



Automating the OODA loop in the age of intelligent machines: reaffirming the role of humans in command-and-control decision-making in the digital age

James Johnson

To cite this article: James Johnson (2022): Automating the OODA loop in the age of intelligent machines: reaffirming the role of humans in command-and-control decision-making in the digital age, *Defence Studies*, DOI: [10.1080/14702436.2022.2102486](https://doi.org/10.1080/14702436.2022.2102486)

To link to this article: <https://doi.org/10.1080/14702436.2022.2102486>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 22 Jul 2022.



Submit your article to this journal [↗](#)



Article views: 636



View related articles [↗](#)



View Crossmark data [↗](#)

Automating the OODA loop in the age of intelligent machines: reaffirming the role of humans in command-and-control decision-making in the digital age

James Johnson 

Department of Politics and International Relations, University of Aberdeen, King's College Aberdeen, United Kingdom

ABSTRACT

This article argues that artificial intelligence (AI) enabled capabilities cannot effectively or reliably compliment (let alone replace) the role of humans in understanding and apprehending the strategic environment to make predictions and judgments that inform strategic decisions. Furthermore, the rapid diffusion of and growing dependency on AI technology at all levels of warfare will have strategic consequences that counterintuitively increase the importance of human involvement in these tasks. Therefore, restricting the use of AI technology to automate decision-making tasks at a tactical level will do little to contain or control the effects of this synthesis at a strategic level of warfare. The article re-visits John Boyd's observation-orientation-decision-action metaphorical decision-making cycle (or "OODA loop") to advance an epistemological critique of AI-enabled capabilities (especially machine learning approaches) to augment command-and-control decision-making processes. In particular, the article draws insights from Boyd's emphasis on "orientation" as a schema to elucidate the role of human cognition (perception, emotion, and heuristics) in defense planning in a non-linear world characterized by complexity, novelty, and uncertainty. It also engages with the Clausewitzian notion of "military genius" – and its role in "mission command" – human cognition, systems, and evolution theory to consider the strategic implications of automating the OODA loop.

ARTICLE HISTORY

Received 6 May 2022
Accepted 13 July 2022

KEYWORDS

Artificial intelligence; machine learning; command and control; OODA loop; military theory; mission command

Introduction

This article argues that artificial intelligence (AI) enabled capabilities cannot effectively, reliably, or safely complement – let alone replace – humans in understanding and apprehending the strategic environment to make predictions and judgments to inform and shape command-and-control (C2) decision-making – the authority and direction assigned to a commander (Margaret 2014, 2014; Bostrom 2014; Cantwell Smith 2019).¹ Moreover, the rapid diffusion of and growing dependency on AI technology (especially machine learning (ML)) (Terrence 2018; Domingos 2012, 85–86; Russell and Norvig 2014) to augment human decision-making at all levels of warfare harbinger strategic consequences that counterintuitively increase the importance of human involvement in these tasks across the entire chain of

CONTACT James Johnson  james.johnson@abdn.ac.uk

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

command.² Because of the confluence of several cognitive, geopolitical, and organizational factors, the line between machines analyzing and synthesizing (i.e. prediction) data that informs humans who make decisions (i.e. judgment) will become increasingly blurred human-machine decision-making continuum. When the handoff between machines and humans becomes incongruous, this slippery slope argument will make efforts to impose boundaries or contain the strategic effects of AI-supported tactical decisions inherently problematic and unintended strategic consequences more likely.

The article re-visits John Boyd's observation-orientation-decision-action metaphorical decision-making cycle (or "OODA loop") to advance an objective epistemological critique of using AI-ML enabled capabilities to augment command-and-control decision-making processes (Boulanin 2020).³ Toward this end, the article draws insights from Boyd's emphasis on "Orientation" (or "The Big O") to elucidate the role of human cognition (perception, emotion, and heuristics) in defense planning and the importance of understanding the broader strategic environment in a non-linear world characterized by complexity, novelty, and uncertainty. It also engages with the Clausewitzian notion of "military genius" (especially its role in "mission command"), (Howard and Paret; Grauer 2016; Beyerchen 1992–1993; Biddle 2004; King 2019), human cognition (Kahneman; Ariely, Kahneman et al. 1982; Baron 2008; Robert 2006), and systems and evolution theory (Jantsch 1980; Prigogine and Stengers 1984; Perrow 1999; Jervis 1997; Thomas 1999), to consider the strategic implications of automating the OODA loop. The article speaks to the growing body of recent literature that considers the strategic impact of adopting AI technology – and autonomous weapons, big data, cyberspace, and other emerging technologies associated with the "fourth industrial revolution" (Barno and Bensahel) – in the military decision-making structures and processes (Raska 2021; Talmadge 2019; Goldfarb and Lindsay 2022).

The article contributes to understanding the implications of AI's growing role in human decision-making in military C2. While the diffusion and adoption of "narrow" AI systems have had some success in nonmilitary domains to make predictions and support – largely linear-based – decision-making (e.g. commercial sector, healthcare, and education), AI in a military context is much more problematic (Agrawal et al. 2018; Furman and Seamans 2018). Specifically, military decision-making in non-linear, complex, and uncertain environments entails much more than copious, cheap datasets and inductive machine logic. In command-and-control decision-making, commanders' intentions, the rules of law and engagement, and ethical and moral leadership are critical to effective and safe decision-making in the application of military force. Because machines cannot perform these intrinsically human traits, thus the role of human agents will become even more critical in future AI-enabled warfare (Payne 2021; Johnson 2021). Moreover, as geostrategic and technological deterministic forces spur militaries to embrace AI systems in the quest for first-mover advantage and reduce their perceived vulnerabilities in the digital age, commanders' intuition, latitude, and flexibility will be demanded to mitigate and manage the unintended consequences, organizational friction, strategic surprise, and dashed expectations associated with the implementation of military innovation (Michael 2010; Stephen 2010).

The article is organized into three sections. The first unpacks Boyd's OODA loop concept and its broader contribution to military theorizing, particularly the pivotal role of cognition in command decision-making to understand and survive in *a priori* strategic

environments within the broader framework of complex adaptive organizational systems dynamic non-linear environment. Section two contextualizes Boyd's loop analogy with nonlinearity, chaos, complexity, and system theories with the recent developments in AI-ML technology to consider the potential impact of integrating AI-enabled tools across the human-machine command-and-control decision-making continuum. This section also considers the potential strategic implications of deploying AI-ML systems in unpredictable and uncertain environments with imperfect information. Will AI alleviate or exacerbate war's "fog" and "friction"?

The second section explores human-machine teaming in high-intensity and dynamic environments. How will AI's cope with novel strategic situations compare to human commanders? It argues that using AI-ML systems to perform even routine operations during complex and fast-moving combat environments is problematic and that tactical leaders exhibiting initiative, flexibility, empathy, and creativity remain critical. This section also contextualizes the technical characteristics of AI technology with the broader external strategic environment. Will AI-enabled tools compliment, supplant, or obviate the role of human "genius" in mission command? The final section considers the implications of AI-ML systems for the relationship between tactical unit leaders and senior commanders. Specifically, it explores the potential impact of AI-enabled tools that improve situational awareness and intelligence, surveillance, and reconnaissance (ISR) on the notion of the twenty-first century "strategic corporals" and juxtaposes the specter of "tactical generals."

The "real" OODA loop is more than just speed

John Boyd's OODA loop has been firmly established in many strategic, business, and military tropes (Grant 2013; pp. 600–602; Osinga 2013, pp. 603–624). Several scholars have criticized Boyd's OODA loop concept as overly simplistic, too abstract, and over-emphasizing speed and information dominance in warfare. Critics argue, for instance, that beyond pure granular tactical considerations (i.e. air-to-air combat), the OODA loop has minimal novelty or utility at a strategic level, for instance, managing nuclear brinkmanship, civil wars, or insurgencies (Hasik 2013, pp. 583–99; Storr 2001, 39–45). Some have also lambasted the loop for lacking originality, as pseudoscience (informed by thermodynamics, quantum mechanics, and human evolution) and lacking scholarly rigor. Moreover, the OODA concept struggles to meet the rigorous social science standards of epistemological validity, theoretical applicability, falsifiability, and robust empirical support.

Others argue that these criticisms misunderstand the nature, rationale, and richness of the OODA concept and thus understate Boyd's contribution to military theorizing, particularly the role of cognition in command decision-making (Osinga 2013, 603–624). In short, the OODA concept is much less a rigorously tested epistemological or ontological model for warfaring (which the author never intended), but instead a helpful analogy – akin to Herman Kahn's "escalation ladder" psychological metaphor (Kahn 1965) – for elucidating C2 decision-making cognitive processes and dynamics of commanders and their organizations as they "adjust or change in order to cope with new and unforeseen circumstances" (Boyd 1982). In other words, Boyd's concept is analogous to organizational and individual psychological experiences of the OODA loop in their respective temporal and spatial journeys across the strategic environment (Osinga 2013, 3).

