Foreign Policy at BROOKINGS

ARTIFICIAL INTELLIGENCE IN WAR: HUMAN JUDGMENT AS AN ORGANIZATIONAL STRENGTH AND A STRATEGIC LIABILITY

AVI GOLDFARB AND JON LINDSAY

NOVEMBER 2020

> We must remind ourselves that it is simply not possible to construct a model for the art of war that can serve as a scaffolding on which the commander can rely for support at any time. Whenever he has to fall back on his innate talent, he will find himself outside the model and in conflict with it; no matter how versatile the code…*talent and genius operate outside the rules, and theory conflicts with this practice*.
>
> —Carl von Clausewitz, *On War*

EXECUTIVE SUMMARY

Artificial intelligence has the potential to change the conduct of war. Recent excitement about AI is driven by advances in the ability to infer predictions from data. Yet this does not necessarily mean that machines can replace human decision makers. The effectiveness of AI depends not only on the sophistication of the technology but also on the ways in which organizations use it for particular tasks. In cases where decision problems are well-defined and plentiful relevant data is available, it may indeed be possible for machines to replace humans. In the military context, however, such situations are rare. Military problems tend to be more ambiguous while reliable data is sparse. Therefore, we expect AI to enhance the need for military personnel to determine which data to collect, which predictions to make, and which decisions to take.

The complementarity of machine prediction and human judgment has important implications for military organizations and strategy. If AI systems will depend heavily on human values and interpretations, then even junior personnel will need to be able to make sense of political considerations and the local context to guide AI in dynamic operational situations. Yet this in turn will generate incentives for adversaries to counter or undermine the human competencies that underwrite AI-enabled military advantages. If AI becomes good at predicting the solution to a given problem, for instance, a savvy adversary will attempt to change the problem. As such, AI-enabled conflicts have the potential to drag on with ambiguous results, embroiled in controversy and plagued by crises of legitimacy. For all of these reasons, we expect that greater reliance on AI for military power will make the human element in war even more important, not less.

INTRODUCTION

We often hear that substituting artificial intelligence (AI) for military personnel will accelerate the tempo of combat operations and increase the chances of inadvertent engagement and escalation. We question these assertions by adopting an economic perspective on decision-making that distinguishes *prediction*, the ability to infer missing data, from *judgment*, the specification of payoffs. Prediction has been made easier and cheaper by advances in machine learning and an abundance of data, yet we suggest that the complementary component—human judgment—is becoming more valuable. Because friction and controversy are inevitable in national security, we expect the military use of AI to make human judgment even more crucial and challenging. Ironically, however, the same organizational capacity that enables judgment, and thereby makes war fighting more predictable and controllable, also has the potential to make conflict more ambiguous and less decisive. In short, the ability to automate aspects of decision-making can make it harder to come to a decision within an organization or on the battlefield.

This development is grounded in long-term historical trends in the organization of military power. Previous episodes of technological substitution in military history—first for power projection and then for intelligence—tended to make complementary human skills more important. Human labor was often shifted into different parts of the military decision-making process, rather than fully eliminated. The implication for AI substitution today is that the individual and organizational capacity for judgment will increasingly become more important, distributed, and challenging. To meet this challenge, militaries will require better-educated, technically savvy personnel, even as organizational coordination and leadership will become more complicated.

As human judgment becomes increasingly important in the deployment of automated military power, it follows that the organizational capacity for judgment will become an increasingly attractive target for enemy action. Most scenarios of AI-enabled rapid war fighting and escalation assume traditional modes of combat, with AI-enabled weapons targeting enemy formations. However, an intelligent adversary facing an AI-enabled force has strong incentives to change the game, both to avoid being targeted by the strengths of AI and to exploit its weaknesses. If judgment is a key organizational source of strength, then it might also be turned into a key strategic liability. We thus expect greater reliance on AI to heighten the salience of strategic threats to judgment such as disinformation, subversion, and moral controversy. At a tactical level, adversaries may attempt to poison the data and training processes of AI systems. At a more strategic level, adversaries will have incentives to attack organizational, alliance, and societal cohesion. War could thus become even more protracted, less decisive, and more fraught with ethical dilemmas.

This essay has four parts. First, we explain why machine learning in general makes prediction cheaper and human judgment more valuable. Next we apply this economic perspective to military affairs to explain how data limitations and judgment challenges define the comparative advantages of humans and machines in war. We then discuss the organizational challenges that arise from distributing judgment throughout an AI-enabled military. Finally, we speculate about the strategic implications of greater reliance on, and political manipulation of, human judgment in war.

