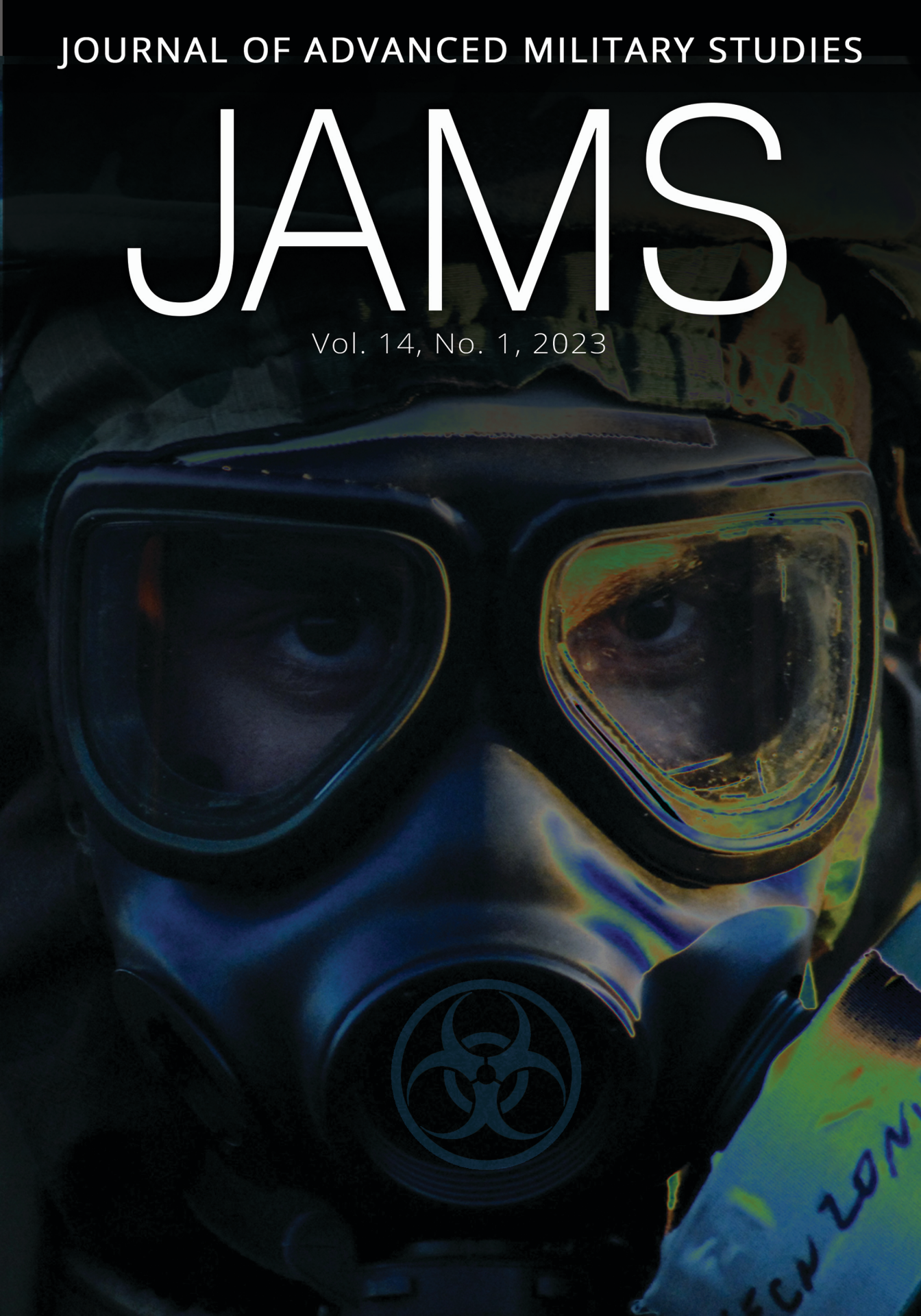
JOURNAL OF ADVANCED MILITARY STUDIES

JAMS

Vol. 14, No. 1, 2023



JOURNAL OF ADVANCED MILITARY STUDIES

JAMS

Published by Marine Corps University Press
2044 Broadway Street | Quantico, VA 22134

Journal of Advanced Military Studies
(Print) ISSN 2770-2596
(Online) ISSN 2770-260X

DISCLAIMER

Established in 2008, MCU Press is an open access publisher that recognizes the importance of an open dialogue between scholars, policy makers, analysts, and military leaders and of crossing civilian-military boundaries to advance knowledge and solve problems. To that end, MCUP launched the *Journal of Advanced Military Studies* (JAMS) to provide a forum for interdisciplinary discussion of national security and international relations issues and how they have an impact on the Department of Defense, the Department of the Navy, and the U.S. Marine Corps directly and indirectly. JAMS is published biannually, with occasional special issues that highlight key topics of interest.