The OODA concept was not designed as comprehensive means to explain the theory of victory at the strategic level. Instead, the concept needs to be viewed as part of a broader cannon on conceptualizations Boyd developed to elucidate the complex, unpredictable, and uncertain dynamics of warfare and strategic interactions (Ibid, p. 173). Moreover, the popularized depictions of the simplified version of Boyd’s concept (or the “simple OODA loop,” see Figure 1) neglects Boyd’s comprehensive rendering of the loop (or the “real OODA loop,” see Figure 2) – with insights from cybernetics, systems theory, chaos and complexity theory, and cognitive science – that Boyd

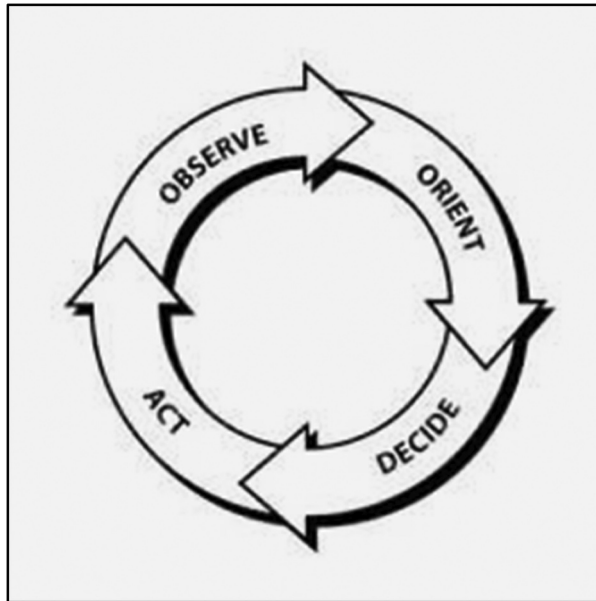


Figure 1. The “simple” OODA loop.

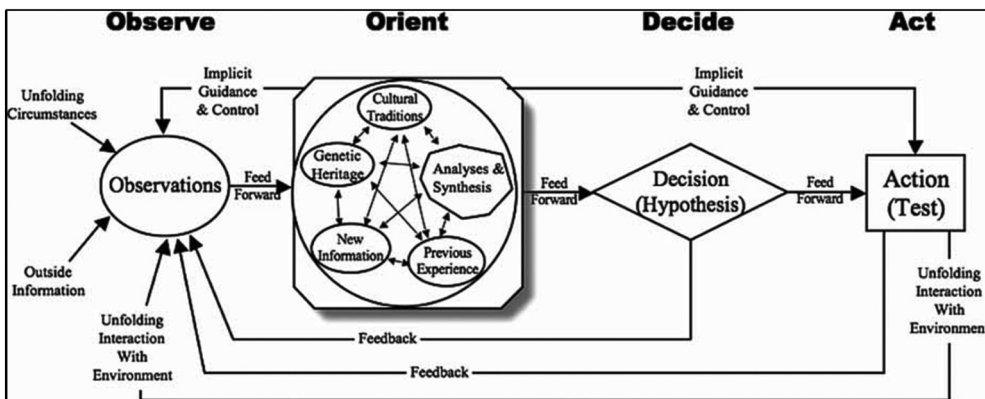


Figure 2. The “real” OODA loop.

introduced in his final presentation, *The Essence of Winning and Losing* (Boyd 1995). The influence of this science and scholarly Zeitgeist most clearly manifests in the OODA loop's vastly overlooked "orientation" element (Ibid, chap. 3 and 4).

The "big o" the center of gravity in warfare

Insights from the cognitive revolution (Polanyi 1969) coupled with Neo-Darwinist research on the environment (Bryant Conant 1964), the Popperian process of adaptation and hypothesis testing (Popper 1968), and complexity and chaos theory (Byrne 1998; Durham 2012), heavily influenced the genesis of Boyd's "orientation" schemata comprising: political and strategic cultural traditions, organizational friction, experience and learning, new and novel information, and the analysis and synthesis of this information (see below) (Gardner 1985; von Neumann 1958; Wiener 1967; Ryle 1966). Viewed this way, "observation" is a function of inputs of external information and understanding the interaction of these inputs with the strategic environment. Boyd wrote, "orientation, seen as a result, represents images, views, or impressions of the world shaped by genetic heritage, cultural traditions, previous experiences, and unfolding circumstances" (Boyd 1987). Without these attributes, humans lack the psychological skills and experience to comprehend and survive in *a priori* strategic environments. The OODA loop concept is a shorthand for these complex dynamics.

Elaborating on this idea, Boyd asserted that orientation "shapes the way we interact with the environment – hence orientation shapes *the way we observe, the way we decide, the way we act*. Orientation shapes the character of present observation-orientation-decision-action loops – while these present loops shape the character of future orientation" (emphasis added) (Ibid., p. 16). In other words, without the contextual understanding provided by "orientation" in analyzing and synthesizing information, "observations" of the world have limited meaning. These assertions attest to the pivotal role of "orientation" in Boyd's OODA decision-making analogy, the critical connecting thread enabling commanders to survive, adapt, and effectively make decisions under uncertainty, ambiguity, complexity, and chaos (Boyd 1987). Boyd's OODA analogy and notions of uncertainty and ambiguous information and knowledge are equally (if not more so) relevant for strategy in the digital information age – combining asymmetric information, mis-disinformation, mutual vulnerabilities, destructive potential, and decentralized, dual-use, and widely diffused AI technology (Kissinger et al. 2021, pp. 166–167; Trinkunas et al. 2020).

According to Boyd, "how orientation shapes observation, shapes decision, shapes action, and in turn is *shaped by the feedback and other phenomena* coming into our sensing or observing window" (emphasis added) (Boyd, p. 5). Thus, the loop metaphor depicts an ongoing, self-correcting, and non-static process of prediction, filtering, correlation, judgment, and action. Scientists including Ian Stewart and James Maxwell have similarly cautioned against assuming that the world is stable, static, and thus predictable (Clerk Maxwell 1969, pp. 440–442). These processes, moreover, contain complex feed-forward and feedback loops (or "double-loop learning") that implicitly influence

“decision” and “action” and, in turn, inform hypothesis testing of future decisions.⁴ For instance, while military theorists often categorize politics as extrinsic to war (i.e. a linear model), the feedback loops which operate from the use of force to politics and from politics to the use of force are intrinsic to war (Roche and Watts; Overy 1981). As a corollary, Boyd argues that C2 systems must embrace (and not diminish) the role of implicit orientation and their attendant feedback loops (Roche and Watts 1991; Overy 1981).

Similarly, Clausewitz argued that the interactive nature of war generates system-driven dynamics comprising human psychological forces and characterized by positive and negative feedback loops, thus leading to potentially limitless versions of ways to seek one-upmanship in competition “to compel our enemy to do our will” (Boyd, p. 24). War, Clausewitz argues, is a “true chameleon,” exhibiting randomness, chance, and different characteristics in every instance (Ibid, p 89). Therefore, it is impossible, as many military theorists tend, to force war into compartmentalized sequential models of action and counter-reaction for theoretical simplicity (Paret 1985, 74–75).⁵ Instead, effective commanders will explore ways to exploit war’s unpredictability and non-linear nature to gain the strategic upper hand (Beyerchen 1992–1993, pp. 74–75). At an organizational level, systems are driven by the behavior of individuals who act based upon their intentions, goals, perceptions, and calculations, an interaction that generates a high degree of complexity and unpredictability (Jervis 1997, 16).

Therefore, a broader interpretation of the OODA loop analogy is best viewed as less a general theory of war than an attempt to depict the strategic behavior of decision-makers within the broader framework of complex adaptive organizational systems in a dynamic non-linear environment (Dooley 1996), pp. 2–3).⁶ According to Clausewitz, general (linearized) theoretical laws used a perspective to explain the non-linear dynamics of war must be heuristic; each war is a “series of actions obeying its own peculiar law” (Von Clausewitz 1976, 80). Colin Gray opined that Boyd’s OODA loop is a grand theory because the concept has an elegant simplicity, a vast application domain, and valuable insights about strategic essentials (Gray 1999, 91).

Intelligent machines in non-linear chaotic warfare

Nonlinearity, chaos theory, complexity theory, and systems theory have been broadly applied to understand organizational behavior and intra-organizational and the inter-organizational dynamics associated with competition and war (Jervis 1997; Perrow 1999). Charles Perrow introduced that organizations can be categorized as either simple/linear (stable, regular, and consistent) or complex/non-linear (unstable, irregular, and inconsistent) systems with tight or loose couplings (Perrow 1999, 90). Non-linear systems do not obey proportionality or additivity; they can behave erratically through disproportionately large or small outputs, exhibiting interactions in which the whole is not necessarily equal to the sum of the parts (Beyerchen 1992–1993, p. 63). The real world has always contained an abundance of non-linear phenomena, for example, fluid turbulence, combustion, breaking or cracking, biological evolution, and biochemical reactions in living organisms (Campbell 1987, pp. 218–262). Perrow’s idea of coupling and complexity is

critical in a system's susceptibility to accidents and failure's speed, severity, and probability. All things being equal, tightly coupled systems such as AI-ML algorithms react faster, unpredictably, with extensive multi-layered interconnections, compared to loosely coupled systems such as universities – which, when everything runs smoothly, afford more time for response and recovery (Perrow 1999, 262).

According to Perrow, the problem arises in cases where a system is both complex and tightly coupled. Whereas complexity generally optimally requires a decentralized response, tight coupling suggests a centralized approach – to ensure a swift response and recovery from accidents *before* tightly coupled processes cause failure. In response to complex and tightly coupled systems, Perrow dismisses the efficacy of decentralizing decision-making at lower levels in an organization (i.e. commanders on the battlefield) because, in complex/tightly coupled systems, potential failures throughout the system are unforeseeable and highly contingent (Ibid., pp. 331–334). Perrow writes that alterations to any one component in a complex system “will either be impossible because some others will not cooperate, or inconsequential because some others will be allowed more vigorous expression” (Ibid., p. 173).