CHEAPER PREDICTION: MORE VALUABLE JUDGMENT

One of the key insights from the literature on the economics of technology is that technological substitution makes human complements more

important.[1] Some assets and skills become more important as a new technology replaces human labor, machines, and processes. Some old jobs are lost to technological substitution while new jobs emerge as complements. There is little reason to believe that AI will break with this pattern.

The recent resurgence of interest in AI and its commercial applications has been driven by rapid advances in machine learning—not by progress toward automating or surpassing the human mind (or what is often described as "artificial general intelligence"). In *Prediction Machines: The Simple Economics of Artificial Intelligence*, Ajay Agrawal, Joshua Gans, and Avi Goldfarb argue that it is useful to understand recent advances in machine learning as better, faster, and cheaper forms of prediction. Prediction is defined in the statistical sense as "the process of filling in missing information" from the available data.[2] In this sense prediction makes it possible to recognize patterns, classify objects or events, and estimate future outcomes. Nearly every impressive AI achievement to date has been a triumph of prediction, including automated route planning, image recognition, text translation, targeted advertising, and beating human champions in games such as Go and Jeopardy!

Prediction technology matters because better prediction is key for effective decision-making, but prediction is only one component of a decision. The process of decision-making has four components: data, prediction, judgment, and action.[3] Military writers often describe this same process as an "OODA loop": observe, orient, decide, act.[4] The first step (data/observe) brings in information about entities and events out in the world. The next step (prediction/orient) combines data about the environment with data in memory to identify patterns and fill in missing information about the current situation. The third step (judgment/decide) makes determinations about what is valuable and desirable. The fourth step (action/act) implements choices to change or maintain the state of the world. These interventions generate more observable data that can be used to adjust future action. In any real organization these phases work in parallel with complicated feedback loops, but the cycle is still a useful heuristic for distinguishing different decision-making functions.

Better, faster, and more efficient machine learning implies that organizations will be able to perform more prediction in the future. Machine prediction doesn't simply replace human prediction; it can also increase the number, accuracy, complexity, and speed of predictions.[5] Yet predictions depend on data, in terms of not only quantity but also quality. If data are biased regarding race, gender, ability, or other features of the task domain, then predictions will be skewed, reflecting those biases.[6] If data are not available on the situation being predicted, one cannot rely on machine predictions to make decisions.[7] If adversaries or competitors manipulate or "poison" the data, then predictions will be useless or worse.[8]

Given enough data of sufficient quality (relevant, unbiased, unmanipulated, and so forth), computers may optimize a utility function better than humans. At the same time, machines lack any understanding about why any of their predictions matter, or the conditions under which they do not. Failure to understand the pragmatic context of human practice may result in AI performance errors and tragedies. In a military context, this may include automated weapons that kill the wrong targets or react prematurely to false warnings. Much of the literature on military AI emphasizes the risks of targeting error and strategic stability that result from automated prediction failure.[9]

When economic goods become less expensive, their complements become more valuable: for example, a drop in the price of bread expands the market for butter. Likewise, as AI makes prediction less expensive, there is more demand for human judgment (in all but a few restrictive categories where it is possible to substitute AI for some well-defined decision

tasks, as discussed below). A larger supply of automated predictions in turn creates opportunities and challenges for organizations and workers. Each prediction requires a human to decide on payoff structures, which are ultimately determined by political, economic, and social values and preferences. The open question for any given AI is who makes the motivating judgments and when. The answers in any given industry or firm will thus have consequences for the labor force.

Generally, if judgment can be prespecified—if the conditions under which specific actions are to be taken can be precisely defined—then it can lead to automation. One act of human judgment could then be formally applied to a large number of situations. This possibility increases the incentives to concentrate decision-making at the top of an organization. The commander commands, and the machines faithfully execute. However, for more complex, nuanced, or idiosyncratic situations, it may not be possible to prespecify judgment. Additional acts of judgment are needed with each prediction.

This means that individuals throughout an organization will need to make more judgment calls, which will tend to decentralize decision-making. In health care, for example, some diagnoses generate clear treatment plans. An AI could provide an accurate diagnosis, the AI-developer could prespecify the implications, and individual doctors would have little role in recommending a treatment. In contrast, some diagnoses and treatments depend on the details of a given patient situation. Diabetes, depression, and hypertension display wide variation in treatment pathways, in part because of different patient needs.[10] With a large number of possible situations and lots of nuance to consider, judgment cannot be feasibly determined in advance, so it is best to show the AI prediction to humans who provide judgment and make a decision.