ARTICLE SUBMISSIONS

The editors are looking for academic articles in the areas of international relations, geopolitical issues, national security and policy, and cybersecurity. To submit an article or to learn more about our submission guidelines, please email MCU_Press@usmcu.edu.

BOOK REVIEWS

Send an email with a brief description of your interests to MCU_Press@usmcu.edu.

SUBSCRIPTIONS

Subscriptions to JAMS are free. To join our subscription list or to obtain back issues of the journal, send your mailing address to MCU_Press@usmcu.edu.

ADDRESS CHANGE

Send address updates to MCU_Press@usmcu.edu to maintain uninterrupted delivery.

INDEXING

The journal is indexed by ProjectMUSE, Scopus, ScienceOpen, EBSCO, ProQuest, OCLC ArticleFirst, Defense Technical Information Center, Journal Seek, IBZ Online, British Library System, Lancaster Index to Defense and International Security Literature, and AU Library Index to Military Periodicals.

FREELY AVAILABLE AT
WWW.USMCU.EDU/MCUPRESS

Contents

Vol. 14, No. 1

From the Editors

The nature of the Marine Corps University's work in national security, history, and military studies is always adapting to meet the needs of the Marine Corps in an ever-changing threat environment. Indeed, the very existence of Marine Corps University Press was born of a need to understand and evolve with the growing needs of a professional military educational (PME) institution determined to offer the finest PME to enlisted and officer servicemembers.

In 2006, the U.S. Marine Corps convened a group of current and former military leaders and scholars to evaluate the status of PME within the Corps. Their focus was on four critical elements: faculty, students, curricula, and facilities. At the end of three months of deliberation, General Charles E. Wilhelm's committee made a series of recommendations to ensure the current and future status of Marine Corps education represented the quality of the Corps it supports. One of those findings was for the creation of a publishing house to support the mission of the university and foster research and discourse that crosses the military-civilian domain to innovate, solve problems, and advance knowledge around the world.[1]

Based on the Wilhelm Committee's report, Marine Corps University Press was created in 2008. By 2014, the structure and mission of the press had solidified and efforts were underway to align the organization with the professional accreditation requirements and scholarly rigor within the Association of University Presses. In 2016, MCUP gained introductory status and full regular membership in 2020, joining the ranks of more than 150 esteemed and long-established university presses.

The year saw other significant changes for the press. The *MCU Journal* was rebranded into the *Journal of Advanced Military Studies* (JAMS), a concept more representative of the content and a better tool for acquisition and development that you will see in the following pages. In addition, the press became a stand-alone directorate reporting to the president of Marine Corps University.

During the last 15 years, the Marine Corps has also seen significant change with the withdrawal of American troops from Afghanistan, the release of the *Commandant's Planning Guidance*, and then the vocal debates created by *Force*

Design 2030, *Training and Education 2030*, and *Talent Management 2030*. Those familiar with these documents should see a common thread from Commandant David H. Berger:

- Our identity is firmly rooted in our warrior ethos. This is the force that will always adapt and overcome no matter what the circumstances are.[2]
- We must transform our traditional models for organizing, training, and equipping the force to meet new desired ends, and do so in full partnership with the Navy.[3]
- The current training and education system is not preparing the Marine Corps with the knowledge and range of skills required for the future operating environment. . . . Reimagining training and education requires the application of information-age learning tools.[4]
- Our success on emerging battlefields will depend on our force being more highly trained, cognitively mature, and operationally experienced.[5]

Whether these reports contribute to the evolution of the Marine Corps and their role within the next generation of warfare remains to be seen. In the pages that follow, the authors explore how the United States can remain competitive in various next-generation conflicts, including gray zones; cyber, hybrid and irregular warfare; biological warfare; rethinking doctrine to align with twenty-first century technologies; and other emerging types of conflict and strategies employed by both state and nonstate actors. The authors offer their views from historical, contemporary, and forward-looking perspectives in an effort to encourage discussion but also offer an honest assessment of military capabilities for today and tomorrow.

Unlike previous issues of JAMS, where we typically begin our discussion using a historical lens, for this issue we have flipped the chronological axis and offer first a two-part piece on future battlefields and the singleton paradox concept from America's newest Service—the U.S. Space Force.

