CONVERSATIONS ON STRATEGY

PODCAST TRANSCRIPT

# Paul Lushenko, Rob Sparrow, and Adam Henschke
# AI, Trust, Culture, and the Military (Part 2)

Contesting Paul Scharre's influential vision of "centaur warfighting" and the idea that autonomous weapon systems will replace human warfighters, this podcast proposes that the manned-unmanned teams of the future are more likely to be minotaurs, teams of humans under the control, supervision, or command of artificial intelligence. It examines the likely composition of the future force and prompts a necessary conversation about the ethical issues raised by minotaur warfighting. The guests also explore culture and trust in relation to AI and the military.

**Keywords: manned-unmanned teaming, centaur warfighting, artificial intelligence, future force, ethics, trust**

E-mail usarmy.carlisle.awc.mbx.parameters@army.mil to give feedback on this podcast or the genesis article.

**Stephanie Crider (Host)**

This is part two of a two-part podcast. You are listening to Conversations on Strategy. The views and opinions expressed in this podcast are those of the authors and are not necessarily those of the Department of the Army, the US Army War College, or any other agency of the US government.

My guests today are Assistant Professor Paul Lushenko, PhD; Professor Rob Sparrow; and Assistant Professor Adam Henschke. Lushenko is the director of special operations and a faculty instructor in the US Army War College's Department of Military Strategy, Planning, and Operations. Sparrow is a professor in the philosophy program at Monash University, Australia. Henschke is an assistant professor in the philosophy section at the University of Twente, Netherlands.

About a year ago, Sparrow and Henschke coauthored "Minotaurs, Not Centaurs: The Future of Manned-Unmanned Teaming." You can find it in the Spring 2023 issue of *Parameters*. Look for volume 53, issue 1. Lushenko and Sparrow took this concept a step further with their piece, "Artificial Intelligence and US Military Cadets' Attitudes about Future War."

So, before we get started, I just want to thank you all for joining me today. Welcome to Conversations on Strategy.

How do we understand[artificial intelligence or] AI in terms of the legitimate use of force?

**Adam Henschke**

I might start this one off. At least with the paper that Rob and I had in *Parameters*, as Robert mentioned earlier, it was contrasting the centaurs view of AI in military practice and suggesting, alright, we might need to think in a different way that the use of AI in military context will be much more like a minotaur. We have got the AI making the decision and then the humans doing the dirty work out in the field.

The reason I mention that is, for a long time, at least, a lot of the discussions around AI—particularly the drones, lethal autonomous weapons systems, et cetera—a lot of the discussions there on legitimacy were focused on decisions about or questions whether AI should be making decisions that might take someone's life. But then there's this deeper moral issue: is a lethal decision something that a robot or a nonhuman should make?

But with the "Minotaurs" ["Minotaurs, Not Centaurs: The Future of Manned-Unmanned Teaming"] paper, Rob first suggested the idea to me, one of the things that really interested me was what does this mean for legitimacy? Because what we're arguing in this "Minotaurs" ["Minotaurs, Not Centaurs: The Future of Manned-Unmanned

# Paul Lushenko, Rob Sparrow, and Adam Henschke
# AI, Trust, Culture, and the Military (Part 2)

Teaming"] paper is the decisions that are going to be made are going to be much more likely at the strategic level rather than the kill or don't-kill level.

And that has, I think, really significant implications for the perception of legitimacy of fighting forces. Because the problem is, if you don't have a civilian community who trusts military—either their own military, or potentially a military who is interacting with their own military—if the civilian community doesn't trust that, that can have a really negative impact on their legitimacy. And particularly if you think of this here. Imagine an invading force. It is known that they are using AI, or AI is making a lot of the strategic decisions for it. It seems to me that there would be very, very high distrust in that civilian community, even if that force followed all the laws of armed conflict and it was engaged in a humanitarian mission or something along those lines. If you met all the various laws of armed conflict and just-war criteria, there would still be, to me, a very high likelihood that, if it came out or if it was known or even believed, that the decisions made were made by AI, there would be very, very low trust in that military and what it's trying to do.

I think here, we've got this really interesting challenge about the overall legitimacy of military operations and the militaries themselves, that can be disrupted by AI. That might be a cultural issue, as in, as we all become more familiar with AI, maybe such questions of legitimacy and trust go away. But I do think that there's a really interesting issue there about the legitimacy of the force itself, rather than the narrower questions about legitimacy and fighting, which is what a lot of the discussion has been so far.