Similar dynamics (randomness, complex interactions, and long and intricate chains) can also be found in ecological systems (Ehrenfeld 1991, 26–39). Because of the intrinsic unpredictability of complex systems, evolutionary biology is – that is, compared to conventional international relations (IR) theories – particularly amenable for understanding IR's such as alliance alignments, the balance of power, signaling deterrence and resolve, assessments of actors intentions, and diplomatic and foreign policy mechanisms and processes (Bernstein et al. 2000, 70). In complex systems, Robert Jervis notes, “problems are almost never solved once and for all; initial policies, no matter how well designed can be definitive [machine generated] solutions will generate unexpected difficulties” (Jervis 1997, 291). Because of the interconnectedness of world politics – imposed on states by the structures of international relations and actors' perceptions of other's intentions and policy preferences – low-intensity disputes can be strategically consequential (Ibid., p 24).

During conflict and crisis, where the information quality is poor (i.e. information asymmetry) and judgment and prediction are challenging, decision-makers must balance these competing requirements. Automation is generally an asset when high information quality can be combined with precise and relatively predictable judgments. Clausewitz's highlights the relationship between the unpredictability of war – caused by interaction, friction, and chance – as a manifestation and contributor of the role of nonlinearity. Clausewitz wrote, for instance, “war is not an exercise of the will directed at inanimate matter . . . or at matter which is animate but passive and yielding . . . In war, *the will is directed at an animate object that reacts*” – and thus, the causes that the outcome of the action cannot be predicted (Von Clausewitz 1976, 149). Inanimate intelligent machines operating and reacting in human-centric (“animate”) environments exhibiting interaction, friction, and chance cannot predict outcomes and thus control. While human decision-making under these circumstances is far from perfect, the unique interaction of factors including, *inter alia*, psychological, social, cultural (individual and organizational), ideology, emotion, experience (i.e. Boyd's “orientation” concept), and luck, give humans a sporting chance to make clear-eyed political and ethical judgments in the chaos (“fog”) and nonlinearity (“friction”) of warfare (Gleick 1987).

The computer revolution and recent AI-ML approaches have made militaries increasingly reliant on statistical probabilities and self-organizing algorithmic rules to solve complex problems in the non-linear world. According to Maxwell, analytical mathematical rules are not always reliable guides to the real world “where things never happen twice” (Clerk Maxwell 1969, p. 440–442). AI-ML techniques (e.g. image recognition, pattern recognition, and natural language processing) inductively (inference from general rules) fill the gaps in missing information to identify patterns and trends; thereby increasing the speed and accuracy of certain standardized military operations, including open-source intelligence collation, satellite navigation, and logistics. However, because these quantitative models are isolated from the broader external strategic environment of probabilities rather than axiomatic certainties characterized by Boyd’s “orientation,” human intervention remains critical to avoid distant analytical abstraction and causal deterministic predictions during the non-linear chaotic war. Several scholars assume that the perceived first-mover benefits of AI-augmented war machines will create self-fulfilling spirals of security dilemma dynamics that will upend deterrence (Michael 2019, pp. 764–788; Kenneth Payne 2021; Johnson 2020).

AI’s exacerbating the “noise” and “friction” of war

Because AI-ML predictions and judgments tend to deteriorate where data is sparse (i.e., nuclear war), low quality (i.e., biased, intelligence is politicized, or data is poisoned or manipulated by mis-disinformation)⁷ (NATO, 2021), military strategy requires Clausewitzian human “genius” to navigate battlefield “fog,” and the political, organizational, and information (or “noise” in the system) “friction” of war (Shaw 1981).⁸ For example, the lack of training data in the nuclear domain means that AI’s would depend on synthetic simulations to predict how adversaries might react during brinkmanship between two or more nuclear-armed states (Johnson 2020). Nuclear deterrence is a nuanced perceptual dance in competition and manipulation between adversaries, “keeping the enemy guessing” by leaving something to chance (Thomas 1960, pp. -199–201).

Therefore, datasets will need to reflect the broader strategic environment military decision-makers face in a military context, including the distinct doctrinal, organizational, and strategic cultural approaches of allies and adversaries. Even where situations closely mirror previous events, a dearth of empirical data to account for war’s contingent, chaotic, and random nature makes statistical probabilistic AI-ML reasoning a very blunt instrument. AI’s predicting and reacting to *a priori* novel situations will increase the risk of mismatch – between algorithmically optimized goals and the evolving strategic environment – and misperception will heighten the risk of accidents (e.g. targeting errors or false alerts) and inadvertent catastrophe. In unpredictable and uncertain environments with imperfect information that requires near-perfect confidence levels, simulations and synthetic data sets are technically limited (Paul and Bracken 2022, 293).

To cope with novel strategic situations and mitigate unintended consequences, human “genius” – the contextual understanding afforded by Boyd’s “orientation” – is needed to finesse multiple flexible, sequential, and resilient policy responses. Jervis writes that

“good generals not only construct fine war plans but also understand that events will not conform to them” (Jervis 1997, 293). Unlike machines, humans use abductive reasoning (or inference to the best explanation) and introspection (or “metacognition”⁹) to think laterally and adapt and innovate in novel situations (Silver 2015, pp. 272–273). Faced with uncertainty or lack of knowledge and information, people adopt a heuristic approach – cognitive short-cuts or rules of thumb derived from experience, learning, and experimentation – promoting intuition and reasoning to solve complex problems (Gladwell and Blink 2005; Herbert 1987, 57–64). While human intuitive thinking heuristics often produces biases and cognitive blind spots, it also offers a very effective means to make quick judgments and decisions under stress in *a priori* situations. AI systems use heuristics derived from vast training datasets to make inferences that inform predictions; they lack, however, human intuition that depends on experience and memory.

Recent empirical studies indicate that the ubiquity of “friction” in AI-ML systems, designed to reduce the “fog” of war and thus improve certainty, can create new or exacerbate legacy, accountability, security, and interoperability issues, thereby generating more “friction” and uncertainty (Michael 2010; Timothy 2013; Nina 2015, pp. 529–553). In theory, there is ample potential for AI-ML tools (e.g. facial and speech recognition, AI-enhanced space satellite navigation, emotion prediction, and translation algorithms) to benefit military operations where vast amounts of disparate information (i.e. “noise”) data and metadata (that labels and describes data) is often incomplete, overlooked, or misdiagnosed. For example, reducing the data processing burden, monitoring multiple data feeds, and highlighting unexpected patterns for intelligence, surveillance, and reconnaissance (ISR) operations that are integral to improving tactical and strategic command decision-making (Timothy 2013; Nina 2015).

In practice, intelligence tasks involve ambiguity, deception, manipulation, and misdisinformation, which require nuanced interpretation, creativity, and tactical flexibility. Like C2 reporting systems, intelligence operations are more of an art than a science, where human understanding of the shifting strategic landscape is critical in enabling commanders to predict and respond to unexpected events, changes in strategic objectives, or intelligence politicization – which will require updated data, thus rendering existing data redundant or misleading. Operational and planning tasks – that inform decision-making – cannot be delegated to AI-ML systems or used in human-machine teaming without commanders being fully mindful of the division of labor and the boundaries between human and machine control (i.e. humans in vs. out of the loop).

Furthermore, human intervention is critical in deciding when and how changes to the algorithm’s configuration (e.g. the tasks it is charged with, the division of labor, and the data it is trained on) are needed, which changes in the strategic environment demand. In other words, rather than complimenting human operators’ linear algorithms trained on static datasets will exacerbate the “noise” in non-linear, contingent, and dynamic scenarios such as tracking insurgents and terrorists or providing targeting information for armed drones and missile guidance systems. Moreover, some argue that AI systems designed to “lift the fog of war” might instead compound the friction within organizations with unintended consequences, particularly when disagreements, bureaucratic inertia, or controversy exists about war aims, procurement, civil-military relations, and the chain of command amongst allies.

“Rapid looping” and the dehumanization of warfare

AI-ML systems that excel at routine and narrow tasks and games (e.g. DeepMind’s StarCraft II and DARPA’s Alpha Dogfight (AlphaStar Team 2019; Newdick 2021) with clearly defined pre-determined parameters in relatively controlled, static, and isolated (i.e. there is no feedback) linear environments – for example, such as logistics, finance, and economics, data collation – are found wanting when it comes to addressing politically and morally strategic questions in the non-linear world of C2 decision-making (Raul 2020). For what national security interests are we prepared to sacrifice soldiers’ lives? What stage on the escalation ladder should a state sue for peace over escalation? When do the advantages of empathy and restraint trump coercion and the pursuit of power? At what point should actors step back from the brink in crisis bargaining? How should states respond to deterrence failure, and what if allies view things differently?

In high-intensity and dynamic combat environments such as densely populated urban warfare – even where well-specified goals and standard operating procedures exist – the latitude and adaptability of “mission command” remains critical, and the functional utility of ML-AI tools for even routine “task orders” (i.e. the opposite of “mission command”) problematic (Kramer 2015). Routine task orders such as standard operating procedures, doctrinal templates, explicit protocols, and logistics performed in dynamics combat settings still have the potential for accidents and risk of life, commanders exhibiting initiative, flexibility, empathy, and creativity are needed (Stephen 2002, 66). Besides, the implicit communication, trust, and a shared outlook that “mission command” imbibes across all levels make micro-management by senior commanders less necessary – that is, it permits tactical units to read their environment and respond within the overall framework of strategic goals defined by senior commanders – thus potentially speeding up the OODA decision cycle. Boyd writes, “the *cycle time increases commensurate with an increase in the level of organization, as one tries to control more levels and issues . . . the faster rhythm of the lower levels must work within the larger and slower rhythm of the higher levels so that the overall system does not lose its cohesion or coherency*” (emphasis added) (Boyd 1987, p. 72).