Dr. Ben Zweibelson, first in "The Singleton Paradox: Defense Considerations on Complexity, Emergent Technology, and the Complete Disruption of Modern Warfare" and then in "Whale Songs of the Wars Not Yet Waged: The Demise of Natural-Born Killers through Human-Machine Teamings Yet to Come" introduces the concept of a singleton as a future artificial intelligent (AI) entity that could assume central decision making for organizations and societies, creating a *singleton paradox* for security affairs, foreign policy, and military organizations, where AI takes more responsibility (or even total control) in warfare and defense decisions, whether tactical or even strategic.

Italian Army lieutenant colonel Alessandro Nalin and Dr. Paolo Tripodi continue this conversation on the ethics of technology on the battlefield in

"Future Warfare and Responsibility Management in the AI-based Military Decision-making Process." Nalin and Tripodi argue the possible ethical implications of AI integration in the military decision-making process and how the particular characteristics of AI systems with machine learning capabilities might interact with human decision-making protocols, where such machines might make ethical decisions that resemble those made by humans.

The picture of the future created by these initial articles should generate concern in the reader for the future. Dr. Brian R. Price considers some of these higher emotions in "Colonel John Boyd's Thoughts on Disruption: A Useful Effects Spiral from Uncertainty to Chaos." Price draws attention to a series of disruptive actions Boyd lists, including uncertainty, doubt, mistrust, confusion, disorder, fear, panic, and chaos. The author also argues that creativity, when coupled with concepts from the effects spiral, can enhance traditional maneuver and combat, triggering an opponent's collapse without the need for annihilation.

In the wake of the destructive effects of the 2019 global pandemic, the world must also focus on alternative attacks that do not look like traditional tactics. Dominik Juling considers how advances in biotechnology and other transformations of the threat environment will increase the risk that North Atlantic Treaty Organization (NATO) forces will be confronted with a biological, particularly a genetically modified, weapon in "Future Bioterror and Biowarfare Threats for NATO's Armed Forces until 2030." Juling presents a bleak outlook on how the security dimension of pathogens has fundamentally changed in the twenty-first century and will evolve even faster in the future.

Drs. Eldan Haber and Lev Topor shift the conversation from the biological to the virtual sphere in "Sovereignty, Cyberspace, and the Emergence of Internet Bubbles," where the cyber domain emerges as the perfect platform for international struggle for power and influence. In reality, these restricted networks, or *internet bubbles*, are already forming within Russia, China, North Korea, and Iran, but liberal democracies like the United States might be at a severe disadvantage against cyber proxy warfare due to legal and constitutional barriers.

José de Arimatéia da Cruz and H. Chris Tecklenburg take this concept of cyber susceptibility a step further in "The Nationalization of Cybersecurity: The Potential Effects of the *Cyberspace Solarium Commission Report* on the Nation's Critical Infrastructure." The authors provide a historical look at the Department of Homeland Security and the *Cyberspace Solarium Commission Report*, including its recommendations to prevent and respond to cyberattacks.[6] However, in the authors' eyes, many of these recommendations attempt to nationalize cybersecurity.

The final set of articles pulls us away from the theoretical and firmly plants us back on military terra firma and the lessons that could and should be taken from recent events. Dr. Glen Segell, in "Including Africa Threat Analysis in *Force Design 2030*," examines the threat analysis across Africa that should have

been included in *Force Design 2030* for when the Marine Corps is deployed landward or seaward to Africa. The Commandant's strategic guidance document examined the threat analysis presented by China, Russia, Korea, Iran, and violent extremist organizations, but left Africa out of the equation, which represents a notable omission given that previous high-level interventions in Africa were not overtly successful.

Athahn Steinback and Dr. Steven Childs offer "The Deficiency Disparity: The Limit of Systemic Theory and the Need for Strategic Studies in the Power Transition Theory," analyzing parity between nation-states and includes case studies on the Russo-Japanese War, the Afghan War, and ongoing war in Ukraine, demonstrating the decisive influence of power projection, strategy, morale, doctrine, geopolitical constraints, and readiness on conflict outcomes.

Dr. Peter Dobias and Kyle Christensen, in "Intermediate Force Capabilities: Countering Adversaries across the Competition Continuum," then present almost two decades of NATO research into nonlethal, intermediate force capabilities and examine the applicability of these capabilities across the competition continuum as key enablers for NATO operations in the gray zone.

