GlobalResearch
Center for Research on Globalization

Worried About the Autonomous Weapons of the Future? Look at What's Already Gone Wrong

By Dr. Ingvild Bode and Dr. Tom Watts
Global Research, April 22, 2021
Bulletin of the Atomic Scientists 21 April 2021
Theme: Militarization and WMD

***

To the casual observer, the words "military AI" have a certain dystopic ring to them, one that's in line with sci-fi movies like "Terminator" that depict artificial intelligence (AI) run amok. And while the "killer robots" cliché does at least provide an entry point into a debate about transformative military technologies, it frames autonomous AI weapons as a challenge for tomorrow, rather than today. But a close look at the history of one common type of weapons package, the air defense systems that militaries employ to defend against missiles and other airborne threats, illuminates how highly automated weaponry is actually a risk the world already faces.

As practical, real-world examples, air defense systems can ground a debate over autonomous weapons that's often abstract and speculative. Heads of state and defense policymakers have made clear their intentions to integrate greater autonomous functionality into weapons (and many other aspects of military operations). And while many policymakers say they want to ensure humans remain in control over lethal force, the example of air defense systems shows that they face large obstacles.

Weapons like the US Army's Patriot missile system, designed to shoot down missiles or planes that threaten protected airspace, include autonomous features that support targeting. These systems now come in many different shapes and sizes and can be typically operated in manual or various automatic modes. In automatic modes, the air defense systems can on their own detect targets and fire on them, relegating human operators to the role of supervising the system's workings and, if necessary, of aborting attacks. The Patriot air defense system, used by 13 countries, is "nearly autonomous, with only the final launch decision requiring human interaction," according to research by the Center for Strategic and International Studies.

Air defense systems have been used by militaries for decades. Researchers began developing some of the first so-called "close-in weapons systems" to provide warships a last line of defense against anti-ship missiles and other high-speed threats in the 1970s. Modernized versions of these systems—including the Phalanx, which entered production in 1978—are still in use on US and allied warships. By one estimate, at least 89 countries operate air defense systems; the weapons have shaped the role of human operators.

Our research on the character of human-machine interaction in air defense systems suggests that over time, their use has incrementally reduced the quality of human oversight in specific targeting decisions. More cognitive functions have been "delegated" to machines, and human operators face incredible difficulties in understanding how the complex computer systems make targeting decisions.

Maintaining appropriate human control over specific targeting decisions is particularly important when thinking about the concept of meaningful human control, which plays a prominent role in the international regulatory discussion on autonomous weapons systems. This is because, as previous research suggests, the brunt of a soldier or the military's obligations under international humanitarian law (such as complying with the principles of distinction, proportionality and precaution enshrined in the Geneva Conventions) apply to specific, battlefield decisions on the use of force, rather than to the development and testing of weapons systems.

A study of air defense systems reveals three real-world challenges to human-machine interaction that automated and autonomous features have already created. These problems are likely to grow worse as militaries incorporate more AI into the high-tech weapons of tomorrow.

Targeting decisions are opaque.

The people who operate air defense systems already have trouble understanding how the automated and autonomous features on the weapons they control make decisions, including how the systems generate target profiles and assessments. In part, that's due to the sheer complexity of the systems' internal workings; how many people understand the algorithms behind the software they use, after all? But high-profile failures of air defense systems also suggest that human operators are not always aware of known system weaknesses.

The history of Patriot systems operated by the US Army, for instance, includes several near-miss so-called "friendly fire" engagements during the First Gulf War in the 1990s and in training exercises. But as John Hawley, an engineering expert working on automation in air defense systems, argued in a 2017 report, the US Army was so convinced of the Patriot system's successes that they did not want to hear words of caution about using the system in automatic mode. Rather than addressing the root-causes of these deficiencies or communicating them to human operators, the military appears to have framed the issues as software problems that could be fixed through technical solutions.

Another problem that operators of air defense systems encounter is that of automation bias and over-trust. Human operators can be overly confident of the reliability and accuracy of the information they see on their screens. They may not question the algorithmic targeting parameters provided to them by the machine. For example, the Patriot system was involved in two well-documented friendly-fire incidents and one near miss during the 2003 Iraq War. When a Patriot system shot down a Royal Air Force Tornado fighter jet over Kuwait in 2003, the British Ministry of Defense's accident report said "the operating protocol was largely automatic, and the operators were trained to trust the system's software." But human operators need a more balanced approach; they need to know when to trust the system and when to question its outputs.

The combat information center on the Vincennes. Credit US Navy.

Operators can lose situational awareness.

As militaries integrate more automated and autonomous features into the critical functions of air defense systems, human operators' roles have changed. They've shifted from actively controlling the weapons systems to monitoring their operations. In real terms, the machines are now performing the bulk of the cognitive skills involved in operating an air defense system, not just the motor and sensory tasks. Human operators are increasingly either overloaded or underloaded with tasks vis-à-vis those delegated to the machine, and they have sometimes lost situational awareness, which the researcher Mica Endsley defines as "the perception of elements in the environment … the comprehension of their meaning, and the projection of their status in the near future." Particularly in the context of high-stress combat situations, this can make it nearly impossible for human operators to question system outputs and to make reasoned deliberations about whether certain targets are appropriate.