**Robert Sparrow**

One way of thinking about this is to think about the implications of AI for winning hearts and minds, to use that old language. And I actually don't think this is straightforward because better targeting practices, for instance; killing fewer noncombatants; reducing civilian casualties; and certainly drones. There's been a long debate about whether precision-strike capabilities enabled by drones lead to overall decreases in noncombatant casualties. Depends a lot on what your comparison case is because if it was gonna be saturation bombing, then clearly the drones look better, but if the alternative was not choosing to attack a particular target when it was close to a civilian area, then that comparison might go differently.

But there's a case to be made that AI might actually help win over a population by making warfighting better in various ways. But also, it is really hard to imagine people responding well to fully autonomous systems commanded by an AI.

Or even, if you're talking to a service member, if there's a soldier standing at a checkpoint, and you ask them, "What are they doing there?" and their answer is, "A computer told me to be here," I think that is likely to be pretty detrimental to winning over a population of a occupied territory.

The other thing I feel strongly about here is that we shouldn't think that the machines themselves are going to be ethical, or start doing ethics. Unsurprisingly, I am hostile to the idea that you can automate philosophy. I think you can build machines that do various things better or worse, and that might include attacking the appropriate targets. But the robots are not going to be, or the computers are not going to be wracked by regret. They're not going to be put on trial for war crimes. And they're not going to be capable of explaining to the grieving relative why their son has come home in a body bag.

And so, I think it's really important to think it's human beings who are going to do ethics and who are going to be making the kind of morally fraught decisions and the machines might carry them out. And then there's a question about the level at which the commands are given. Maybe there's a future where you say, "Occupy this territory," or "Invade this territory," or "Win this war." My sense is that's not the right level of decision making to hand over to AI.

# Paul Lushenko, Rob Sparrow, and Adam Henschke
# AI, Trust, Culture, and the Military (Part 2)

**Paul Lushenko**

I appreciate how both Adam and Rob problematized this. And I'm gonna take it a step further, having done a fair amount of research here. And I think it comes down to the way that we define the use, purpose, constraint, location of AI capabilities in terms of force employment. But also, importantly, how we define legitimacy, which is often conceived in terms of rightful wartime conduct, which is the moral aspect of legitimacy. But there's also a legal dimension here, which it says compliance with international humanitarian law.

And so, on the former point, if I draw from my research on drone warfare, what I've seen is that there are these broad, hand-wave definitions of what drone warfare is. Is it about the platform itself? Is it just about counterterrorism or high-value targeting? Is it about one form of conflict, which is irregular warfare? Is it about remote warfare? Often, drones are conflated with remote warfare, which is about a broader respatialization of force that includes things like proxy forces or, in fact, artificial intelligence.

And so, in order to really answer this question and test it empirically, I think we have to get down to the attribute level of AI. Which is why, to the extent this gives us some leverage, the typology that Rob and I introduce in this recent piece is so very important to me because it talks about different attributes, which are, over time, testable in terms of these methods.

The othe r point, too, is the difference between moral and legal legitimacy. Again, in my research, I've mostly focused on perceptions of rightful wartime conduct, which is based upon intuitions of morality. The legal aspect is also, similarly, very, very important. And I give you one example about where this really hits home. I've done a recent survey experiment among a very rare sample of chaplains in the United States Army. All chaplains in the United States Army, in fact, were given the opportunity to take this survey where I presented variation in the use of drones that were either loosely or stringently constrained through different standards of targeting. And then furthermore, where the capability was used, whether it was a nonactive or active conflict zone, which we also call "undeclared" theater of operations, like Pakistan, or "declared," like Iraq and Afghanistan.

And what I find is that, yes, legal and moral perceptions on legitimacy can covary, but they can also deviate—and quite widely—in terms of especially the type of theater of operation. And then, furthermore, they're also, here in this context, as a legitimacy paradox in which case they would think something is legitimate but not support it as much.

And so, Rob's absolutely correct. This is a very, very challenging space. And for me, the devil is really in the details. As someone who researches this empirically, we have to define what we mean by AI. And we have to define what we mean by legitimacy. And hopefully, if we buy into the assumption that public opinion exercises an important shaping effect on democratic accountability and oversight through public officials, this would have an effect on our policy formation, strategy formation, and indeed potentially even military modernization.

And so, I guess to put a pin in it, we need to do a lot more research to figure out how different attributes of AI really affect perceptions of legitimacy in these two different veins of legal and moral.

**Henschke**

Just a question for you, to follow up on that, Paul. I think you're right in marking the distinction between the ethical or moral legitimacy and the legal legitimacy question. But have you looked at, or are you interested in something that we might consider more political legitimacy? And this goes a bit more to the just cause, right intention, those sorts of considerations.