War is not a game; instead, it is intrinsically structurally unstable; an adversary rarely plays by the same rules and, to achieve victory, often attempts to change the rules that do exist or invent new ones. The diffusion of AI-ML will unlikely assuage this ambiguity in a myopic and likely ephemeral quest, as many have noted, to speed up and compress the command-and-control OODA decision cycle – or Boyd’s “rapid looping.” Instead, policymakers risk being blind-sided by the potential tactical utility – where speed, scale, precision, and lethality coalesce to improve situational awareness – offered by AI-augmented capabilities, without sufficient regard for the potential strategic implications of artificially imposing non-human agents on the fundamentally human endeavor of warfare. Moreover, the appeal of “rapid looping” may persuade soldiers operating in a high-stress environment with large amounts of data (or “data tsunami”) to use AI tools as a means to offload cognitively, thus placing undue confidence and trust in machines – known as “automation bias” (Skitka et al. 1999, pp. 991–1006). Recent studies demonstrate that the more cognitively demanding, time-pressured, and stressful a situation is, the more likely humans are to defer to machine judgments (Mary 2004, pp. 557–562).

NATO's Supreme Allied Command Transformation is working with a team at John Hopkins University to develop an AI-enabled "digital triage assistant" to attend to injured combatants – trained on injury datasets, casualty scoring systems, predictive modeling, and inputs of a patient's condition – to decide who should receive prioritized care during conflict and mass casualty events (e.g. the Russian-Ukrainian conflict) where resources are limited (Graham 2021). On the one hand, AI-enabled digital assistants can make quick decisions intense, complex, and fast-moving situations using algorithms and data (especially much-vaunted "big-data" sources), and arguably, removing human biases that could reduce human error – caused by cognitive bias, fatigue, and stress, for example – and potentially save lives (DARPA, 2022).

On the other hand, critics are concerned about how these algorithms will cause some combatants (allies, adversaries, civilian conscripts, volunteers, etc.) to get prioritized for care over others (Verma 2022). If, for example, there was a large explosion and civilians were among the people harmed (e.g. the Kabul Airport bombing in 2021), would they get less priority, even if they were severely injured? Would soldiers defer to an algorithm's judgment regardless of whether facts on the ground suggested otherwise during an intense situation? Further, if the algorithm plays a part in someone dying, who would be held responsible?

These ethical conundrums are compounded by mounting evidence of bias – such as when algorithms in health care prioritize white patients over black ones for getting care – in AI datasets that can perpetuate biased decision-making (Simonite 2019). In contexts where judgment and decisions directly (or indirectly) can affect human safety, algorithmic designers cannot remove entirely unforeseen biases or prepare AI's to cope with a priori situations. Besides, an optimized algorithm (even if human engineers determine these goals) (Goldfarb and Lindsay 2022, 22, and 30–31) is unable to encode the broad range of values and issues humans care about, such as empathy, ethics, compassion, and mercy – not to mention Clausewitzian courage, *coup d'oeil*, primordial emotion, violence, hatred, and enmity – which are critical for strategic thinking. In short, when human life is at stake, the ethical, trust, and moral bar for technology will always be higher than those we set for accident-prone humans.

Notwithstanding the many vexing ethical and morals about the intersection between people, algorithms, and ethics (Applin 2018, pp. 101–102; Heather 2019, pp. 124–140; Schwarz 2018) introducing non-human agents onto the modern battlefield *in extremis* risks atrophying the vital feedback – or Boyd's "double-looping learning" of empathy, correlation, and rejection – in "mission command" between tactical unit leaders who interpret and execute war plans crafted by generals – or the "strategic corporals" vs. "tactical generals" problem discussed below (US Department of the Army 2012). In this symbiotic relationship, mismatches, miscommunication, or accidents would critically undermine the role of "genius" in mission command on the modern battlefield – combining AI-ML technology, asymmetrical information and capabilities, and multi-domain operations where it is in highest demand.

Butterfly effects, unintended consequences, and accidents

Even a well-running optimized algorithm is vulnerable to adversarial attacks which may corrupt its data, embed biases, or become a target of novel tactics that seek to exploit blind spots in a system's architecture (or "going beyond the training set"), which the AI cannot predict and thus effectively counter (Biggio and Roli, pp. 317–331; Goodfellow et al. 2014). Moreover, algorithms optimized to fulfill a specific goal in unfamiliar domains (i.e. nuclear war) and contexts – or if deployed inappropriately – false positives are possible, which inadvertently spark escalatory spirals (Saalman 2018). Other technical shortcomings of AI-ML systems (especially newer unsupervised models) tested in dynamic non-linear contexts such as autonomous vehicles include: (1) algorithmic inaccuracies; (2) misclassification of data and anomalies in data inputs and behavior; (3) vulnerability to adversarial manipulation (e.g. data-poisoning, false-flags, or spoofing); and (4) erratic behavior and ambiguous decision-making in new and novel interactions (James 2010). These problems are structural, not bugs that can be patched or easily circumvented.

In war, much like other domains like economics and politics, there is a new problem for every solution that AI's (or social scientists) can conceive. Thus, algorithmic recommendations which may look technically correct and inductively sound may have unintended consequences unless they are accompanied by novel strategies authored by policymakers who are (in theory) psychologically and politically prepared to cope with these consequences with flexibility, resilience, and creativity – or the notion of "genius" in mission command discussed below. Chaos and complexity theories can help to elucidate the potential impact of the interaction of algorithms with the real world. Specific interactions in physics, biology, and chemistry, for instance, do not produce more properties as a summing between them; instead, the opposite occurs; that is, a non-additive chemical phenomenon that does not have a numerical value equal to the sum of values for the component parts (Kamo et al. 2015, 20–26). Combining two medical treatments, for example, can produce considerably more than a double dose effect in the patient or even a single but unexpected outcome. Moreover, recent research has found that computers cannot capture the behavior of the complexity of real-world chaotic dynamical events such as climate change (Boghossian et al. 2019, 1–8).

The coalescence of multiple (supervised, unsupervised, reinforcement and deep learning etc.), complex (military and civilian datasets), and tightly coupled and compressed (convolution neural networks, artificial neural networks, Bayesian networks, etc.) ML algorithms, sensitive to small changes in the non-linear real world's initial conditions, could generate a vast amount of stochastic behavior (or "butterfly" effects), thus increasing the risk of unintended consequences and accidents. Therefore, without an improved *technical* understanding of how these systems interact (extracting patterns from the world) and connect with the real world (the "explainability" problem) (Deeks et al. 2019, 1–25) as well as a consensus between militaries (allies/partners and adversaries) and other stakeholders on how AI-ML might be *normatively* aligned with human values, ethics, and goals. This synthesis

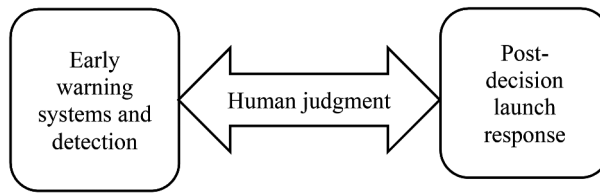


Figure 3. The human-machine decision-making continuum.

will require robust boundaries and controls that delineate the anomaly generating noisy data of the virtual world from the friction and chaos of war (Russell 2019; Singh Gill 2019, pp. 169–179; Sauer 2021, 4–29; Schmitt 2013, 1–37).

Conceptually speaking, boundaries might be placed between AI’s analyzing and synthesizing (prediction) data that inform humans who make decisions (judgment); for example, through recruitment, the use of simulations and wargaming exercises, and training combatants, contractors, algorithm engineers, and policymakers in human-machine teaming. However, the confluence of several factors will likely blur these boundaries along the human-machine decision-making continuum (see Figure 3): cognitive (automation bias, cognitive offloading, and anthropomorphizing) (Watson 2019); organizational (intelligence politicization, bureaucratic inertia, unstable civil-military relations) (Richard 2007; Sagan 1996); geopolitical (first-mover pressures, security dilemma dynamics) (Lieber 2000; Jervis 1978); and divergent attitudes (of both allies and adversaries) to military ethics, risk, deterrence, escalation, and misaligned algorithms (Acton et al. 2017; Morgan et al. 2008; Dobos 2020).

AI’s isolated from the broader external strategic environment (i.e. the political, ethical, cultural, and organizational contexts depicted in Boyd’s “real OODA loop”) are no substitute for human judgment and decisions in chaotic and non-linear situations. Even in situations where algorithms are functionally aligned with human decision-makers – that is, with knowledge of crucial human decision-making attributes – human-machine teaming risks diminishing the role of human “genius” where it is in high demand. Consequently, commanders are less psychologically, ethically, and politically prepared to respond to nonlinearity, uncertainty, and chaos with flexibility, creativity, and adaptivity. Because of the non-binary nature of tactical and strategic decision-making – tactical decisions are not made in a vacuum and invariably have strategic effects – using AI-enabled digital devices to complement human decisions will have strategic consequences that increase the importance of human involvement in these tasks.

How much does a human decision to use military force derived from data mined, synthesized, and interpreted by AI-ML algorithms possess more human agency than a decision executed fully autonomously by a machine? AI researcher Stuart Russell argues that through conditioning what and how data is presented to humans’ decision-makers – without disclosing what has been omitted or rejected – AI’s possesses “power” over human’s “cognitive intake” (Russell 2021). This “power” runs in opposition to one of AI-ML’s most touted benefits: reducing the cognitive load on humans in high-stress environments.