The tragic 1988 downing of an Iranian Air flight carrying 290 passengers and crew by a US Navy warship, the *Vincennes*, illustrates how human operators in the midst of combat can misinterpret computer outputs and make fatal mistakes. The Vincennes, a ship so advanced it was jokingly called a "Robocruiser" because of its AEGIS air defense system, was designed to handle the type of threat the Soviet Navy might pose on the high seas. It could track and respond to hundreds of airborne threats at a time. But during a skirmish with a few light

Iranian gunboats, the crew of the *Vincennes* misinterpreted data on their computer screens and identified an Iranian Airlines Airbus as an F-14 fighter plane descending to attack. A 1992 *Newsweek* investigation found that senior personnel on the *Vincennes* were unfamiliar or uncomfortable operating the AEGIS's complex combat computer system.

The wreckage of a Ukraine International Airlines passenger plane. Fars News Agency. CC BY 4.0.

War is already too fast.

Improvements in the speed and maneuverability of modern weaponry continue to reduce the time human operators have to decide whether to authorize the use of force. Take what happened to an unfortunate Ukraine International Airlines jet as a recent example. The Iranian operators of a Tor-M1 system near Tehran's airport shot down the civilian plane carrying 176 passengers and crew members in January 2020, only minutes after the plane took off. Iran blamed human error for the incident, saying the missile defense system hadn't been recalibrated after being repositioned. Operating without a full picture of known traffic in Iranian airspace at the time, they mistook the plane for an incoming American cruise missile. According to Amir Ali Hajizadeh, commander of Aerospace Force of the Islamic Revolutionary Guard Corps, the operators of the Tor-M1 had 10 seconds to decide whether to fire or not. The point here is not to excuse this tragedy but to highlight the almost impossible demands that such a timeframe represents for critical deliberation in high stress combat scenarios.

When taken together, these three challenges call into question the level to which humans can exercise meaningful control over specific situations in existing systems that rely on autonomy in targeting. While these tragedies have prompted episodic introspection, they have not necessarily led to a more fundamental reassessment of whether it is appropriate to further integrate automated and autonomous features into air defense systems.

Failures of air defense systems typically arise from the complexities inherent with human-machine interaction. But when things go wrong, it's frequently the individual human operators at the bottom of the chain of command who bear responsibility for what really are structural failures. Focusing on "human error" shifts attention away from a closer scrutiny of how the use of automated and autonomous technology structures the use of force.

Regulating autonomous weapons.

In our assessment, the decades long process of integrating automated and autonomous features into the critical functions of air defense systems has contributed toward an emerging norm governing the use of air defense systems. The norm is that humans have a reduced role in use of force decisions. Unfortunately, much of the international debate on autonomous weapons systems has yet to acknowledge or scrutinize this norm, which likely will apply to future systems, too.

Countries have been debating possible regulations on lethal autonomous weapons systems at the United Nations since 2014. Many states have agreed in principle that human responsibility for using weapons systems has to be retained to ensure that autonomous weapons systems are used in compliance with international humanitarian law. But this raises two questions. First, how can human control over the use of force be defined; and second, how can such control be measured to ensure that it is people, not machines, who retain ultimate control over the use of force?

Almost a decade after a nonprofit called Article 36 introduced the concept of meaningful human control, there is no agreement on what exactly makes human control *meaningful.* Not only does this lack of a shared framework complicate efforts to regulate autonomous weapons development; in a more practical sense, it also makes it difficult to assess whether the control humans have over various weapon systems meets the necessary legal and moral standards on a case-by case-basis.

Policymakers should analyze the precedents that the use of highly automated air defense systems and other existing weapons systems with automated or autonomous features in their targeting functions (such as active protection systems, counter-drone systems, and loitering munitions) have set and the ways in which these weapons are altering the relationship between humans and technology. Too often, incrementally integrating more and more autonomous features into weapons systems is presented as either an inevitable trajectory of technological progress or as a reaction to what adversaries are doing. The current crop of more-or-less autonomous weapons has created norms for human control over lethal force, and policymakers need to understand how these may undermine any (potential) international efforts to regulate autonomous weapons systems.

*

Note to readers: please click the share buttons above or below. Forward this article to your email lists. Crosspost on your blog site, internet forums. etc.

Dr Ingvild Bode is Associate Professor at the Centre for War Studies, University of Southern Denmark.

Dr Tom Watts is a Lecturer in Politics and International Relations at Hertfordshire University.

Featured image: Patriot missiles in Israel target an Iraqi Scud missile. Alpert Nathan / Government Press Office. CC BY-NC-SA 2.0.

Articles by: Dr. Ingvild Bode
and Dr. Tom Watts