To summarize the main thrust of the emerging economics literature on AI in organizations, the feasibility of automated prediction depends on the data available to support it and the difficulty of the judgments involved. What does this mean for national security? Because judgment in war is notoriously difficult, and combat information is notoriously unreliable, we expect the human complements to AI to be as important as ever in war.

MILITARY AUTOMATION: MORE HUMAN JUDGMENT

The automation of military tasks has been under way for well over a century. Different parts of the decision cycle (OODA loop) have been automated in turn. First, mechanization in the industrial era improved the ability of humans to move around the battlefield and to deliver fire from a distance (that is, act). Second, digitization of information enhanced the ability of military forces to perceive what was happening far away and to communicate with dispersed units, thereby automating the collection and distribution of data (observe). A third substitution of technology for humans is now under way with the advent of AI, this time affecting the interpretation of information, or what we have described as prediction (orient).

Each historical substitution of machines for humans in different parts of the OODA loop has tended to create stress in other parts that often have led to the need for more human labor. Industrial age militaries used mass to compensate for uncertainty—volleys of musket fire could compensate for the inaccuracy of individual infantryman, and area bombing could compensate for the inaccuracy of unguided munitions—but this required large and destructive forces. Digital age forces have attempted to substitute information for mass by using smaller but smarter forces, armed with guided munitions that can attack targets precisely. However, doing more with less requires more complex command

and control (C2), as the orient and decide steps of the OODA loop depend on human understanding. Militaries today are starting to substitute AI for human prediction to deal with an information overload produced by a revolution in intelligence, reconnaissance, and surveillance (ISR).[11] Yet this substitution will require investing even more organizational effort in the judgment tasks that guide AI performance.

Although the information revolution has significantly improved battlefield awareness, the fog of war remains. In modern data-intensive warfare, personnel struggle to connect systems, negotiate data access, customize software, navigate configuration glitches, and improve information security. In *Information Technology and Military Power*, Jon Lindsay explains why the means of reducing uncertainty, ironically enough, have become new sources of uncertainty. [12] There is little reason to believe that the adoption of AI will reverse these trends, and there are many reasons to believe that it will exacerbate them by increasing the complexity of military operations and data systems.

The essential tasks of military command include the definition of the mission, the management of operations, and the motivation of personnel.[13] The consummate act of judgment, as we have described it here, is the definition of the utility function. In a military context this means the specification of positive objectives and negative limits, or rules of engagement. These are complicated matters of political preferences and moral values where AI is of little help. Managerial tasks that plan and implement military operations, by contrast, involve a mixture of judgment and prediction, so AI might provide some help here. Yet another function of leadership is motivating subordinates and promoting common purposes, values, identities, and meanings throughout the force. This is especially important if commanders expect subordinates to be able to take independent initiative in battle yet remain coordinated with each other and focused on common goals. Leadership enables personnel to understand *why* they are fighting so that they can exercise good judgment on the battlefield. In this respect, AI cannot replace leadership.

Unfortunately, judgment is hard. Judgment in morally fraught situations where the stakes are life or death is even harder. A vast literature on political psychology highlights numerous obstacles to clear and objective judgment.[14] Fraught civil-military relations or strained alliances make consensus judgments even more difficult to achieve.[15] War, moreover, is plagued by missing and unreliable data and difficult or controversial judgments. More and better data and machine learning techniques can provide predictions that support determinations of what matters and why, but they cannot replace judgment.

This does not mean that AI is useless in war, but it does mean that its usefulness will vary with the difficulty of judgment and the availability of data for any given task. Table 1 summarizes the interaction of these factors. AI has a comparative advantage in a world of good and plentiful data with well-defined goals and objectives. Logistics and administration are good candidates because they are well structured by peacetime bureaucracy. Anything that bureaucracies can do well, AI can probably help them to do better. By contrast,

TABLE 1. COMPARATIVE ADVANTAGES OF MACHINES AND HUMANS IN DECISION-MAKING

	Plentiful data	**Sparse data**
Defined judgment	Automation advantage Example: logistics and administration	Automation risk Example: targeting error
Difficult judgment	Human-machine teaming Example: intelligence and planning	Human advantage Example: strategy and command

people have a comparative advantage in a world of poor and missing data with difficult judgment problems. The classic challenges of strategy and command are most pronounced in realms full of fog, friction, controversy, and moral ambiguity.

