AI ON THE BATTLEFIELD? – IT'S ALREADY HERE

By Paul Springer and Jacqueline Whitt February 20, 2020
https://warroom.armywarcollege.edu/podcasts/autonomous-warfare/

Jacqueline Whitt: Hello, and welcome to A Better Peace the War Room podcast. I'm **Jacqueline Whitt** Professor of Strategy at the U.S. Army War College and the Editor for A Better Peace. We're glad you've joined us for another episode. Thinking about the future of warfare occupies an awful lot of time and energy amongst national security and military professionals and one of the most robust areas for discussion is the development and use of autonomous systems in war. Military strategists and futurists are thinking about what's possible now or in the future as well as what's desirable. Introducing autonomous systems into warfare forces us to think about important questions about the role of technology in war and the relationship between humans and technology. I've asked **Dr. Paul Springer** to come join me in the studio. He's the Chair of the Department of Research at the Air Command and Staff College in Montgomery, Alabama. He's going to talk with us about the future of autonomous systems. He is the author of a book *Outsourcing War to Machines: The Military Robotics Revolution,* and his research agenda is increasingly focused on issues related to cyber and robotic technology and the future of war. Paul, welcome to the War Room.

Paul Springer: Thank you very much for having me. I am excited to be here.

JW: I end up starting lots of podcasts with definitional questions. So, you get a definitional question to start. What is autonomous warfare?

PS: The notion of autonomy is the question of what level of decision-making an individual or in this case a machine is enabled to engage in. When we use the term robot, for example, m9ost definitions of robots require a machine that's capable of sensing its environment and making decisions of how it will behave on the basis of its sensing. So, a fully autonomous machine does not have external inputs once its original programming is effectively placed into the system. It does not have a human controller telling it what to do. It's making its own decisions on the basis of the input that it has received both in terms of its programming and in terms of the environment that it's currently operating in. Fully autonomous machines, or in the case of the military, fully

autonomous weapons, are those that are capable of choosing when, where and how to engage in lethal activity against potential enemies and that's the subject at hand today.

JW: Okay. That seems terrifying. Is there a concrete example of an autonomous system that we can think about in every day usage right now?

PS: There's actually a lot of autonomous systems that are in current operation. Some of them have lethal capabilities and some don't. Probably one of the more well-known that's used for intelligence surveillance and reconnaissance is the RQ-4 Global Hawk. This is a device that's capable of flying enormous distances and collecting almost limitless amounts of data through a very wide sensor suite package, but it actually flies itself. It does not have a remote operator that's telling it where to go and how to fly largely because that's what works best for it. There are also some lethal options and they've been in the field for a long period of time. A lot of listeners might be familiar, for example, with the Patriot Missile Battery. The Patriot Missile Battery has a fully autonomous mode. Essentially, you can tell it that it has the capability to engage enemy targets and fire upon them without human inputs and that's because sometimes the only way to defend a location of point-defense is through the reaction speed of machines. We also have all kinds of other point-defense systems such as the CIWS which is the basically last-ditch defense against inbound missiles for naval vessels. There is a land version. That's the Counter-Rocket, Artillery and Mortar system that I liked characterize as R2-D2 with a Gatling gun strapped to him. This thing is capable of firing thousands of rounds per minute and doing so to shoot down inbound mortars and small rockets and there's just no mechanism for a human to try to track those targets and choose whether to engage or not. I would argue probably the most well-known similar system in the world is the Israeli system Iron Dome. Iron Dome essentially tracks inbound projectiles, makes its own determination of where it thinks they're going to land and only engages them if shooting them down actually represents a better approach to the safety and security of Israeli citizens than allowing them to continue on their course. They may very well be shelled by hundreds of rockets in a given attack, but Iron Dome will only engage the ones that are likely to strike populated areas.