In a recent Joint All-Domain Command and Control (JADC2) report, the US Department of Defense (DoD) proposed integrating AI-ML technology into C2 capabilities across all domains to exploit AI-enhanced remote sensors, intelligence assets, and open sources to “sense and integrate” (i.e. Boyd’s “observe” and “orientate”) information to “make sense” (i.e. analysis, synthesis, and predict) of the strategic environment, so the “decision cycle” operates “faster relative to adversary abilities (US Department of Defense 2022). The JADC2 report obfuscates the possible strategic implications of automating the OODA loop for tactical gains and the illusionary clarity of certainty. In a similar vein, researchers in China’s *PLA Daily* (the official newspaper of China’s People’s Liberation Army) argue that advances in AI technology will automate the OODA loop for command decision-making of autonomous weapons and drive the broader trend toward machines replacing human observation, judgment, prediction, and action (Johnson 2018).¹⁰ While the *PLA Daily* authors stress the importance of training and human-machine “interfacing,” like the DoD’s JADC2, they also omit consideration of the strategic implications of this technologically determined “profound change” (Yan et al. 2022).

Static, pre-defined, and isolated algorithms are not the answer to the quintessentially non-linear, chaotic, and analytically unpredictable nature of war. Therefore, an undue focus on speed and decisive tactical outcomes (or completing the “kill-chain”) in the decision-making loop underplays AI’s influence in command-and-control decision-making activities across the full spectrum of operations and domains. As the US-led Multinational Capability Development Campaign (MCDC) notes: “Whatever our C2 models, systems and behaviours of the future will look like, *they must not be linear, deterministic and static*. They must be agile, autonomously self-adaptive and self-regulating” (Brose 2020).

AI-empowered “strategic corporals” vs. “tactical generals”

US Army Gen. Charles Krulak coined the term “strategic corporal” to describe the strategic implications which flow from the increasing responsibilities and pressures placed on small-unit tactical leaders due to rapid technological diffusion and the resulting operational complexity in modern warfare that followed the information revolution-based revolution in military affairs in the late-1990’s (Charles 1999). Krulak argues that recruitment, training, and mentorship will empower junior officers to exercise judgment, leadership, and restraint to become effective “strategic corporals.”

On the digitized battlefield, tactical leaders will need to judge the reliability of AI-ML predictions, determine algorithmic outputs’ ethical and moral veracity, and judge in real-time whether, why, and to what degree AI systems should be recalibrated to reflect changes to human-machine teaming and the broader strategic environment. In other words, “strategic corporals” will need to become military, political, and technological “geniuses.” While junior officers have displayed practical bottom-up creativity and innovation in using technology in the past, the new multi-directional pressures from AI systems will unlikely be resolved by training and recruiting practices. Instead, pressures to make decisions in high-intensity, fast-moving, data-centric, multi-domain human-teaming environments might undermine the critical role of “mission command,” which connects tactical leaders with the political-strategic leadership – namely, the

decentralized, lateral, and two-way (explicit and implicit) communication between senior command and tactical units. Despite the laudable efforts of the DoD to instill tactical leaders with the “tenants of mission command” to leverage AI-ML in joint force multi-domain environments, the technical and cognitive burden on subordinate commanders – singularly tasked with completing the entire OODA loop – will likely be too great (Nina 2015, 529–553). The US Army’s reliance on electronic communications

Under tactical pressures to compress the decision-making, reduce “friction,” and speed up the OODA loop, tactical leaders may make unauthorized modifications to algorithms (e.g. reconfigure human-machine teaming, ignore AI-ML recommendations),¹¹ or launch cross-domain countermeasures in response to adversarial attacks) that puts them in direct conflict with other parts of the organization or contradicts the strategic objectives of political leaders. According to Boyd, the breakdown of the implicit communication and bonds of trust that defines “mission command” will produce “confusion and disorder, which impedes vigorous or directed activity, hence, magnifies friction or entropy” – precisely the outcome of the empowerment of small-group tactical leaders on the twenty-first-century battlefield was intended to prevent (Boyd 1987, pp. 20–21). This breakdown may also be precipitated by adversarial attacks on the electronic communications (e.g. electronic warfare jammers and cyber-attacks) that advanced militaries rely on, thus forcing tactical leaders to fend for themselves (Johnson and Krabill 2020). In 2009, for example, the Shiite militia fighters used off-the-shelf software to hack into un-secured video-feed from US Predator drones. Militarily advanced near-peer adversaries such as China and Russia, have well-equipped electronic warfare (EW) for offensive jamming (or triangulating their sources for bombardment), sophisticated hackers, and drones to target precision weapons (Freedberg 2017). To avoid this outcome, Boyd advocates a highly decentralized hierarchical structure that allows tactical commanders initiative while insisting that senior command resist the temptation to overly cognitively invest in and interfere with tactical decisions (Boyd 1982, p. 128).

On the other side of the command spectrum, AI-ML augmented ISR, autonomous weapons, and real-time situational awareness might produce a juxtaposed yet contingent phenomenon, the rise of “tactical generals” (Peter 2009). As senior commanders gain unprecedented access to tactical information, the temptation to micro-manage and directly intervene in tactical decisions from afar will rise. Who understands the commander’s intent better than the generals themselves? While AI-ML enhancements can certainly help senior commanders become better informed and take personal responsibility for high-intensity situations as they unfold, the thin line between timely intervention and obsessive micro-management is a fine one. Centralizing the decision-making process – and contrary to both Boyd’s guidance and the notion of “strategic corporals” – and creating a new breed of “tactical generals” micro-managing theater commanders from afar, AI-ML enhancements might compound the additional pressures placed on tactical unit leaders to speed up the OODA loop, and become political, and technological “geniuses.” This dynamic may increase uncertainty and confusion and amplify friction and entropy.

Micromanaging or taking control of tactical decision-making could also mean that young officers lack the experience in making complex tactical solutions in the field, which might cause confusion or misperceptions in the event communications are compromised and the “genius” of “strategic corporals” is demanded. For example, interpreting sensor

data (e.g. from warning systems, electronic support measures, and optronic sensors) from different platforms, monitoring intelligence (e.g. open-source data from geospatial and social media), and using historical data (e.g. data on the thematic bases for producing temporal geo-spatialization of an activity) relating to the strategic environment of previous operations (US Office of Naval Research 2014).

US then-Army Chief of Staff, General Mark Milley, recently stated that the Army is “overly centralized, overly bureaucratic, and overly risk averse, which is the opposite of what we’re going to need in any type of warfare” (Freedberg 2017). Milley stressed the need to decentralize leadership, upend the Army’s deep culture of micromanagement, empower junior officers, and thus strengthen mission command. Further, the breakdown in lateral relations and communication across the command chain would also risk diminishing the crucial tactical insights that inform and shape strategizing, potentially impairing political and military leadership, undermining civil-military relations, and causing strategic-tactical mismatches, which during a crisis may spark an inadvertent escalation.

Tactical units operating in the field using cloud computing technology (or “tactical cloud”) coupled with other AI-enabling sensors and effectors to ensure interoperability, resilience, and digital security, and close to combat situations such as rugged, urban, complex terrains – where logistics and communications lines are put under intense stress – would be well placed to guide and inform strategic decision-making. Thus, senior commanders would indubitably benefit from the speed and richness of two-way information flows as the “tactical cloud” matures; for example, information on topography, the civilian environment, 3D images of the combat distribution of airspace volumes, compliment C2 strike and reconnaissance tracking and targeting for drone swarming and missile strike systems and verifying open-source intelligence and debunking of mis-disinformation (Gros 2019).

Absent the verification and supervision of machine decisions by tactical units (e.g. the troop movement of friendly and enemy forces), if an AI system gave the green light to an operation in a fast-moving combat scenario – when closer examination of algorithmic inputs and outputs are tactically costly – false positives from automated warning systems, mis-disinformation, or an adversarial attack, would have dire consequences (John 2017).¹² Moreover, tactical units executing orders received from brigade headquarters – and assuming a concomitant erosion of two-way communication flows – may not only diminish the providence and fidelity of information received by senior commanders deliberately from their ivory towers but also result in unit leaders following orders blindly and eschewing moral, ethical, or even legal concerns (Fabre 2022). Psychology research has found that humans tend to perform poorly at setting the appropriate objectives and are predisposed to harm others if ordered by an “expert” (or “trust and forget” (Brewer and Crano 1994). As a corollary, human operators may begin to view AI systems as agents of authority (i.e. more intelligent and more authoritative than humans) and thus be more inclined to follow their recommendations blindly; even in the face of information (e.g. that debunks mis-disinformation) that indicates they would be wiser not to.

Whether the rise of “tactical generals” compliments, subsume, or conflicts with Krulak’s vision of the twenty-first century “strategic corporals, and the impact of this interaction on battlefield decision-making are open questions. The literature on systems,

complexity, and nonlinearity suggests that the coalesce of centralized, hierarchical structures, tight-coupled systems, and complexity generated by this paradigm would make accidents more likely to occur and much less easy to anticipate (Perrow 1999, pp. 331–334). The erosion of “mission command” this phenomenon augurs would dramatically reduce the prospects of officers down the chain of command overriding or pushing back against top-down tactical decisions; for example, in 1983, when a Soviet air force lieutenant, Stanislav Petrov, overrode a nuclear launch directive from an automated warning system that mistook light reflecting off clouds for an inbound ballistic missile (Lockie 2018). AI-ML data synthesis and analysis capacity can enhance the decision-making process’s prediction (or “observation”) element. However, it still needs commanders at all levels to recognize the strengths and limitations of AI intuition, reasoning, and foresight (Baum et al. 2018, 19–20). Professional military training and education of military professionals will be critical in tightly integrated, non-linear, and interdependent tasks where the notion of handoffs between machines and humans becomes incongruous.