Obviously they're moral considerations, but they're also quite political. And it seems to me that one of the aspects that would be really worth looking at is impacts of AI, perceptions on AI, on those perceptions of political legitimacy. Again, whether it's our own forces or potential combative forces.

# Paul Lushenko, Rob Sparrow, and Adam Henschke
# AI, Trust, Culture, and the Military (Part 2)

**Lushenko**

That's a really good question. And I think, to the extent, Adam, that I do that is through this notion of microfoundations. And Rob and I do this a little bit in our piece, as well, is to attempt to shed new light on values and beliefs that would underline these overall attitudes. But what I find is that there are multiple different underlining considerations that are instrumental, that are normative, that are operational for at least chaplains in the setting that I study.

And if I recall correctly, what I find is that it comes down to the perceived veracity of intelligence, which can also be very politicized. It comes down to perceptions of threat. It gets back to securitization theory in the way that we legitimate a threat to, in this case, the United States homeland as a predicate for the use and approval of force abroad. And then, what would be no secret to both of you as philosophers in this space is the degree to which we can reduce the potential for civilian casualties on the battlefield.

These things are kind of an ecumenical spread of different considerations that, quite frankly, can be subconscious—unless or until you ask someone to explain their intuitions and then they draw heavily on these different microfoundations.

**Host**

We normally see generational shifts in attitudes towards things like new technology. Why is this a problem with AI?

**Sparrow**

I'm not sure it is a problem with AI. It's something that we need to pay attention to. I don't know, when machine guns were introduced there were some human-factor issues—in terms of people's willingness to use those weapons, the concepts of operations around them—that was important to pay attention to. And typically what happens is people kind of muddle through, and then eventually you get people start to understand how best to use some new weapon or tactic and the moral questions that come along with it.

It's very clear at the moment, this is the very early stages of the potentially very radical transformation of the way war is fought. We need to pay attention to this stuff. That's not necessarily to say that differences in attitudes are themselves a huge problem. There's potentially a problem.

If the people who make decisions about introducing this technology don't understand how the people who are using it—who will be using it—feel about it, then the plans may go awry. If you want to fight wars alongside of AI well you need to pay attention to the human side of the human-machine teaming. And you can't assume that your attitudes as a senior commander, or senior scholar, that you can extrapolate from that to how younger people are . . . are likely to feel and behave.

**Henschke**

That's a really important point. I remember, probably 10 years ago now, I was at a workshop looking at issues to do with cyberwar. And I am older now than I was then—I wasn't super old then, but I was still of a generation who existed before the Internet. I remember when I used to look up things in a library and not have to use a library computer. That sort of thing.

But the thing that was really striking to me at that workshop—I was talking to someone, and she was a colleague from the Norwegian military. And she brought to my attention the notion of the digital native and how my thoughts, my attitudes, my experiences are probably going to be significantly different to someone who is of this digital-native generation. And now, the couple of generations of digital-native people who have been immersed in computer-based technologies that potentially make their perceptions and understandings of these things significantly different to something that I've had.

So, I think this is where Paul and Rob's work—actually looking at the trust attitudes and potential generational

differences in trust attitudes—is so important. 'Cause I've got a bunch of ideas and thoughts about these things. But if it turns out that's all crap, like if I'm saying, "Ah, well, you know, people shouldn't trust this." And it turns out that they don't, then there's no problem there. And so, we need to be really attentive to not just these cultural issues but the significant generational shifts that occur within communities and in militaries as well.

**Lushenko**

I may just add here that, first of all, what I'm often asked when I talk to people about drones is, "What gives? What's the difference between other standoff capabilities like jets, bombers, attack helicopters, artillery, so on and so forth?"

I think what I see with AI is a magnification of the normative concerns that people have. If you take a look at the work by Neil Renic, a colleague of all of ours, used to be at University of Queensland, now at University of Copenhagen. It's a really interesting book on the radical asymmetry of drones—that drones impose such asymmetry disadvantage against an adversary that you actually transition into morally problematic killing. You remove the right for someone to defend themselves and give due notice to do so.

And I think with AI, you get a magnification across these legal, moral, and ethical considerations for distinction, for responsibility and accountability, in ways that you potentially don't even have with semiautonomous drones right now. In that respect, I think AI is a little bit different. And layer on top of that everything that Adam just talked about in terms of generational differences, because junior people—people younger than us, if you will—have such greater fluency with gaming, with social media, things of that nature.

And so, I think there is a little bit of difference here, as far as I can tell, in my research.

**Sparrow**

I just want to come back to this issue of the digital natives because when that's deployed, clearly there are now one or two generations of people who have grown up with digital technologies. They've been interacting with computers all their lives. But people often move very quickly from that to the idea that we should defer to digital natives.