JW: I think this gets into other areas where we think about why, and you've talked a little bit about why autonomous systems are desirable, so reaction time, this decision-making capability, why are humans attracted to the idea of autonomous systems in warfare? I think it really goes back to the notion of wanting to be able to engage in the use of force without putting one's own forces at risk. And so, technology has often taken the place of the brute force aspects of military combat and whether it's armoring your troops to make them less vulnerable to enemy weapons, whether it's putting them inside vehicles or putting them inside aircraft to fly above the enemy, this notion of being able to essentially bring down violence and punishment upon one's enemies is at the heart of warfare and conflict. But particularly for the United States, one of the limiting factors of the ability to use military violence is the possibility of absorbing casualties. There is a

perception that the United States' public will not accept heavy casualties in war. I disagree, I think the public is willing and able to accept heavy casualties if it believes that the objective of the war is worth paying that price. But when you look at things like World War II, the United States is engaged in that war on two fronts across the globe and is going to lose about 400,000 service personnel. The Soviets are also engaged in that particular war and their casualties are measured in the millions. We don't even have a particularly precise number for the number of Soviets that died in World War II because their entire approach to conflict was very human-based, whereas the United States had started to use technology to try to compensate for numbers.

JW: So, when we think about which militaries, which defense systems maybe are particularly attracted to this sort of high-technology maybe standoff forms of warfare, do we see patterns in who is really going after research and development of these types of systems?

PS: Trying to build these kinds of systems requires a significant investment. The start-up costs are enormous quite frankly, but once you've actually created these systems, replicating them and producing more is relatively easy and there's a lot of things that make them particularly attractive. One is of course you can change the programming of these machines. You don't face problems of morale. Robots don't desert, robots don't walk away to the enemy, robots just do what they're told effectively. They follow their programming. There's also some of the factors that we don't necessarily think about it the forefront. If the United States' President, regardless of who that president is, wants to send U.S. forces into harm's way, the War Powers Resolution requires that individual to notify Congress of what they're doing. That individual doesn't necessarily have to notify Congress of anything if all they're doing is sending machines into conflict because they're not bound by the same systems and there's no grieving family if you lose a robot. And so, I would argue that this actually makes warfare more attractive as a means of national interaction with rivals and potential enemies because the political costs of risking your robots are almost nil.

JW: I think that's an interesting problem about the assumptions, about cost, the assumptions about effectiveness that we make in the political arena, as well as in the military arena on the ground. If we think about the future and if we could make the tech perfect, if we could get the computer code exactly right and the machines are making exactly the decisions we want them to make, are there still challenges or obstacles to using autonomous systems in warfare?

PS: In a perfect world, I suppose, in a perfect world we probably wouldn't have war, but let's assume that in a much better future world, we still have war but that war is characterized by a struggle of machines and all of the combatants agree that whichever side is successful using only their autonomous machines will be declared the victor, that I suppose, would be a substantial improvement in the outcomes of war. It doesn't really match human history though. We've had plenty of examples.

JW: I feel like we will find a way to mess that up.

PS: Well, we've had plenty of examples where one side has a significant technological advantage over the other and the side that was at the disadvantage essentially tried to offset technological inferiority with human wave attacks or numbers, and because in some cases, the technologically inferior foe has won the war against the technologically superior foe, that gives hope to future fighters. So, I can envision a scenario in which first-world powers fight against one another using autonomous machines. What I can't envision is either side surrendering until they have absorbed enormous human costs. So, what I think you would wind up having is one side triumphs on the machine battlefield and then winds up having to unleash those extremely lethal, very efficient machines on the human population of the other side in order to compel their surrender. Nation states don't surrender easily. They tend to be able to absorb enormous amounts of punishment and very large amounts of casualties particularly in wars of national survival. And so, I can very easily envision, effectively a holocaust or a genocide, being brought about by these machines if one side is absolutely determined to conquer and overwhelm the other.

JW: Because the human element at some point is still there. If nation states or if states are still comprised of humans, political actors, they're still political decisions to be made. When we think about maybe the handoff, at what point do autonomous systems or can autonomous systems take over decision-making?

PS: In a lot of ways, autonomous systems have already taken over decision-making in our everyday lives and we don't spend a lot of time thinking about it. Algorithms rule so much of the information that we take in on any given day. Machines attempt to influence us on a daily basis whether it's through advertising, whether it's through opportunities that are presented to us. And in warfare that's also true. But everyday life is full of all kinds of autonomy that we don't really even take notice of any more. Anyone that flies on a regular basis has probably, actually been on an aircraft that was landed by the machine, not by the human pilot. Machines actually tend to do a lot of those tasks better and more safely, particularly in difficult weather. Machines have other sensory capabilities that humans aren't capable of utilizing, and their reaction speed of course is so much faster that they are often more capable of responding to the unexpected than humans are.

