involvement is required in lethal attacks. Some have called for “meaningful human control” over attacks. Others have suggested a standard of “appropriate human judgment.” Regardless of the terms used, many agree that humans should be involved in lethal force decisions at some level.

Automation and autonomy are already used for a wide range of functions in weapons, including searching for and identifying potential targets, tracking targets, cueing them to human operators, timing when to fire, and homing in on targets once launched. Generally, humans decide which specific targets are to be engaged.

Humans currently perform three kinds of roles with respect to target selection and engagement, sometimes performing multiple roles simultaneously.

**Increasing
autonomy
in weapons
raises the
question of
how much
human
involvement
is required
in lethal
attacks.**



The human as essential operator:

The weapon system cannot accurately and effectively complete engagements without the human operator.



The human as moral agent:

The human operator makes value-based judgments about whether the use of force is appropriate—for example, whether the military necessity of destroying a particular target in a particular situation outweighs the potential collateral damage.



The human as fail-safe:

The human operator has the ability to intervene and alter or halt the weapon system's operation if the weapon begins to fail or if circumstances change such that the engagement is no longer appropriate.

As machine intelligence capabilities increase, humans may not be needed as essential operators of future weapon systems. However, humans still may be needed to make value-based moral decisions and/or act as fail-safes in the event that the weapon system fails.

Rather than focus on where humans are needed today based on existing technology, we ought to ask: If technology was sufficiently advanced to make lethal force decisions in a lawful manner, are there tasks for which human judgment still would be required because they involve decisions that only humans should make? Do some decisions require uniquely human judgment? If so, why?



Phalanx MK-15 Close In Weapons System (CIWS) (U.S. Navy)

Defensive Human-Supervised Autonomous Weapons

At least 30 nations currently have defensive human-supervised autonomous weapons to defend against short-warning saturation attacks from incoming missiles and rockets. These are required for circumstances in which the speed of engagements can overwhelm human operators.

Once activated and placed into an autonomous mode, these weapon systems can search for, select, and engage incoming threats on their own without further human intervention. Humans supervise the weapon's operation and can intervene to halt the engagements if necessary. To date, human-supervised autonomous weapons have been used in limited circumstances to defend human-occupied vehicles or bases. Because humans are co-located with the system and have physical access, in principle they could manually disable the system if necessary if it begins malfunctioning.

Why might militaries build autonomous weapons?

A frequent argument for autonomous weapons is that they may behave more lawfully than human soldiers on the battlefield, reducing civilian casualties. It is true that machines' abilities at object recognition are rapidly improving and may soon surpass humans' ability to accurately identify objects. Automated target identification tools may help to more accurately discriminate military from civilian objects on the battlefield, reducing civilian casualties. Adding automation does not necessarily equate to removing human involvement, however. Militaries could design weapons with more advanced automatic target identification and still retain a human in the loop to exercise human judgment over each engagement.

Adding automation does not necessarily equate to removing human involvement.

Autonomous weapons may be desired because of their advantages in speed. Without a human in the loop, machines potentially could engage enemy targets faster. At least 30 nations already use human-supervised autonomous weapons to defend human-occupied vehicles and bases for this reason.

Militaries also could desire autonomous weapons so that uninhabited ("unmanned") vehicles can select and engage targets even if they are out of communications with human controllers. A number of major military powers are building combat drones to operate in contested areas where communications with human controllers may be jammed.

Does automation increase or decrease human control?

Automation changes the nature of human control over a task or a process by delegating the performance of that task to a machine. In situations where the machine can perform the task with greater reliability or precision than a person, this can actually increase human control over the final outcome. For example in a household thermostat, by delegating the task of turning on and off heat and air conditioning, humans improve their control over the outcome: the temperature in the home.

Some tasks lend themselves more easily to automation. Tasks that have an objectively correct or optimal outcome and that occur in controlled or predictable situations may be suitable for automation. Tasks that lack a clear “right” answer, that depend on context, or that occur in uncontrolled and unpredictable environments may not be as suitable for automation.

Some engagement-related tasks in war have a “right” answer. Object identification is a potentially fruitful area for automation. Recent advances in machine intelligence have shown tremendous promise in object recognition.

Some engagement-related tasks may not have an objectively right answer and may depend on context or value-based decisions. Whether a person is a combatant may depend heavily on context and their specific actions. Proportionality considerations may similarly depend on ethical and moral judgment.

Some engagement-related tasks may not have an objectively right answer.

To learn more, visit www.cnas.org/ethicalautonomy

Additional CNAS reports on autonomous weapons

An Introduction to Autonomy in Weapon Systems,
by Paul Scharre and Michael C. Horowitz (February 2015)

Meaningful Human Control in Weapon Systems: A Primer,
by Michael C. Horowitz and Paul Scharre (March 2015)

Autonomous Weapons at the UN: A Primer for Delegates,
by Paul Scharre, Michael C. Horowitz, and Kelley Sayler (April 2015)

Autonomous Weapons and Operational Risk,
by Paul Scharre (February 2016)

ABOUT THE CNAS ETHICAL AUTONOMY PROJECT

The Center for a New American Security's Ethical Autonomy project is made possible through the generous support of the John D. and Catherine T. MacArthur Foundation. CNAS is an independent, nonprofit research organization. CNAS does not take institutional positions.