The mixed categories are more complicated. Automation is risky when data are poor but objective functions are clearly defined. In this case, machines are given specific goals, but limited or unreliable data generate prediction errors and, potentially, counterproductive action. For example, targeting false positives might lead lethal autonomous weapon systems to inadvertently fire on noncombatants or friendly forces. By contrast, in the quadrant labeled "Human-machine teaming," people can use AI to enhance decision-making, but they must exercise a lot of effort to guide and audit AI performance.[16] This is AI for decision support (a complement) rather than mere automation (a substitute). Many intelligence and operational examples fall into this category. ISR and C2 reporting systems generate a large mass of potentially relevant data, but the data become hard to interpret, and metadata are missing or misleading. As the operational environment or mission objectives change, the data on hand may no longer be relevant. Human judgment is needed to understand when this occurs and why.

AI is a general-purpose technology that can be used across the entire military enterprise, but we expect its performance to vary for different military tasks. When data are plentiful and goals are clear, we expect to see full automation of many organizational tasks, military and civilian. This is most likely in the realms of logistics and administration. Unfortunately, the conditions that are most challenging for prediction machines—poor data and difficult judgment—are also the conditions that are most likely during military operations. Many military tasks feature mixed conditions that call for combining the strengths and weaknesses of humans and machines. AI can provide decision support for human beings who make the ultimate judgment on what to value and what to do. In particular, machine learning has the potential to help organizations cope with the information overload created by pervasive ISR and operational complexity. Yet this latest technological substitution for prediction only makes the human complement of judgment more important. More reliance on AI in war makes humans even more important for military power.

ORGANIZATIONAL CHALLENGES: MORE DISTRIBUTED JUDGMENT

Judgment has already become more distributed in military organizations as a result of the ISR revolution, as reflected in larger headquarters elements and more educated personnel.[17] We expect human capital requirements for personnel to continue increasing, as will requirements for continuous collaboration across units, not to mention endless debugging of command and control systems. To the extent that AI systems can make some operations faster or more precise, this will tend to increase the requirements for coordination and synchronization. As more parts of the organization, and different government organizations and coalition partners, come into contact, so will different interpretations, values, priorities, and understandings. These all pose problems for collective judgment, exacerbated by the growing scope and complexity of military operations made possible—ironically—by AI and other information technologies.

We do not expect AI to substitute for human personnel across the board and to lead to the complete automation of war. Instead, we anticipate that human judgment will continue to become more difficult, distributed, and complex in military organizations. Even junior personnel will have to make sense of a developing political situation and tailor operations to the local context.

General Charles Krulak famously highlighted the importance of "the strategic corporal" on the battlefields of the twenty-first century.[18] Krulak argued that operational and political complexity tended to make tactical actions more strategically consequential, for better or for worse. This in turn placed a higher premium on the character and leadership ability of junior personnel—on the quality of their judgment. The importance of judgment among junior personnel has been highlighted by embarrassing public scandals from Iraq and Afghanistan involving targeting errors, prisoner abuse, and other lapses in judgment.

AI systems that further extend the tactical reach of junior personnel in complex political situations will also further increase the importance of judgment. Forward personnel will need to be able to see the predictions from AI systems, assess whether the data that created the prediction are reliable, and make value judgments about how and why automated systems can advance the mission. It is imperative that the human capital requirements of this task not be underestimated. The strategic corporal in the age of AI must be not only a Clausewitzian genius, able to see the way through the fog of war, but also a talented hacker, able to reprogram systems in the face of unforeseen problems. Human judgment is needed not only to direct AI-enabled operations but also to constantly reconfigure and repair complex AI systems themselves. This implies that organizational capacity for judgment is key to the assessment of military power. For instance, China may be deploying a great deal of AI-enabled weaponry that draws on a variety of domestic surveillance data, but does the People's Liberation Army have the requisite technical skill, organizational flexibility, and operational experience to put AI to work in wartime conditions?[19]

STRATEGIC CHALLENGES: MORE CONTROVERSIAL JUDGMENT

War is ultimately a struggle of power and will between rival organizations and societies. Strategic adversaries have incentives to avoid playing to enemy strengths, and to undermine them if possible. In short, if judgment becomes a source of strength for an AI-enabled military organization, then an intelligent adversary will make judgment more difficult.