But when it comes to history and technology change, there's actually two different narratives possible. One is a narrative of the gaining of wisdom, that you develop a better understanding and all the old fears, moral panic, people used to worry about how the Internet was affecting our brains. Now that we've got 20, 30 years of experience, now we know that all that stuff was silly. And we really understand things.

That's one possible narrative. But the other narrative is actually a narrative of corruption, and enculturation, such that all the old fears were right, but people don't notice anymore because that's part of what this technology being ubiquitous means.

Yes, you can grow up with something and be used to it and, as a result, actually be much less critical of it than you should be. Yes, there is this generational difference, but what's really important to recognize that doesn't mean that you should just defer to the attitudes of digital natives because part of what becoming a digital native might involve is actually giving up on some valuable stuff that was available to people who weren't fully socialized by these technologies.

Now, those are open questions. What I would insist was that people are used to something doesn't necessarily mean that they have a good understanding of it.

**Lushenko**

Well, Rob, I think we see that in our research. On the one hand, senior officers have been enculturated to support capabilities they don't trust. Because there's a broader narrative among the military brass that this is game-changing and we ought to employ it to maintain lethal overmatch of an adversary. And on the other hand, from our research, we find that cadets can be trained to trust and support minotaur warfare to degrees that are similar to their trust

right now in capabilities that are used for strategic decision making. More to your point, they could be as much as digital immigrants, as we talk about, as well as natives.

**Henschke**

This goes back to one of the things we were discussing before: the important distinction between trust attitudes and whether those trust judgments are worthy. The difference between trust attitudes and whether the targets of that trust are worthy of that trust.

And again, this is where the ethical, legal, perhaps even psychological-political analysis of these things is really important to try and work out some measures, some ways of both assessing and communicating when these things are worthy of that trust.

Exactly to Rob's point, I agree very much with that simply because there's a bunch of generations who have a particular set of attitudes—whether it's the older generations or the younger generations—that alone doesn't necessarily tell us the whole story. We need to be looking at whether the targets of those trust attitudes are worthy of that trust or not. And they might actually warrant greater trust than we're giving them, but they may also warrant far less trust than we give them. And so we need to be looking at those criteria and the communication of those criteria when the targets of trust are worthy of trust or not.

**Host**

How can military ethics include training to develop the skills and confidence and to develop a culture that encourages the questioning of AI commands?

**Henschke**

I think we can take lessons from both the evolution of military ethics and the evolution of training in military law. Perhaps, I'm not sure if this is 100 percent true, but I assume 100 years or so ago, when a lot of the laws of armed conflict were being developed, a lot of people involved in the military were like, "Ah, there's no way we can train people to do this. This is all too hard or undermining our fighting forces and our capacity to win," et cetera, et cetera.

We've seen, in the past 100 plus years, that training can happen both with the legal aspects of military activity and also the ethical aspects. And I think we can look at the ways in which those educational programs were developed—particularly around issues of when to question commands and when to accept that a command is probably going to be a decent command to follow.

This is very, very different content but a very similar thing here—especially if we're thinking of AI that is at this higher strategic level—to question those AI commands and to know when it is good to question them and then to know also whether it is good to go ahead and trust them.

**Sparrow**

My sense is the fundamental ethics, the reasons for not directly targeting civilians don't change, whether it's you doing it on your own initiative or whether you're doing it because of some sort of targeting algorithm has told you to do it.

What I do think people should be paying attention to here is institutional culture and process of institutional socialization. There's quite a long history of human-factors research in relation to automation. And we actually know quite a lot about how to operate well with automation. And so, there's particular industries and sectors, like the airline industry, that have done that well. And then there's particular industries and sectors that have done that poorly. And one of the differences is precisely in training people to consider the possibility that the information provided by the instruments is wrong, to keep reinforcing the need for the maintenance of the skills that are necessary to do that because de-skilling is a huge problem in this context. You might tell people before you fire,

# Paul Lushenko, Rob Sparrow, and Adam Henschke
# AI, Trust, Culture, and the Military (Part 2)

"Check whether the information you have is likely to be accurate." But if people don't train without these systems and they don't maintain the human skills necessary to check the information provided by AIs, then they can't do it when they need to.

But this isn't ethics, this is organizational culture and socialization. And so, some of this I think is amenable to change or design, but you have to be quite conscious that just importing a whole lot of new tech and imagining that it will work without changing the way that you operate, that's naive.

So, the human-factor stuff is really important. In a sense, that's where people should be paying attention. I think the fundamental ethical questions are probably the same.