Air Force master sergeant Bonnie L. Rushing and Dr. Kyleanne Hunter continue that threat assessment in "The Human Weapon System in Gray Zone Competition." The authors argue that as the United States considers the next generation of warfare, managing the human weapon system must be a primary concern, particularly how it will shape a military force to successfully compete in gray zone operations with Russia and China, because without that basic comprehension, all technological, doctrinal, or strategic advancements will be useless.

The final article by Dr. Richard Shuster rounds out the discussion using a historical approach and an analysis of military actions in France during World War II. " 'Trying Not to Lose It': Allied Disaster in France and the Low Countries, 1940" highlights the Allies critical point of failure in France and the Low Countries due to a military plan that ignored key tenets of operational art and planning.

As Marine Corps University Press celebrates its 15th year supporting PME and the mission of Marine Corps University, we invite readers to join the conversation, either by following us on our social media accounts or by submitting work that amplifies national security and international relations topics. JAMS offers several such opportunities during 2023–24 with our forthcoming Fall 2023 issue on Russia, NATO, and the conflict in Ukraine, but also with calls for articles on the militarization of space (Spring 2024) and amphibious operations and the evolution of the military Services (Fall 2024). We look forward to hearing your thoughts on these topics and to your future participation as an author, reviewer, or reader. Find us online on our LinkedIn page (https://tinyurl.com/y38oxnp5), at MC UPress on Facebook, MC_UPress on Twitter, and MCUPress on Instagram or contact us via email at MCU_Press@usmcu.edu.

Endnotes

1. Gen Charles E. Wilhelm et al., *U.S. Marine Corps Officer Professional Military Education: 2006 Study and Findings* (Quantico, VA: Marine Corps University, 2006), 42.
2. Gen David H. Berger, *Commandant's Planning Guidance: 38th Commandant of the Marine Corps* (Washington, DC: Headquarters Marine Corps, 2019), 2.
3. Gen David H. Berger, *Force Design 2030* (Washington, DC: Headquarters Marine Corps, 2020), 2.
4. Gen David H. Berger, *Training and Education 2030* (Washington, DC: Headquarters Marine Corps, 2023).
5. Gen David H. Berger, *Talent Management 2030, Update* (Washington, DC: Headquarters Marine Corps, 2023).
6. Angus King and Mike Gallagher, *United States of America Cyberspace Solarium Commission Report* (Washington, DC: U.S. Cyberspace Commission, 2020).

PART I
The Singleton Paradox
On the Future of Human-Machine Teaming and Potential Disruption of War Itself

Ben Zweibelson, PhD

Abstract: Technological innovation has historically been applied in war and security affairs as a new tool or means to accomplish clear political or societal goals. The rise of artificial intelligence posits a new, uncharted way forward that may be entirely unlike previous arms races and advancements in warfare, including nuclear weapons and quantum technology. This article introduces the concept of a singleton as a future artificial intelligent entity that could assume central decision making for entire organizations and even societies. In turn, this presents what is termed a "singleton paradox" for security affairs, foreign policy, and military organizations. An AI singleton could usher in a revolutionary new world free of war and conflict for all of human civilization or trigger a catastrophic new war between those with a functioning singleton entity against those attempting to develop one, along with myriad other risks, opportunities, and emergent consequences.

Keywords: singleton, singularity, transhumanism, artificial intelligence, AI, war studies, security affairs

Machines were first created by humans to shift physical labor from muscle and natural sources (wind, water) and in the last century to shift cognitive labor as well. The history of invention, technology,

Dr. Ben Zweibelson is the director of the U.S. Space Command's Strategic Innovation Group at Peterson Space Force Base, CO. A retired Army infantry officer with combat tours in Iraq and Afghanistan, he earned the Combat Infantryman Badge, Master Parachutist Badge, Pathfinder Badge, Air Assault Badge, the Ranger Tab, four Bronze Star medals, and various awards and citations in his 22 years combined service. He previously worked for U.S. Special Operations Command for seven years, running all design education, theory, and outreach for the Joint Special Operations University. He has a doctorate in philosophy, three master's degrees, and an undergraduate degree in graphic design. He has two design books forthcoming in the summer of 2023.