Many commentators expect AI to make war faster and more volatile. This is possible, but our focus on strategic interaction suggests that AI-enabled conflict also has the potential to become more protracted and less decisive. We anticipate that making particular aspects of military operations more certain will make the entire enterprise less certain.

Reliance on AI could so diffuse responsibility for action that personal responsibility and accountability is undermined.[20] As judgment becomes more distributed, everyone is responsible, and no one is. This can make it harder for the organization to develop a clear collective understanding of what it is doing, and why. Furthermore, judgment is likely to be distributed not only within a military organization but also across the civil-military divide. If judgment is more distributed, then the "unequal dialogue"[21] of civil military relations will necessarily extend into more tactical and technical realms.

This is not a new trend with AI—senior leaders have been able to intervene in decisions about tactical targeting for several decades—but more AI will tend to further blur the distinction between political ends and military means. On one hand, civilians will have to develop a better understanding of the military consequences of the judgments that inform AI prediction. On the other, military personnel will have to become more involved in political conversations about

goals and values to understand what judgments need to be made.

Yet this is also a recipe for politicization from above and below. The distribution of judgment connects more veto players for any given decision. This could promote second-guessing and a reticence to take bold action (decision paralysis), or log-rolling and manipulation to promote parochial organizational and political interests in the name of national security. The strategic corporal becomes a political corporal while the political leader becomes a tactical leader.

The inherent complexity of distributed judgment becomes even more problematic through interaction with a strategic adversary. Strategy links political ends to military means. Ideally, the objectives of military missions should be well defined, clearly related to the national interest, and integrated with other economic or diplomatic tools of statecraft. Strong leadership within military organizations should socialize goals and values among personnel, empowering them to accomplish the mission while respecting normative limits. In this way, judgment becomes a strategic virtue.

If clear judgment is a source of organizational strength, however, then it also becomes a political vulnerability. The clarity and legitimacy of political-military strategies, and the cohesion of enabling coalitions and constituencies, become increasingly attractive targets for adversaries. This has always been true, of course, as adversaries have long been interested in finding ways to attack political will rather than focusing exclusively on fighting capabilities. Yet if material capabilities, such as AI-enabled prediction, become even more dependent on judgment, then it follows that strategies to counter or undermine judgment become even more attractive as well. AI exacerbates longstanding problems of strategic decision-making.

If AI becomes good at optimizing the solution to any given problem, a savvy adversary will attempt to change the problem. In the parlance of AI, the enemy will go beyond the training set.[22] The enemy has incentives to innovate new tactics that are hard to detect, pursue aims that are not anticipated, or poison the data upon which machine prediction relies. Just as the historical evolution of automation has shifted human effort into different phases of the decision cycle, the military advantages of AI will shift adversaries into different strategies to counter them. If prediction takes advantage of the increased availability of data, then adversaries will find ways to manipulate the data.

If prediction machines provide better information, then adversaries will produce more disinformation. If prediction enhances intelligence, adversaries will engage in more devious counterintelligence. If prediction enables more efficient targeting, then adversaries will present more controversial and morally fraught targets.[23] If automated systems operate under tightly controlled rules of engagement, then adversaries will attempt to change the normative frameworks that legitimize the use of force.[24] AI-enabled conflicts have the potential to drag on with ambiguous results, embroiled in controversy and plagued by crises of legitimacy.

We have only begun to outline the organizational and political complexities of automated war here. As AI automates prediction tasks in military affairs, there will be an increased need for human personnel to determine which data to collect, which predictions to make, and which decisions to take. This means that the social complements to AI technology—human capital, leadership, doctrine, policy, and culture—will become key factors sorting winners from losers in security competition. In the end, we expect judgment to become more widely distributed in organizations. At the same time, we expect strategic competition between political actors to make judgment more politically fraught. Greater reliance on AI for military power will make the human element in war even more important, not less.

REFERENCES

1 See, for example, Timothy F. Bresnahan, Erik Brynjolfsson, and Lorin M. Hitt, "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence," *Quarterly Journal of Economics* 117, no. 1 (February 1, 2002): 339–76, https://doi.org/10.1162/003355302753399526; Abigail J. Sellen and Richard H. R. Harper, *The Myth of the Paperless Office* (MIT Press, 2002); David Noble, *Forces of Production: A Social History of Industrial Automation* (New York: Taylor & Francis, 2011).