Defining human control and autonomy's role in weapons



A HUMAN OPERATOR LAUNCHES THE FULLY AUTONOMOUS WEAPON TO ENGAGE POTENTIAL TARGETS OVER A BROAD AREA

WEAPON FLIES IN A SEARCH PATTERN LOOKING FOR TARGETS

FULLY AUTONOMOUS WEAPONS

FULLY AUTONOMOUS WEAPONS

A fully autonomous weapon would operate differently than a semi-autonomous weapon. The weapon still would be programmed by humans and launched by a human. However, the human launching the weapon would not need to launch it at a specific target. The weapon would have more freedom (autonomy) to search for potential targets over a wider area in space and time, select the target(s), and engage them without asking for human permission.

Fully autonomous weapons could be loitering munitions that hunt for their own targets or could be recoverable platforms such as future drones that use lethal force on their own.

Autonomous weapons need not be “learning” or sentient. They are merely weapons that search for, select, and engage targets on their own according to pre-programmed parameters.

A HUMAN OPERATOR LAUNCHES THE SEMI-AUTONOMOUS WEAPON AT A SPECIFIC TARGET. A TARGET COULD CONSIST OF SEVERAL OBJECTS, SUCH AS A COLUMN OF TANKS

ONCE LAUNCHED, SOME WEAPONS CAN BE RECALLED OR RETARGETED IN FLIGHT. OTHERS, SO-CALLED “FIRE AND FORGET” WEAPONS, CANNOT

SEMI-AUTONOMOUS WEAPONS

SEMI-AUTONOMOUS WEAPONS

Semi-autonomous weapons, such as homing missiles and torpedoes, have existed in various forms for over 70 years and are widely used by every modern military. In semi-autonomous weapons, the weapon has some autonomy, but the human operator who launches the weapon still makes a conscious decision to engage and destroy a specific target.

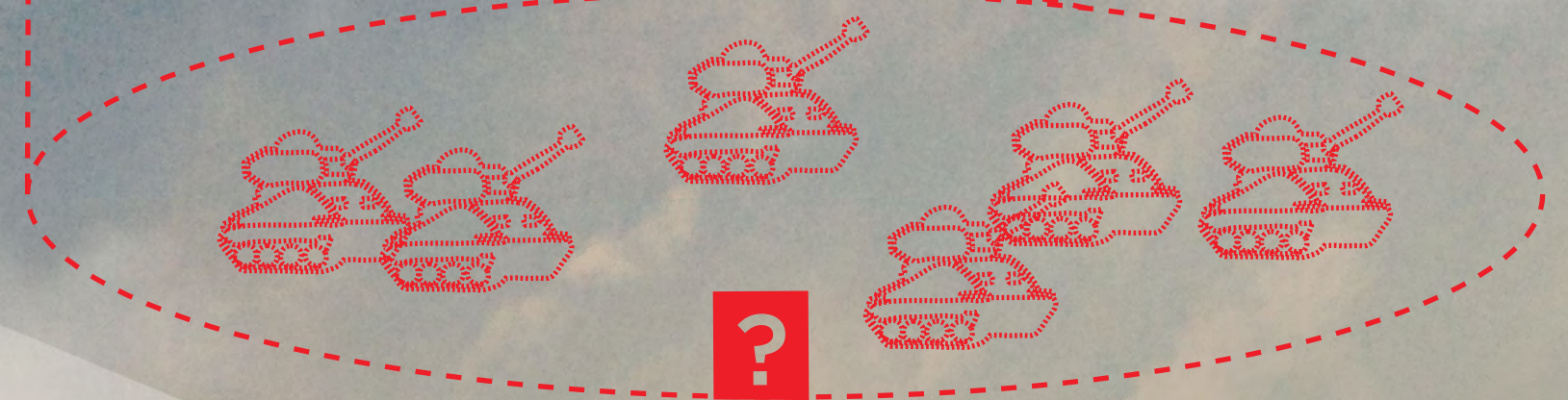
Semi-autonomous weapons could be munitions or could be recoverable platforms such as drones that keep a human “in the loop” on lethal force decisions.

How is human control exercised in an attack with a semi-autonomous weapon?

The human operator may not have control over the weapon once launched. The weapon may have some autonomy. However, the autonomy is limited. The human exercises control over the attack by launching the weapon at a known enemy target at a specific place and time. There is a risk that the weapon homes in on the wrong target. However, this risk is mitigated by limiting the time and space the weapon has to search for the target. Because the weapon's autonomy is limited, the human can ensure that this specific attack – against this enemy target, at this place, at this time – is lawful.



THE WEAPON EMPLOYS VARIOUS HOMING OR GUIDANCE MECHANISMS TO CORRECT FOR AIMING ERRORS IN ORDER TO HIT THE SPECIFIC TARGET THE HUMAN OPERATOR INTENDED



POTENTIAL TARGETS

How is human control exercised in an attack with a fully autonomous weapon?

Humans decide when to launch the weapon, the type of target the weapon is searching for, and the weapon's search area in time and space. However, because the weapon has more freedom (autonomy), human control over the attack is different. For a fully autonomous weapon, the human does not know the specific target that is to be engaged, only the general class of targets over a specified area in space and time.

The human launching the weapon still must ensure that his or her actions are lawful. This means that the human operator must ensure, before launching the weapon, that any potential target that the weapon might select and engage within the search area would be a lawful engagement. Target recognition technology is good enough today to accurately discriminate some types of targets, such as enemy radars. The legal requirements for proportionality and precautions in attack could be met by: only using the weapon in an area devoid of civilians; using a small warhead to minimize potential collateral damage; and/or including sensors on the weapon itself to sense and avoid civilians.