Journal of Advanced Military Studies vol. 14, no. 1
Spring 2023
www.usmcu.edu/mcupress
https://doi.org/10.21140/mcuj.20231401001

and civilization provides an astounding roadmap from the earliest wheel to today's advanced satellite in geostationary orbit. Woven intricately throughout all of these developments is the never-ending dynamic of organized violence in human affairs. War creates demand for new technology and opportunities, while new technology and opportunity often pave the way for subsequent military applications. Radical shifts in what types of war occurs and how such warfare is exercised often relate to profound technological innovations and scientific discoveries. This relationship is dynamic, but it remains a human-designed, human-controlled one regardless of whether war is waged with edged weapons on horseback or in an all-domain, technologically dense, joint military endeavor against cunning and sophistically enabled adversaries. Today, most discussions on artificial intelligence (AI) and human decision making orbit a specific, tactical, and technologically immediate perspective that may be blinding institutions from greater disruption further afield.[1]

Frequently, too, the rush to implement new constructs exceeds the necessary wisdom and curiosity for how such innovation may require new ways of conceptualizing war, strategy, and military transformation in the wake of such developments.[2] This seems true in how AI is rapidly integrated into modern security applications, doctrine, methods, and tactics *without* essential debate across the military profession on what this means and how future warfare might differ from past historically grounded and institutionally recognized patterns. According to Haridimos Tsoukas,

> Too heavy an influence by the *past* results in incapacity to see what has changed in the present and what is the likely shape of things to come. This is a problem inherent in formal organization. The latter tends to perceive the world predominantly in terms of its own cognitive categories, which are necessarily derived from past experiences. The world may be changing but the cognitive system underlying formal organization, a system that reflects and is based on past experiences, changes slowly.[3]

With the profound developments today in human-machine teaming, the Department of Defense excels at fielding prototypes and experimental gear at the cutting edge of tactical and technological excellence. Where are the deeper discussions on ethics, organizational change, and potential disruption of how war itself is understood? We must "draw our attention to the need to shift from thinking about processes in organization [and knowledge therein] to 'how we should be thinking about processes'."[4] Some sacred cows must be led to conceptual slaughter, if only to prevent such devastation from happening on future battlefields beyond our institutionally regulated limits of understanding.

Human-machine teaming as a concept is hardly a new area for military contemplation, in that the combination of human and machine decision making dates back to mechanical computation machines of the early nineteenth century and analog computers that would eventually aid military cryptology

and ship gun laying in World War II. The Cold War would become defined by a cybernetic drive to reform military operations as interlocking systems of humans and machines obeying formalized rules in a hierarchical cycle of formulated, often rigid decision making.[5] This extends into contemporary warfare where human-machine teaming is a prominent area of focus for new technology, organizational form and function, and operational planning. The origin of civilization is considered to start somewhere between 4000–3000 BCE, and war over 40 centuries features a gradual shift in humans directly controlling and operating analog machines of war toward different variations of human-machine teaming where intelligent machines gain new and potentially dominant roles in whether warfighting effects are applied, including when and where they occur (or do not occur).[6] Played forward, the obvious shift of muscle to analog machine suggests that superior AI may one day exceed human thought on future battlefields, including strategic and organizational considerations. Such an AI development could represent a singleton of defense and security activities for whatever nation or group develops and implements it. How might such a change disrupt future warfare or redefine war itself entirely?

Whether considering the ancient chariots in Greek or Roman warfare or weaponized drones used in the ongoing Russo-Ukrainian War in 2023, these mechanical tools work for the human operator, even if in recent decades the human is repositioned to respond after activities occurred, or program in advance how the team should respond in conditions beyond human comprehension.[7] Weapons strike their target through human senses, whether directly involved or informed by artificial enhancement and depiction. Thus, war remains a human designed, human conceptualized, and ultimately a human experienced and controlled form of organized violence. Today's artificially intelligent war tools remain as such, but tomorrow's may not. It is in this area that vigorous debate must occur, beyond the technological or tactical, and in ways that break with most all established war conventions of battles past. Today's smart weapon requires human decision making, while future ones require new ways of framing, including potentially the entire arrangement of decision making in war. For the first time in history, modern militaries may be at the event horizon of a singleton paradox for war.

What is a singleton, and how does it relate to AI, human-machine teaming, and complex warfare? This article introduces the unfamiliar notion that in the future, potential *general* intelligence machines built for security challenges may force the reconceptualization of what human-machine teams are, including how future wars might be waged or prevented.[8] While this may seem fantastical and wildly impractical for the coming decade, readers might remember that, in 1903, a few short weeks before Orville and Wilbur Wright flew their first flight at Kitty Hawk, North Carolina, the *New York Times* published an informed, rational article declaring airplanes would take another 10 million years for humans to technologically realize.[9] Nick Bostrom, a philosopher focused on technology, first formed the hypothesis that Earth-originating intelligent life

will form a singleton that comprehensively manages everything for civilization. This will be explained in detail, but the primary reason no single government, authoritative dictator, or group has yet to accomplish any cohesive and permanent singleton manifestation is that the human species seems incapable of reaching and employing sufficient intelligence to provide anything but flawed, questionably sufficient, and regularly faulty decision making writ large. Humanity forever exists in the paradox of possessing world-changing curiosity and intellect but coupled with the fact that the species in general routinely demonstrates unintelligent behaviors and frequently makes irrational decisions with dire consequences. Or in the words of an anonymous Yosemite National Park ranger, "There is considerable overlap between the intelligence of the smartest bears and the dumbest tourists" when asked about the difficulty in designing bear-proof garbage cans.[10] Might intelligent machines provide a new on-ramp to decision making and strategic developments beyond human limits?