Conclusion

Whether or when AI will continue to trail, match, or surpass human intelligence and, in turn, compliment or replace humans in decision-making is a necessary speculative endeavor but a vital issue to deductively analyze, nonetheless. That is, develop a robust theory-driven hypothesis to guide analysis of the empirical data about the impact of the diffusion and synthesis of “narrow” AI technology in command-and-control decision-making processes and structures. This endeavor has a clear pedagogical utility, guiding future military professional education and training in the need to balance data and intuition as human-machine teaming matures. A policy-centric utility includes adapting militaries (doctrine, civil-military relations, and innovation) to the changing character of war and the likely effects of AI-enabled capabilities – which already exist or are being developed – on war’s enduring chaotic, non-linear, and uncertain nature. Absent fundamental changes to our understanding of the impact (cognitive effects, organizational, and technical) of AI’s on the human-machine relationship, we risk not only failing to harness AI’s transformative potential but, more dangerously, misaligning AI capabilities with human values, ethics, and norms of warfare that spark unintended strategic consequences. Future research would be welcome on how to optimize C2 processes and structures to cope with evolving human-machine relations and adapt military education to prepare commanders (at all levels) for this paradigm shift.

The article argued that static, pre-defined, and isolated AI-ML algorithms do not answer war’s non-linear, chaotic, and analytically unpredictable nature. Therefore, an undue focus on speed and tactical outcomes to complete the decision-making loop (or Boyd’s “rapid OODA looping”) understates the permeation of AI in command-and-control decision-making across the full spectrum of operations and domains. Insights from Boyd’s “orientation” schemata – particularly as they relate to human cognition in command decisions and the broader strategic environment – coupled with systems, chaos, and complexity theories still have a pedagogical utility in understanding these dynamics. The article argued that a) speeding up warfare and compressing decision-making (or automating the OODA loop) will have strategic implications that cannot be

easily anticipated (by humans or AI's) contained; b) the line between the handoff from humans to machines will become increasingly blurred; and thus, c) the synthesis and diffusion of AI into the military decision-making process increases the importance of human agency across the entire chain of command.

While the article agrees with the conventional wisdom that AI is already and will continue to have a transformative impact on warfare, it finds fault with the prevailing focus of militaries on harnessing the tactical potential of AI-enabled capabilities in the pursuit of speed and rapid decision-making. This undue focus risk blind-siding policymakers without sufficient regard for the potential strategic implications of artificially and inappropriately imposing non-human agents on the fundamentally human nature of warfare. Misunderstanding the human-machine relationship during fast-moving, dynamic, complex battlefield scenarios will likely undermine the critical symbiosis between senior commanders and tactical units (or "mission command"), which increase the risk of mismatches accidents, and inadvertent escalation. *In extremis*, the rise of "tactical generals" (empowered with AI tools making tactical decisions from afar) and the concomitant atrophy of "strategic corporals" (junior officers exercising judgment, leadership, and restraint) will create highly centralized and tight-coupled systems that make accidents more probable less predictable. Paradoxically, therefore, AI's used to complement and support humans in decision-making might obviate the role of human "genius" in mission command when it is most demanded.

Although the article focused on "narrow" task-specific AI, the prospects of AGI – and aside from the hype surrounding what a "third-wave" of genuine intelligence might look like, whether it is even technically possible, and broader debates about the nature of "intelligence" and machine consciousness – are confounding. Conceptually speaking, AGI systems would be able to complete the entire OODA loop without human intervention – and, depending on the goals it is programmed by humans to optimize – would be able to out-fox, manipulate, deceive, and overwhelm any human adversary. Pursuing these goals at all costs and absent an off switch (which adversaries would presumably seek to activate), AGI's would unlikely attach much import to the various ethical and moral features of warfare that humans care about most (Bostrom 2014). In imagined future wars between rival AI's who define their own objectives and possess a sense of existential threat to their survival, the role of humans in warfare – aside from suffering the physical and virtual consequences of dehumanized autonomous hyper-war – is unclear. In this scenario, "strategic corporals" and "tactical generals" would become obsolete, and machine "genius" – however that might look – would fundamentally change Clausewitz's nature of war.

Notes

1. There are three different types of AI: artificial "narrow" intelligence (ANI), artificial general intelligence (AGI) that matches human levels of intelligence, and artificial superintelligence (ASI) (or "superintelligence") that exceeds human intelligence. The debate about when and whether AGI (let alone ASI) will emerge is highly contested and thus inconclusive. This

- article focuses on task-specific “narrow AI” (or “weak AI”), which is rapidly diffusing and maturing (e.g. facial recognition, natural language processing, navigation, and digital assistants such as Siri and Google Assistant).
2. Machine learning (ML) – closely associated with “narrow AI” – is a widely used form of statistical prediction that applies vast training datasets (e.g. metadata and Big-Data) to inductively generate missing information to improve the quantity, and accuracy, complexity, and speed of predictions. There are three main machine learning approaches – supervised, unsupervised, and reinforcement – differentiated by the type of feedback that contributes to the algorithm learning process: (1) in “supervised” learning, an algorithm is trained to produce hypotheses or take specific actions to pursue pre-determined objectives or outputs based on specific inputs (or labels) (e.g. image, text, and speech recognition); (2) in “unsupervised” learning (including neural networks and probabilistic methods) the algorithm has no set parameters or labels; instead, the system learns by finding patterns in the data (e.g. protein folding and DNA sequencing); (3) in “reinforcement learning” the system uses feedback loops to reinforce the algorithms learning process through a reward-and-punishment process (e.g. autonomous vehicles and robotics) to optimize the overall “reward.”
 3. Many AI-enabled weapons have already been deployed, operational, or developed by militaries. Examples include: Israel’s autonomous Harpy loitering munition; China’s “intelligentized” cruise missiles, AI-enhanced cyber capabilities, and AI-augmented hypersonic weapons; Russia’s armed and unarmed autonomous unmanned vehicles and robotics; and US “loyal wingman” human-machine teaming (unmanned F-16 with a manned F-35 or F-22) program, intelligence, surveillance, and reconnaissance (ISR) space-based systems, and various AI-ML-infused command and control support systems.
 4. However, some scholars contend that learning from complexity and uncertainty from past events is ambiguous.
 5. In contrast with Clausewitz’s non-linear approach, other classical military theorists such as Antoine- Henri de Jomini and Heinrich von Bulow adopt implicit linear approaches to manipulate empirical evidence selectively.
 6. Complex Adaptive System (CAS) is a concept that refers to a group of semi-autonomous agents who interact in interdependent ways to produce system-wide patterns, such that those patterns then influence the behavior of the agents. In human systems at all scales, you see patterns that emerge from agents’ interactions in that system.
 7. Several military organizations are testing alternative AI-ML approaches to compensate for the lack of labeled data (i.e. real-world information from the battlefield), which is needed to train existing supervised ML systems. These new approaches combine supervised ML with unsupervised deep-learning approaches, which work with a limited amount of annotated data. “Unsupervised machine learning in the military domain,” *NATO Science & Technology Organization*, 27 May 2021, <https://www.sto.nato.int/Lists/STONewsArchive/displaynewsitem.aspx?ID=642>.
 8. According to information theory, the more possibilities and information a system has, the greater amount of “friction” and “noise” it embodies.
 9. “Metacognition” is a familiar concept to master chess players who can shift their thought concentration to execute moves when faced with complex problems and trade-offs.
 10. Other countries, including the UK (e.g. The British Army’s Project Theia), France (e.g. the French Air Force’s Connect@aero program), and NATO (e.g. IST-ET-113 Exploratory Team) have also started developing and testing AI technology to support C2 and military decision-making; these efforts, however, remain limited in scale and scope. “Unsupervised machine learning in the military domain,” *NATO Science & Technology Organization*, 27 May 2021, <https://www.sto.nato.int/Lists/STONewsArchive/displaynewsitem.aspx?ID=642>.; and “Digging deeper into THEIA,” *The British Army*, 8 July 2021, <https://www.army.mod.uk/news-and-events/news/2021/07/digging-deeper-into-theia/>.; and Philippe Gros, “The ‘tactical cloud,’ a key element of the future combat air system,” *Fondation Pour La Recherche Stratégique*, no. 19:19, 2 October 2019.

11. The judgments (i.e. outputs) of AI-ML algorithms cannot be determined in advance because it would take too long to specify all possible contingencies; thus, human judgment is required to interpret the system's prediction to inform judgment and decision-making.
12. For example, in 2003, a MIM-104 Patriot surface-to-air missile's automated system misidentified a friendly aircraft as an adversary that human operators failed to correct, leading to the death by friendly fire of a US F-18 pilot. Other potential causes of misguided or erroneous algorithmic outputs include: open-source mis-disinformation, adversarial attacks, data-poisoning, or simply the use of outdated, malfunctioning, or biased data training sets.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Dr James Johnson is a Lecturer in Strategic Studies at the University of Aberdeen. James is also an Honorary Fellow at the University of Leicester, a Non-Resident Associate on the ERC-funded Towards a Third Nuclear Age Project, and a Mid-Career Cadre with the Center for Strategic Studies (CSIS) Project on Nuclear Issues. James is the author of *Artificial Intelligence and the Future of Warfare: USA, China Strategic Stability*. His latest book project with Oxford University Press is entitled *AI & the Bomb: Nuclear Strategy and Risk in the Digital Age*.