JW: At the same time, we want the human pilot there, right?

PS: Well, we're kind of gradually being weaned away from that. In aircraft right now, sure, yes, everybody wants to see the pilot and wants to know that that pilot is capable of flying that aircraft and that there's a co-pilot also capable of flying that aircraft. But when you look at things like metro systems, subway systems, an awful lot of those have started to become fully

autonomous and people are becoming more comfortable with it. When we look at things like self-driving cars, people are becoming more comfortable with that notion and so gradually, people are becoming more comfortable with the idea that you maybe don't need a human in the loop, particularly when we read about humans that make terrible errors, terrible judgments, the train conductor that's playing on their phone rather than paying attention to their job, for example, causes us to call for new forms of safety, new forms of security. We also see aircraft accidents that are deliberately caused by individual pilots. One can look at the German Airwings disaster of a few years ago, for example, where a pilot effectively decided to commit suicide and took everyone else on the plane with them. Had that aircraft been equipped with essentially an autopilot override, if the machine had superiority over the human, he might not have been able to crash the aircraft.

JW: I keep thinking about our insistence that humans be in the loop on certain decisions when it comes to warfare and when we think about remotely piloted vehicles, lethal systems we want humans in that decision-making process. Is there good reason for that?

PS: Personally, I like having a human involved in the ultimate decision to take lives because there's no way to punish a machine for making a mistake. There's no way to hold a machine accountable for making a mistake. When you're talking about actually taking a human life, that's the one thing that we don't have any mechanism to provide restitution for if we get it wrong. And for me, what I like to refer to as the dark triad, is a machine that's fully autonomous, that’s lethal and that has offensive intent. I'm perfectly comfortable with machines having any two of those elements but once it has the capability to go seek targets and kill them without human input, now, I start to get very nervous. We have designed but not fielded those systems. You can envision those systems though. Even in the 1990s, we had the LOCAAS system which was fully autonomous, sought its own targets, it was actually looking for Soviet-designed hardware. It was basically a tank killer and it would fire essentially multiple rounds before the main body of the system became the fourth and final strike platform. It didn't work very well. It tended to actually hit the same targets again and again. It didn't have the processing power to keep track of which targets had already been struck and could be bypassed. By now, 25 years later on the other hand, between cloud computing and distributed processing, it's entirely possible for these machines to effectively work together and have extremely complex behaviors that are derived from very simple programming. It's possible to take a small group of robots, about a hundred robots, and map out a city the size of Chicago over a period of about one week. You only really need to give them three commands: one, move around and report back what you see; two, stay away from other robots; and three, continue to move so if you get trapped, essentially backtrack and move around some more. Just doing that they will autonomously map the entire city without any kind of a centralized control determining where each one goes.

JW: So, nobody has to plan it. Nobody has to program each individual one to cover a certain sector. With those three things, they'll figure out…

PS: That's right. We actually derive a lot of that type of programming from what's called biomimetics. We essentially follow the behavior of very simple creatures on earth. When you see a swarm of birds flying through the sky and they seem to engage in very complex maneuvers, but they don't collide with one another, it causes you to realize that, in their in their general behavior, it tends to be, fly in the direction that the flock is moving but don't collide with one another. Very simple programming if you will, though they are biological creatures and very complex behavior for the outside observer.

JW: With this dark triad, it sounds like the right-choice triangles, right. Pick any two, but you add the third one and you're in trouble or you're not going to get it. Which one is furthest away from military planning?

PS: Well, when we look at those three, I would say that the one that's probably the furthest away from military planning is the autonomous aspect of things. Military planning tends to be offensive-minded and it tends to have lethality in mind and the autonomy concept… Military forces prefer to maintain control over their own operations and their own environment and one can argue that that's kind of the first objective in war is to seize control whether it's of a geographic position, whether it's of the pace of operations, whether it's of the style of conflict that one is choosing to pursue. So, the idea of turning that control over to some other entity, to an autonomous system, I think is the one that's furthest away from the way that the military prefers to operate. But these systems, they're very alluring because they do offer enormous tactical advantages in particular, and I worry that we're going to field a lot of these systems without really thinking through both the legality and the morality of putting them into the field.