Readers should take warning that the fantastic and the pragmatic are connected in unusual ways. Modern militaries insist that innovation is key, flexibility in ideas and adaptation are paramount, yet in the same breath many pragmatic professionals then ask for simplicity and uniformity.[11] Complex problems are expected to be "solved" using traditional, linear, mechanistic modes of inquiry espoused in modern doctrine and practice.[12] Henry Mintzberg terms this *machine bureaucracy* where complex reality is inappropriately simplified so that bureaucratic processes are permitted to operate, despite their often-glaring insufficiencies in addressing complex, dynamic systems.[13] Incremental, logical, and linear progress is desired in such institutionalized bureaucracy, so that control remains well in hand of those charged with safeguarding not just the future of the organization, but also the legacy and entrenched belief systems that represent identity and purpose.[14] This "problem-solution" logic dismisses complexity so that courses of action paired with optimized analysis imposes a simplification of reality instead.[15] Elizabeth Kinsella elaborates:

> Practitioners set the problems that they go about solving, and such problem setting is a form of worldmaking that often falls outside the realm of the technical knowledge learned in professional schools. Problem setting often begins when one's usual understanding of the world bumps up against a disorienting dilemma or problematic situation that falls outside of one's usual frames. . . . In this way the practitioner is viewed *as setting the problem within a world of his or her own making* [emphasis added].[16]

Innovation in military organizations is expected to be accomplished in largely the same way that traditional planning occurs, and all innovative activities must also comply with most all institutionally protected and coveted content so that the organization does not experience disruption or uncertainty while changing.[17] Yet, neither innovation (nor planning in complexity) works this way.[18] Neither does the arrival of a new paradigm for war, science, or oth-

er discipline where the institutional frame for reality is defeated and replaced with an entirely novel one. Thomas S. Kuhn wrote of new scientific paradigms radically disrupting the legacy one, replacing it entirely, and in the wake of that disruption, witnessing a migration of people that adapt to the new paradigm while those unwilling or unable to do so fade off into irrelevance.[19] The last cavalry charge occurred at least one generation of warfighters too late, and it was not led by disruptive innovators. The practical debates on AI and human-machine teaming are necessary for today's current conflicts. Yet, to engage in where future conflicts might radically depart from established norms, militaries must move away from the practical to the fantastic area of AI and human-machine teaming debates. Only in the abstraction of the fantastic might new insights and illumination occur that provide clearer yet novel perspectives for tomorrow's unrealized conflict.

AI today remains *narrowly* exceptional, in that intelligent machines can outperform humans in very specific tasks, such as analyzing thousands of images in seconds to isolate a specific facial pattern or identifying and targeting the trajectory and point of origin of a mortar fired at friendly forces so that immediate counterbattery occurs in seconds. Machines now are superior to humans in chess, trivia games, and many other areas of mention, yet machines remain utterly wedded to the coding that provides them select (narrow) super-human abilities, if only those same humans refine and update the code accordingly. Changing one simple rule in *Jeopardy!* would eliminate IBM's Watson from the contest until programmers adjusted the code. Today's intelligent war machines remain entirely dependent on human operators and designers. AI systems provide amazing, game-changing capabilities in strictly narrow applications in warfare, where the human decision makers, operators, and machine designers largely remain completely in control.[20] Hence the term *human-machine teaming* positions the human first in order of importance.

Security affairs and war studies discussions abound with supposed "game-changing" concepts, yet all too often these immediately become hyperbole or fixate on isolated technological developments within warfare with the aforementioned disregard for complexity theory and overemphasis on institutionalized, largely Newtonian war frames.[21] If one assumes the metaphor of "game" for how technology is positioned in advancing or changing warfare to the advantage of the technological innovator, the deeper implication is that new technology permits the user to gain new control or dominance over a lesser equipped opponent who is fighting according to some shared rules and patterns.[22] Artificial intelligence, once able to reach levels of equivalency or superiority with how humans demonstrate general intelligence, may end up changing the game in ways the creators will not recognize. This could put both human competitors into situations where control and dominance are no longer exercised in the historical patterns of past conflict. Indeed, truly advanced AI might for all intents and purposes break the human war paradigm entirely. This is a radical, likely absurd notion, particularly for the pragmatics and realists within

the military institution. Then again, in 1903 before the Wright brothers made history, many readers of the *New York Times* would likely be seen as rational, reasonable, and well-informed people able to distinguish clearly between what is potentially game changing, and what could not possibly happen for another 10 million years.