ORCID

James Johnson  <http://orcid.org/0000-0002-5203-8583>

References

- Acton, James M., et al., eds. 2017. *Entanglement: Russian and Chinese Perspectives on Non-Nuclear Weapons and Nuclear Risks*. Washington, DC: Carnegie Endowment for International Peace.
- Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Cambridge, Mass.: Harvard Business Review Press.
- Applin, Sally. 2018. "They Sow, They Reap: How Humans are Becoming Algorithm Chow." *IEEE Consumer Electronics Magazine* 7 (2): 101–102. doi:10.1109/MCE.2017.2776468.
- Baron, Jonathan. 2008. *Thinking and Deciding*. 4th ed. Cambridge: Cambridge University Press.
- Baum, Seth D., Robert de Neufville, and Anthony M. Barrett. "A Model for the Probability of Nuclear War." *Global Catastrophic Risk Institute, Global Catastrophic Risk Institute Working Paper* 18–1 (March 2018), pp. 19–20.
- Bernstein, Steven, Richard Ned Lebow, Janice Gross Stein, and Steven Weber. 2000. "God Gave Physics the Easy Problems: Adapting Social Science to an Unpredictable World." *European Journal of International Relations* 6 (1): 43–76. doi:10.1177/1354066100006001003.
- Beyerchen, Alan. 1992–1993. "Clausewitz, Nonlinearity, and the Unpredictability of War." *International Security* 17 (3): 74–75. doi:10.2307/2539130.
- Biddle, Stephen. 2004. *Military Power: Explaining Victory and Defeat in Modern Battle*. Princeton, NJ: Princeton University Press.
- Biggio, Battista, and Fabio Roli. "Wild Patterns: Ten Years after the Rise of Adversarial Machine."
- Boghossian, Bruce M., P V. Coveney, and H. Wang 2019. "A New Pathology in the Simulation of Chaotic Dynamical Systems on Digital Computers." *Advanced Theory and Simulations* 2 (12): 1–8. DOI:10.1002/adts.201900125.

- Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- Boulanin, Vincent. edited by. *Artificial Intelligence, Strategic Stability and Nuclear*. Stockholm: SIPRI Publications June 2020
- Boyd, John, *Organic Design for Command and Control* (Unpublished presentation, 1987).
- Boyd, John, *Patterns of Conflict* (unpublished presentation, draft version, 1982).
- Boyd, John, *Strategic Game? and ?* (Unpublished presentation, 1987).
- Boyd, John, *The Essence of Winning and Losing* (Unpublished presentation, 1995).
- Bradshaw, Jeffrey, Marco Carvalho, et al., “Coactive Emergence as a Sensemaking Strategy for Cyber Operations,” *IHMC Technical Report*, October 2012, pp. 1–24.
- Brewer, Marilyn B., and William D. Crano. 1994. *Social Psychology*. New York, NY: West .
- Brose, Christian. 2020. *The Kill Chain: Defending America in the Future of High-Tech Warfare*. New York, NY: Hachette.
- Bryant Conant, James. 1964. *Two Modes of Thought*. New York: Trident Press.
- Byrne, David. 1998. *Complexity Theory and the Social Sciences, an Introduction*. London: Routledge.
- Campbell, David. 1987. “Nonlinear Science: From Paradigms to Practicalities.” *Los Alamos Science* 15: 218–262.
- Cantwell Smith, Brian. 2019. *The Promise of Artificial Intelligence: Reckoning and Judgment*. Cambridge: Massachusetts Institute of Technology Press.
- Charles, C. Krulak. “The Strategic Corporal: Leadership in the Three Block War,” *Marines Magazine*, January 1999.
- Clerk Maxwell, James. 1969. Lewis Campbell and William Garnett “Science and Free Will.” *The Life of James Clerk Maxwell*. New York: Johnson Reprint Corporation.
- Deeks, Ashley, Noam Lubell, and Daragh Murray. 2019. “Machine Learning, Artificial Intelligence, and the Use of Force by States,” 10:1.” *Journal of National Security Law & Policy* 1–25.
- “Developing Algorithms that Make Decisions Aligned with Human Experts.” *DARPA*.3 March 2022
- Dobos, Ned. 2020. *Ethics, Security, and the War-Machine: The True Cost of the Military*. Oxford: Oxford University Press.
- Domingos, Pedro. 2012. “A Few Useful Things to Know about Machine Learning.” *Communications of the ACM* 55 (10): 78–87. doi:10.1145/2347736.2347755.
- Dooley, Kevin. 1996. “A Nominal Definition of Complex Adaptive Systems.” *Chaos Network* 8 (1): 2–3.
- Durham, Susanne. 2012. *Chaos Theory For The Practical Military Mind*. New York: Biblioscholar.
- Ehrenfeld, David. 1991. “The Management of Diversity: A Conservation Paradox.” In *Ecology, Economics, Ethics: The Broken Circle*, edited by Bormann, F. Herbert and Stephen Kellert, 26–39. New Haven: Yale University Press.
- Fabre, Cecile. 2022. *Spying through a Glass Darkly*. London: Oxford University Press.
- Freedberg, Sydney, Jr. “Let Leaders Of the Electronic Leash: CSA Milley, Breaking Defense, 5 May 2017, <https://breakingdefense.com/2017/05/let-leaders-off-the-electronic-leash-csa-milley/>
- Furman, Jason, and Robert Seamans. 2018. “AI and the Economy.” In *Innovation Policy and the Economy*, edited by Lerner, Josh and Scott Stern, 161–191. Vol. 19. Chicago: University of Chicago Press.
- Gardner, Howard. 1985. *The Mind’s New Science, A History of the Cognitive Revolution*. New York: Basic Books.
- Gladwell, Malcolm, and Blink. 2005. *The Power of Thinking without Thinking*. New York, NY: Little Brown and Company.
- Gleick, James. 1987. *Chaos: The Making of a New Science*. New York: Viking.
- Goldfarb, Avi, and Jon Lindsay. 2022. “Prediction and Judgment, Why Artificial Intelligence Increases the Importance of Humans in War.” *International Security* 46 (3): 7–50. doi:10.1162/isec_a_00425.
- Goodfellow, Ian, Jonathon Shlens, and Christian Szegedy. Explaining and Harnessing Adversarial Examples. December 20 2014. *arXiv preprint arXiv:1412.6572*

- Graham, Catherine. "Undergrads Partner with NATO to Reduce Combat Casualties." [*The Hub*]. 20 August 2021.
- Grant, T. Hammond. 2013. "Reflections on the Legacy of John Boyd." *Contemporary Security Policy* 34 (3): 600–602. doi:10.1080/13523260.2013.842297.
- Grauer, Ryan. 2016. *Commanding Military Power: Organizing for Victory and Defeat on the Battlefield*. Cambridge: Cambridge University Press.
- Gray, Colin. 1999. *Modern Strategy*. London: Oxford University Press.
- Gros, Philippe. 2 October 2019. "The "Tactical Cloud," a Key Element of the Future Combat Air System." *Fondation Pour La Recherche Stratégique* (19): 19.
- Hasik, James. 2013. "Beyond the Briefing: Theoretical and Practical Problems in the Works and Legacy of John Boyd." *Contemporary Security Policy* 34 (3, December): 583–599. doi:10.1080/13523260.2013.839257.
- Heather, M. Roff. 2019. "Artificial Intelligence: Power to the People." *Ethics & International Affairs* 33 (2): 124–140.
- Herbert, A. Simon. 1987. "Making Management Decisions: The Role of Intuition and Emotions." *The Academy of Management Executive* 1 (1, February): 57–64.
- Hersman, Rebecca. "Wormhole Escalation in the New Nuclear Age". *Texas National Security Review*. 3:3 Autumn 2020 99–110
- James, A. Russell. 2010. *Innovation, Transformation, and War: Counterinsurgency Operations in Anbar and Ninewa Provinces, Iraq, 2005–2007*. Stanford, Calif.: Stanford University Press.
- Jantsch. "The Self-Organizing Universe, Scientific and Human Implications of the Emerging Paradigm of Evolution Prigogine and Stengers, Order Out of Chaos; Lissack, "Complexity." *Emergence*. 110–126.
- Jantsch, Erich. 1980. *The Self-Organizing Universe, Scientific and Human Implications of the Emerging Paradigm of Evolution*. Oxford: Pergamon Press.
- Jervis, Robert. 1978. "Cooperation under the Security Dilemma." *World Politics* 30 (2): 167–214. doi:10.2307/2009958.
- Jervis, Robert. 1997. *System Effects, Complexity in Political and Social Life*. Princeton: Princeton University Press.
- John, K. Hawley. 2017. *Patriot Wars: Automation and the Patriot Air and Missile Defense Systems*. Washington, DC: CNAS. January.
- Johnson, James. 2018. "China's Vision of the Future network-centric Battlefield: Cyber, Space and Electromagnetic Asymmetric Challenges to the United States." *Comparative Strategy* 37 (5): 373–390. doi:10.1080/01495933.2018.1526563.
- Johnson, James, and Eleanor Krabill. 2020. "AI, Cyberspace, and Nuclear Weapons." *War on the Rocks*. January.
- Johnson, James. 2021. *Artificial Intelligence & the Future of Warfare: USA, China, and Strategic Stability*. Manchester: Manchester University Press.
- Kahn, Herman. 1965. *On Escalation: Metaphors and Scenarios*. New York: Harvard University Press.
- Kahneman, Daniel, Paul Slovic, and Amos Tversky, (edited by). 1982. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Kamo, M, Yokomizo, and H. Yokomizo. 2015. "Explanation of non-additive Effects in Mixtures of a Similar Mode of Action Chemicals." *Toxicology* 1 (335): 20–26. doi:10.1016/j.tox.2015.06.008.
- Kenneth Payne, I. 2021. *Warbot: The Dawn of Artificially Intelligent Conflict*. New York: Oxford University Press.
- King, Anthony. 2019. *Command: The Twenty-First-Century General*. Cambridge: Cambridge University Press.
- Kissinger, Henry A., Eric Schmidt, and Daniel Huttenlocher. 2021. *The Age of AI and Our Human Future*. London: John Murray.
- Kramer, Eric-Hans, and E-H. Kramer. 2015. "Mission Command in the Information Age: A Normal Accidents Perspective on Networked Military Operations." *Journal of Strategic Studies* 38 (4): 445–466. doi:10.1080/01402390.2013.844127.