JW: Are there new legal and moral questions and quandaries that we need to think about or can old ways of thinking and frameworks for ethical and moral judgment be applied in these cases?

PS: There are folks that would argue that the existing framework works just fine. I happen to disagree. I think that these machines constitute a revolution in military affairs. They're fundamentally altering the nature and the character of warfare, and as a result, the old paradigms don't apply. I would argue that we live in the Westphalian state still, and that's largely derived because of fundamental changes in warfare that occurred in the 17th century that forced us to rethink our previous assumptions about norms and behaviors. I think that these raise a whole host of challenges, once again, in part because you cannot hold machines accountable in the same way that you can humans. And I'm not sure that you can hold a human commander accountable for a misbehaving robot in the same way that you might hold them accountable for a misbehaving enlisted soldier.

JW: When we talk about misbehaving robots, what does that mean? Does that mean there's a mistake in the code? Is there a glitch in the system? Is it just that machines breakdown and sometimes they mess up?

PS: Sometimes it's any of those things and sometimes it's a judgment situation. Humans still make better judgments in really difficult and challenging situations. Machines could do a better job of facial recognition than humans can because they're not fooled by a lot of the same systems that easily confuse humans, but when it comes to what a facial expression means, for example, humans tend to do a much better job.

JW: Reading emotion, thinking about reactions and responses.

PS: Sure, and when you think about some of the worst incidents in American military history, the things that we are least proud of, whether it's the My Lai Massacre or the massacre at Wounded Knee, I suppose I'm going down the massacre route here which might be a bit of a scare tactic, but when you think about those, those are examples of people doing things that we knew were wrong and that we condemned as wrong but we struggle to even hold humans accountable for those types of behaviors. When young Lieutenant Calley led the massacre at My Lai, 400 civilians are killed, but Calley himself is only going to spend a few years under house arrest. How do you even remotely account for that type of an incident if it's carried out by a malfunctioning machine? And we do see these machines malfunction on a fairly regular basis. They break down like anything else does. We see humans malfunction too, but we have the mechanisms to account for that. There was a very famous incident about 10 years ago of an autonomous air defense gun that was being demonstrated in South Africa. Essentially, they had a whole bunch of their military leadership to come see the new gun and how well it worked. The plan was to fly some target drones overhead. It would shoot down the drones with live ammunition to demonstrate its capabilities. When they turned the gun on into its fully autonomous mode, it immediately pivoted itself toward the bleacher where all of the leadership were sitting and open fired. It was later determined that it had locked on to essentially the fan in a portable toilet facility that was operating at the same RPMs as the target drones it was supposed to engage. As far as it was concerned, it was doing exactly what it was programmed to do. Unfortunately, it had no mechanism to realize that shooting through humans on its own side to get at the target was an inherently bad thing. Humans would presumably not have made the same mistake.

JW: For sure. I'm still slightly terrified. This is a constant sort of state of affairs I think for some people in the national security arena. On the whole are you an optimist or pessimist about the future of autonomous systems in war?

PS: I am 100% a pessimist. I suppose that's just my worldly outlook. I can envision a future in which essentially political leaders can rely upon what amounts to, essentially, a military composed of machine mercenaries that will do exactly what they're told without a single question of why or how or whether this is a constitutional behavior. I can envision a scenario in which a president chooses to engage in widespread war with very little if any accountability or control. One of the things that keeps me awake at night is the notion that if the United States is to engage in conflict, if we ever declare war again, and I'm not sure that we ever will, but if we ever do, imagine a scenario in which the president refuses to negotiate any kind of an armistice. Quite frankly, when we declare war, we are effectively turning the American President into a dictator with very broad powers, the ability in some cases to effectively suspend elements of the Constitution. So, I can envision scenarios in which politicians would not have to convince humans to go along because they have the keys to the kingdom, autonomously speaking, and they no longer need that difficult challenge of leading. They're in a pure command mode instead.

JW: Really difficult and significant questions about technology, about morality, about ethics, about accountability, the political costs and benefits of machines, of robotics in war and the nature and character of war. We've covered some pretty significant bases in the study of war and thinking about what the future might hold. So, Paul, thanks so much for coming and joining me today on A Better Peace.

PS: Thanks so much for the opportunity.