2 Ajay Agrawal, Joshua Gans, and Avi Goldfarb, *Prediction Machines: The Simple Economics of Artificial Intelligence* (Brighton, Mass.: Harvard Business Review Press, 2018), 24.

3 For an introduction to decision theory see Itzhak Gilboa, *Making Better Decisions: Decision Theory in Practice* (Malden, Mass.: John Wiley & Sons, 2010).

4 All of these words have different connotations in different practices. As we use them, the OODA loop and decision cycle refer to the same basic cybernetic functions of decision; see John Steinbruner, *The Cybernetic Theory of Decision: New Dimensions of Political Analysis* (Princeton University Press, 1974). On the OODA loop construct see James Hasík, "Beyond the Briefing: Theoretical and Practical Problems in the Works and Legacy of John Boyd," *Contemporary Security Policy* 34, no. 3 (December 1, 2013): 583–99 (https://doi.org/10.1080/13523260.2013.839257).

5 Although not necessarily all at the same time. Judgment is needed to make tradeoffs.

6 Jon Kleinberg and others, "Discrimination in the Age of Algorithms," *Journal of Legal Analysis* 10 (December 31, 2018): 113–74, https://doi.org/10.1093/jla/laz001; Bo Cowgill and Catherine E. Tucker, "Algorithmic Fairness and Economics" *Journal of Economic Perspectives* (forthcoming; https://doi.org/10.2139/ssrn.3361280).

7 There is a rich technical literature on these ideas, rooted in the Rubin causal model. For a nontechnical summary, see Agrawal, Gans, and Goldfarb, *Prediction Machines*, 61–65.

8 Agrawal, Gans, and Goldfarb, *Prediction Machines*, 64, 204–5; David Danks, "How Adversarial Attacks Could Destabilize Military AI Systems," *IEEE Spectrum*, February 26, 2020 (https://spectrum.ieee.org/automaton/artificial-intelligence/embedded-ai/adversarial-attacks-and-ai-systems).

9 See, for example, Heather M. Roff, "The Strategic Robot Problem: Lethal Autonomous Weapons in War," *Journal of Military Ethics* 13, no. 3 (July 3, 2014): 211–27; Paul Scharre, "Autonomous Weapons and Operational Risk" (Washington: Center for a New American Security, February 2016); Michael C. Horowitz, "When Speed Kills: Lethal Autonomous Weapon Systems, Deterrence and Stability," *Journal of Strategic Studies* 42, no. 6 (September 19, 2019): 764–88 (https://doi.org/10.1080/01402390.2019.1621174); James S. Johnson, "Artificial Intelligence: A Threat to Strategic Stability," *Strategic Studies Quarterly* 14, no. 1 (2020): 16–39.

10 George Hripcsak and others, "Characterizing Treatment Pathways at Scale Using the OHDSI Network," *Proceedings of the National Academy of Sciences of the United States of America* 113, no. 27 (July 5, 2016): 7329–36. For work that emphasizes how health-care treatments depend on the particular context, see George Hripcsak, David J. Albers, and Adler Perotte, "Exploiting Time in Electronic Health Record Correlations," *Journal of the American Medical Informatics Association* 18,

no. Suppl. 1 (December 2011): i109–15; Marzyeh Ghassemi and others, "A Review of Challenges and Opportunities in Machine Learning for Health," *ArXiv:1806.00388 [Cs, Stat]*, December 5, 2019 (http://arxiv.org/abs/1806.00388).

11 The most well-known program is Project Maven, an AI "tool that could process drone footage quickly and in a useful way." See Kelsey Atherton, "Targeting the Future of the DoD's Controversial Project Maven Initiative," *C4ISRNET*, December 17, 2018 (www.c4isrnet.com/it-networks/2018/07/27/targeting-the-future-of-the-dods-controversial-project-maven-initiative/).

12 Jon R. Lindsay, *Information Technology and Military Power* (Cornell University Press, 2020).

13 Anthony King, *Command: The Twenty-First-Century General* (Cambridge University Press, 2019), 438. See also Martin van Creveld, *Command in War* (Harvard University Press, 1985); Michael A. Palmer, *Command at Sea: Naval Command and Control since the Sixteenth Century* (Harvard University Press, 2005).