- Lieber, Keir. 2000. "Grasping the Technological Peace: The Offense-Defense Balance and International Security." *International Security* 25 (1): 71–104. doi:10.1162/016228800560390.
- Lockie, Alex. "The Real Story of Stanislav Petrov, the Soviet Officer Who 'Saved' the World from Nuclear War." *Business Insider*. 26 September 2018
- Margaret, A. Boden. 2014. "AI Its Nature and Future (Oxford: Oxford University Press, 2016)." In David Vernon, *Artificial Cognitive Systems: A Primer*. Cambridge, MA: MIT Press.
- Mary, L. Cummings. "Automation Bias in Intelligent Time-Critical Decision Support Systems," *AIAA 1st Intelligent Systems Technical Conference, 2004*, 557562557562
- Michael, C. Horowitz. 2010. *The Diffusion of Military Power: Causes and Consequences for International Politics*. Princeton, NJ: Princeton University Press.
- Michael, C. Horowitz. 2019. "When Speed Kills: Lethal Autonomous Weapon Systems, Deterrence and Stability." *Journal of Strategic Studies* 42 (6): 764–788. doi:10.1080/01402390.2019.1621174.
- Morgan, Forrest E., et al. 2008. *Dangerous Thresholds: Managing Escalation in the 21st Century*. Santa Monica, CA: RAND Corporation.
- "Multinational Capability Development Campaign (MCDC)." *Final Study Report on Information Age Command and Control Concepts*. 8 February 2019
- Newdick, Thomas. "AI-Controlled F-16s are Now Working as a Team in DARPA's Alpha Dogfights." *The Drive*. 22 March 2021
- Nina, A. Kollars. 2015. "War's Horizon: Soldier-Led Adaptation in Iraq and Vietnam." *Journal of Strategic Studies* 38 (4): 529–553. doi:10.1080/01402390.2014.971947.
- of the Army, Department. "ADP 6-0: Mission Command: Command and Control of Army Forces, Army Doctrine Publication." *US Department of the Army*. 17 May 2012
- Osinga, Frans. 2013. "Getting' A Discourse on Winning and Losing: A Primer on Boyd's 'Theory of Intellectual Evolution.'" *Contemporary Security Policy* 34 (3): 603–624. doi:10.1080/13523260.2013.849154.
- Overy, Richard. 1981. *The Air War 1939-1945*. New York: Potomac Books.
- Paret, Peter. 1985. *Clausewitz, and the State: The Man, His Theories and His Times*. Princeton: Princeton University Press.
- Paul, K. Davis, and Paul Bracken. 2022. "Artificial Intelligence for Wargaming and Modeling." *The Journal of Defense Modeling and Simulation*.
- Perrow, Charles. 1999. *Normal Accidents*. Princeton: Princeton University Press.
- Peter, W. Singer. "Robots and the Rise of Tactical Generals." *Brookings*. 9 March 2009
- Polanyi, Michael. 1969. *Knowing and Being*. London: Routledge.
- Popper, Karl. 1968. *The Logic of Scientific Discovery*. New York.
- Prigogine, Ilya, and Isabella Stengers, *Order out of Chaos* (London: Penguin Random House. (1984). "'Complexity: The Science, Its Vocabulary, and Its Relation to Organizations.'" *Emergence* 1: 1, pp. 110–126. Michael Lissack.
- Raska, M. 2021. "Michael Raska "The Sixth RMA Wave: Disruption in Military Affairs?" *Journal of Strategic Studies* 44 (4): 456–479. doi:10.1080/01402390.2020.1848818.
- Raul, S. Ferreira. "Machine Learning in A Nonlinear World: A Linear Explanation through the Domain of the Autonomous Vehicles." *European Training Network for Safer Autonomous Systems*. 9 January 2020
- Richard, K. Betts. 2007. *Enemies of Intelligence: Knowledge and Power in American National Security*. New York: Columbia University Press.
- Robert, B. Cialdini. 2006. *Influence: The Psychology of Persuasion*. New York: Harper Business. rev. ed
- Roche, James, and Barry Watts. 1991. "Choosing Analytic Measures." *Journal of Strategic Studies* 13 (2): 165–209. doi:10.1080/01402399108437447.
- Russell, Stuart, and Peter Norvig. 2014. *Artificial Intelligence: A Modern Approach*. 3rd ed (Pearson ed. Education: Harlow.
- Russell, Stuart. 2019. *Human Compatible*. New York: Viking Press.
- Russell, Stuart. "'2021 Reith Lectures 2021: Living with Artificial Intelligence," 2021.
- Ryle, Gilbert. 1966. *The Concept of Mind*. London: Hutchinson.

- Saalman, Lora. "Fear of False Negatives: AI and China's Nuclear Posture." *Bulletin of the Atomic Scientists*. 24 April 2018
- Sagan, Scott. 1996. "Why Do States Build Nuclear Weapons? Three Models in Search of a Bomb." *International Security* 21 (3): 54–86. doi:10.2307/2539273.
- Sauer, Frank, and F. Sauer. 2021. "How (Not) to Stop the Killer Robots: A Comparative Analysis of Humanitarian Disarmament Campaign Strategies." *Contemporary Security Policy* 42 (1): 4–29. doi:10.1080/13523260.2020.1771508.
- Schmitt, Michael. 2013. "Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics." *Harvard National Security Journal* 4: 1–37.
- Schwarz, Elke. 2018. *Death Machines: The Ethics of Violent Technologies*. Manchester: Manchester University Press.
- Shaw, Robert. 1981. "Strange Attractors, Chaotic Behavior, and Information Flow." *Zeitschrift der Naturforschung* 36 (1): 80–112. doi:10.1515/zna-1981-0115.
- Silver, Nate. 2015. *The Signal and the Noise: Why so Many Predictions Fail: But Some Don't*, 272–273. New York, NY: Penguin Books.
- Simonite, Tom. "A Health Care Algorithm Offered Less Care to Black Patients." *Wired*. 24 October 2019
- Singh Gill, Amandeep. 2019. "Artificial Intelligence and International Security: The Long View." *Ethics & International Affairs* 33 (2): 169–179. doi:10.1017/S0892679419000145.
- Skitka, Linda J., Kathleen L. Mosier, and Mark Burdick. 1999. "Does Automation Bias Decision-Making?" *International Journal of Human-Computer Studies* 51 (5): 991–1006. doi:10.1006/ijhc.1999.0252.
- Stephen, J. Cimbala. 2002. *The Dead Volcano: The Background and Effects of Nuclear War Complacency*. New York, NY: Praeger.
- Stephen, P. Rosen. 2010. "The Impact of the Office of Net Assessment on the American Military in the Matter of the Revolution in Military Affairs." *Journal of Strategic Studies* 33 (4): 469–482. doi:10.1080/01402390.2010.489704.
- Storr, Jim. 2001. "Neither Art nor Science – Towards a Discipline of Warfare." *RUSI Journal* 146 (2, April): 39–45. doi:10.1080/03071840108446627.
- Talmadge, Caitlin. 2019. "Emerging Technology and Intra-War Escalation Risks: Evidence from the Cold War, Implications for Today." *Journal of Strategic Studies* 42 (6): 864–887. doi:10.1080/01402390.2019.1631811.
- Team, AlphaStar. "AlphaStar: Mastering the Real-Time Strategy Game Starcraft II." *DeepMind Blog*. 24 January 2019
- Terrence, J. Sejnowski. 2018. *The Deep Learning Revolution*. Cambridge: Massachusetts Institute of Technology Press.
- Thomas, C. Schelling. 1960. *The Strategy of Conflict*, 199–201. Cambridge, MA: Harvard University Press.
- Thomas, J. Czerwinski. 1999. *Coping with the Bounds, Speculations on Nonlinearity in Military Affairs*. Washington, DC: National Defence University Press.
- Timothy, S. Wolters. 2013. *Information at Sea: Shipboard Command and Control in the US Navy, from Mobile Bay to Okinawa*. Baltimore, Md.: Johns Hopkins University Press.
- Trinkunas, Harold, Herbert Lin, and Benjamin Loehrke. 2020. *Three Tweets to Midnight: Effects of the Global Information Ecosystem on the Risk of Nuclear Conflict*. Stanford, CA: Hoover Institution Press.
- US Department of Defense, "Summary of the Joint All-Domain Command and Control (JADC2) Strategy," March 2022, <https://media.defense.gov/2022/Mar/17/2002958406/-1/-1/1/SUMMARY-OF-THE-JOINT-ALL-DOMAIN-COMMAND-AND-CONTROL-STRATEGY.PDF>
- US Office of Naval Research, *Data Focused Naval Tactical Cloud (DF-NTC)*, ONR Information Package, 24 June 2014.
- Verma, Pranshu. "The Military Wants AI to Replace Human decision-making in Battle," *The Washington Post*, 29 March 2022.
- Von Clausewitz, Carl. 1976. *On War*, ed. Howard, Michael and Peter Paret. Princeton: Princeton University Press.

- von Neumann, John. 1958. *The Computer and the Brain*. New Haven, CT: Yale University Press.
- Watson, David. 2019. “The Rhetoric and Reality of Anthropomorphism in Artificial Intelligence.” *Minds and Machines* 29 (3): 417–440. doi:10.1007/s11023-019-09506-6.
- Wiener, Norbert. 1967. *The Human Use of Human Beings: Cybernetics and Society*. New York: Avon Books.
- Yan, Ke, Yang Kuo, and Shi Hongbo. 2022. “Human-on-the-Loop: The Development Trend of Intelligentized Command Systems.” *PLA Daily*. March 17.