**Lushenko**

I agree as well. I think there's a durability to our current frameworks that are applicable to the space of AI. I would also add that I think Daniel Brunstetter has got some great work on jus ad vim, the use of force for limited purposes that would be helpful in the space as well.

But to pick up on a thread, briefly, that Rob talked about, one of the key learning outcomes in our seminar here in the war college is to show the implications of de-skilling for higher-order thinking that goes into strategy making, as well as policy formulation. And so, we've done exercises here where we have segmented a class into three different groups. One has gone through a military-planning process without any augmentation from AI. One has used AI to spot-check if there's anything they've missed. In other words, asking AI questions about other considerations. And then a final group has used AI from the start to actually formalize a policy in and of itself.

And what we found in bringing everybody together and synthesizing their feedback is the importance of abduction, which is to say, experiences and expertise that can call BS, for lack of a better term, on the outcomes of AI. And so, to Rob's point about de-skilling, I think what you have to do is contextualize these systems into a broader training regime at different ranks and grades—lieutenant, captain, major, all the way to colonel—where you bring people through rote processes of planning long-form. And then you can augment their learning with these AI systems to show the importance of mastery of concepts.

I think Rob's point is very, very powerful and something that we've impressed upon our emerging senior leaders here at the War College.

**Host**

Do you have any concluding thoughts you'd like to share before we go?

**Sparrow**

I guess this is the beginning of a process. It's a really important period in which there's an opportunity, maybe, to get just slightly ahead of the curve and make sure that we plan for a world in which these AI tools are available. There's going to be a real temptation to use them in military contexts. And the relationships between human beings and machines are themselves something that might evolve, and we need to understand how that might happen.

**Henschke**

For me, probably the most important thing—I already had this belief beforehand, but it's been heavily reinforced by the discussion today—is the importance of attending to culture and, particularly, the cultures within militaries. Obviously, we spoke a lot about trust. But, how do you develop a culture such that we can make better decisions with AI? I was interested in the example that Paul mentioned of how to anticipate the risks of de-skilling. And that really points to, you know, having a mixture of not ignoring AI but not going completely, "Let's let AI do all warfare." Trying to develop a set of cultures so that the AI is used in the best ways that it can be but not in a way that

# Paul Lushenko, Rob Sparrow, and Adam Henschke
# AI, Trust, Culture, and the Military (Part 2)

undermines the capacity for humans to reflect on, to criticize, and to understand the implications of the decisions that are being made.

And some of the work that I've been doing is an interesting area of some of the conceptualization of trust, which sees trust as an unquestioning attitude.

And, as a final comment, often when we look at and think about trust, it carries a bit of a positive judgment with it. If we say, "Oh, we trust that," it kind of implies something good is going on there. But there can be many situations where trust is poorly placed, still. You're placing trust in the wrong thing or in the wrong person. If we think of trust, instead, as an unquestioning attitude, put that into the context of developing and evolving military cultures in relation to AI, we want to encourage questioning attitudes. We've done this before with ways in which we brought in legal aspects of fighting wars, ethical aspects, to know when to ask questions and to know the sorts of questions to ask.

And I think here we want to try and encourage a similar thing. To encourage questioning about AI. It's not that we shouldn't trust it, but we should also see trust can be dangerous when it is an unquestioning attitude in relation to particularly important things that AI might be impacting in military decision making.

**Lushenko**

Rob and Adam are the world-renowned philosophers here. I am the hard-core security-scholar perspective. And so, I just may embrace that and say that, as a pragmatic question, to what degree do these capabilities change the character or nature of warfare? I think we often—not among our crowd here, but defense experts even—often conflate these two concepts. One is immutable: war is political. It's intensely human. It's a clash of wills. Whereas, the character of warfare evolves with the emergence of different technologies.

It seems to me that what we're seeing right now is the implications of AI across different strategic contexts; we ought to study it that way. And that's what we talk with our students about on a day-to-day basis here at the War College.

**Host**

Definitely feels like we came away with more questions than answers. Great chat though. Thank you all for being here. I hope we get to do this again sometime. It seems like there's a lot of opportunities that we might be able to do this again.

**Sparrow**

Thank you, Steph.

**Henschke**

Thanks, Steph.

**Lushenko**

Thanks a lot, Steph.

**Host**

Listeners, for more Army War College podcasts, check out Decisive Point, SSI Live, and A Better Peace.

**Listen to the podcast:**

https://ssi.armywarcollege.edu/SSI-Media/Podcasts-Lectures-and-Panels/Conversations-on-Strategy-Podcast/mod/67381/player/581/audio/80678