Defining the Singleton

Bostrom introduced the concept of a singleton not as social commentary on political systems, ideologies, or why most governments are perpetually dysfunctional bureaucracies on the edge of corruptive ruin. He wanted to pair the failure of optimized societal decision making with that of AI and demonstrate how technology might open a Pandora's box unlike anything previously experienced. This is not to be construed as technological fearmongering, yet humanity does illustrate a strong pattern of developing and implementing new ideas without realizing the consequences. New intelligent weapons designed to augment the human operator represent a familiar manner to extend past mechanical, analog war tools for the warrior in battle. New intelligent systems that can form strategies, war theories, formulate diplomacy, and manage entire defense departments in superior ways beyond the most intelligent human is entirely different. Such developments may be multiple decades away or possibly closer than assumed. That there is no serious military debate on such matters is potentially more terrifying.[23]

Bostrom suggests that a singleton could manifest in a political or ideological group that offers a new world order that actually succeeds in some form, yet Bostrom's original singleton construct suggests that standard human intelligence and abilities for an individual dictator or group of leaders has thus far been proven insufficient. Throughout more than 40 centuries of human history, there has yet to be an ideology, culture, belief system, or group of people capable of executing a singleton beyond that of an empire, nation-state, or some sort of organization that has an expiration date as well as an inability to extend fully to all of civilization.[24] Arguably, some individuals or groups have shown limited singleton abilities to select populations and geographical areas over periods of time, but none have been enduring nor has any entity assumed productive unification of the entire human civilization. Humans with current cognitive and communicative abilities just have not yet realized or implemented any meaningful (or enduring) singleton. Bostrom illustrates that "[a singleton's] defining characteristic . . . is some form of agency that can solve all major global coordination problems. It may, but need not, resemble any familiar form of human governance."[25] While human-machine teaming is typically framed only in tactical military contexts, a singleton is the manifestation of such an arrangement at the grand strategic, national, or ultimately internationally collective level for civilization. This is systemic teaming at the level of networks, ecosystems, and entire species at full realization.

Bostrom explains that a singleton is an entity that becomes the single

decision-making authority at the highest level of human organization. This assumes the entirety of human civilization, often confined to Earth for the near term.[26] Such an entity is considered "a set with only one member," yet this requires further information.[27] First, a singleton is something that is able to take total control of human civilization, or at least those that are reachable and able to be controlled, so that a world order is instituted by the design of the singleton and executed through complete realization. An AI singleton would, if able to reach *general* and then superintelligence, potentially self-develop into an intellect hundreds of thousands of times beyond even the smartest human. Suggesting such an entity would be able to take the mantle of controlling all of civilization raises all sorts of ethical, moral, and existential questions that Bostrom addresses in his book in myriad ways. Ultimately, were such a powerful intellect developed, humans would face significant challenges in containing it, utilizing it effectively, and also anticipating adversarial attempts to develop their own singleton entity first for their own interests or security goals. Such a competition might dwarf the space and nuclear races, given the long-term potential impacts. However, the glide path from the arrival of an AI singleton entity and this realization of total implementation/exercised control is an area requiring further serious research, debate, and strategic contemplation. A singleton entity, by virtue of assumed total control of all aspects of a society, would directly control all security apparatuses, including nuclear strategies.[28]

In *Superintelligence*, Bostrom introduces this concept of a superintelligent singleton, potentially an artificial intelligence, but not necessarily. Superintelligence is defined as "any intellect that greatly exceeds the cognitive performance of humans in virtually all domains of interest."[29] Understandably, many military professionals when considering AI and security applications leap to this concern of a superintelligent AI creation becoming a threat to the human creators, and thus untrustworthy for any critical or existential systems such as nuclear weaponry as well as control of essential services such as power or information. Bostrom makes compelling arguments that a singleton could become realized through some sort of superintelligent entity, whether an artificial intelligence system, a genetically modified human with cognitive abilities so advanced it may no longer qualify as the same species (the first *Supra sapien*, perhaps), or potentially a cybernetically enhanced human.[30]