14 Janice Gross Stein, "The Micro-Foundations of International Relations Theory: Psychology and Behavioral Economics," *International Organization* 71, no. S1 (April 2017): S249–63.

15 Risa Brooks, "Technology and Future War Will Test U.S. Civil-Military Relations," *War on the Rocks* (blog), November 26, 2018 (https://warontherocks.com/2018/11/technology-and-future-war-will-test-u-s-civil-military-relations/); Erik Lin-Greenberg, "Allies and Artificial Intelligence: Obstacles to Operations and Decision-Making," *Texas National Security Review* 3, no. 2 (2020).

16 Andrew Herr, "Will Humans Matter in the Wars of 2030?," *Joint Force Quarterly* 77 (2015): 76–83; Mary L. Cummings, "Artificial Intelligence and the Future of Warfare," research paper (London: Chatham House, January 26, 2017); Chiefs of Staff Development, Concepts and Doctrine Centre, "Human-Machine Teaming," Joint Concept Note (London: U.K. Ministry of Defence, May 2018); Ryan Mick, "Extending the Intellectual Edge with Artificial Intelligence," *Australian Journal of Defense and Strategic Studies* 1, no. 1 (2019): 23–40.

17 Daniel R. Lake, *The Pursuit of Technological Superiority and the Shrinking American Military* (New York: Palgrave Macmillan, 2019); Andrea Asoni and others, "A Mercenary Army of the Poor? Technological Change and the Demographic Composition of the Post-9/11 U.S. Military," *Journal of Strategic Studies* (published online, January 30, 2020): 1–47 (https://doi.org/10.1080/01402390.2019.1692660).

18 Charles C. Krulak, "The Strategic Corporal: Leadership in the Three Block War," *Marines Magazine*, January 1999, 1–5.

19 Elsa B. Kania, "Chinese Military Innovation in the AI Revolution," *RUSI Journal* 164, no. 5–6 (September 19, 2019): 26–34 (https://doi.org/10.1080/03071847.2019.1693803).

20 Heather M. Roff, "Killing in War: Responsibility, Liability, and Lethal Autonomous Robots," in Fritz Allhoff, *Routledge Handbook of Ethics and War: Just War Theory in the Twentieth Century*, pp. 352–64 (New York: Routledge, 2013).

21 Eliot A. Cohen, *Supreme Command: Soldiers, Statesmen and Leadership in Wartime* (New York: Simon & Schuster, 2002).

22 A lot of practical machine learning relies on supervised learning algorithms to generalize from a dataset of known values. Values that are not represented in the training set may not be classified appropriately in the wild. Unsupervised techniques, which attempt to cluster patterns in observed data, may also make classification errors when new data do not fit the pattern.

23 Erik Gartzke, "Blood and Robots: How Remotely Piloted Vehicles and Related Technologies Affect the Politics of Violence," *Journal of Strategic Studies* (published online, October 3, 2019): 1–31 (www.tandfonline.com/doi/full/10.1080/01402390.2019.1643329).

24 AI has the potential to exacerbate several normative problems that have emerged in the precision targeting era, as discussed by Ryder McKeown, "Legal Asymmetries in Asymmetric War," *Review of International Studies* 41, no. 1 (January 2015): 117–38 (https://doi.org/10.1017/S0260210514000096); Ryder McKeown, "International Law and Its Discontents: Exploring the Dark Sides of International Law in International Relations," *Review of International Studies* 43, no. 3 (July 2017): 430–52 (https://doi.org/10.1017/S0260210517000092).

ABOUT THE AUTHORS

Avi Goldfarb (avi.goldfarb@rotman.utoronto.ca) holds the Rotman Chair in Artificial Intelligence and Healthcare at the Rotman School of Management at the University of Toronto, where he is also a professor of marketing and Chief Data Scientist at the Creative Destruction Lab, a non-profit organization that helps science-based startups to scale. (Disclosure: Some of the start-up companies the Creative Destruction Lab helps aim to serve the defense industry.)

Jon Lindsay (jon.lindsay@utoronto.ca) is Assistant Professor at the Munk School of Global Affairs and Public Policy and in the Department of Political Science at the University of Toronto.

The Brookings Institution is a nonprofit organization devoted to independent research and policy solutions. Its mission is to conduct high-quality, independent research and, based on that research, to provide innovative, practical recommendations for policymakers and the public. The conclusions and recommendations of any Brookings publication are solely those of its author(s), and do not reflect the views of the Institution, its management, or its other scholars.