These ideas seem fantastic, and with the current state of artificial intelligence development in 2023, they likely are. However, this may not be the case in less than a century depending on technological advances in computing, particularly quantum computing as well as genetics, nanotechnology, and robotics. A singleton with superintelligence would likely conceptualize on a level incomprehensible and alien to even the smartest humans. This is nicely summarized by the fictional superintelligent character Dr. Manhattan in *The Watchmen* who remarked: "The world's smartest man poses no more threat to me than does its smartest termite."[31] This suggests that whether the superintelligent singleton arrives in the form of an AI system, a genetically enhanced human (or humans),

or cybernetically enhanced humans—these all are areas of significant military research and development at a primitive level of singleton potential.[32] They are decades if not centuries away, yet within popular culture and science fiction stories these concepts are already deep within the societal zeitgeist as an instrument of fear and distrust.

Of significance to this article is the deeper question of whether our modern framing of war and warfare is insufficient for what a potential superintelligent singleton might produce in security affairs. For recorded human history, and particularly in the last three centuries of Western scientific development, military philosophers have granted war a natural, timeless, and universal ordering (albeit a chaotic, passionate, dynamic one for some theorists), with warfare a perpetually changing character where scientific methods could take hold and offer some reliable sense of direction in the fog and friction.[33] War was not always conceptualized as such, nor today do all societies and competitors subscribe to the same war paradigm.[34] While it is highly controversial to challenge such base premises in contemporary American and partnered military communities, a minority of theorists do so. Unfortunately, such debates often occur well outside established military training, doctrine, or educational settings. What is most significant here is not whether one human-designed war frame is superior or inferior to another, but that all of them are of human design, and all of war is a human creation. Given that all war theory is conceptualized by human minds, is there not a potential that AI in a future and potentially advanced configuration might develop dissimilar concepts? Furthermore, were an AI singleton to develop new war theory and practices, could human minds fully appreciate them if they required either intelligence beyond the human limits, or merely nonhuman thinking to forge a conceptual path to them? If this is the case that AI entities would be alone in comprehending and directing such new concepts, how would human operators continue to participate in some sort of decision-making loop of human and machine teaming?

There are many technological, ethical, moral, legal, and strategic questions concerning AI and weaponization, yet most of them orient toward human beings still able to make decisions within the loop, or perhaps "on top of the loop" where AI can produce lethal effects based on previously established human parameters and limits designed by humans for machines to rapidly operate within.[35] The singleton offers the profound possibility that this entire shared, socially constructed notion of war could be shattered and eclipsed by something beyond our reasoning and comprehension. Regular AI may challenge both the assumed character and nature of future war, while a superintelligent singleton might break it completely.[36] Even if this were to occur, would humans be cognizant of such developments, or would they be satisfied with the tangible effects of either successful security affairs or some elimination of violence and conflict?

Incomplete and Misleading Notions of Singletons

Singletons are popular in modern entertainment, whether in science fiction

stories, movies, television, or other similar modes of entertainment. Indeed, advanced societies grapple with the paradoxical challenges of technology and prosperity and whether such designs are doing more harm than good. While existential fears abound with human-controlled weapons of mass destruction, the fear that something nonhuman might be even more existentially dangerous is where killer robots and inhuman logic taken to absurdity evokes great science fiction horror stories. Again, humans as a species feature a long and complex relationship with technology dating back to the earliest recorded history, in that contextually any cutting-edge technological development inspires awe as well as fear. The ancient Greeks used the story of Icarus inventing a flying contraption to warn of recklessness and impulsive behaviors regarding technological developments that distract society too far from established norms and values. Icarus, in his own exuberating thrill of flying it, gets too close to the sun and perishes. Today, when Boston Dynamics uploads new videos of their Atlas robot online, Twitter feeds are flooded with admiration and also snarky comments on the end of human civilization at the hands of robot overlords.[37]

Modern technologically inspired stories extend from far older myths and narratives that draw from basic human desires, values, and wants.[38] Not all industrialized societies feel this way, as notably Japanese culture readily embraces advanced technology, robotics, and significant human-machine teaming with little of the technophobia found in American pop culture such as *The Terminator*, *Wargames*, *Star Trek*, and fantasy cartoons such as *Rick and Morty*.[39] Indeed, Japan is often far ahead of the rest of the world in experimenting with AI and robots in real-world applications, whether with AI engagements in hotels, nursing homes, or for a host of social applications in the home.[40] However, in much of the Anglo-Saxon world of largely Western European origin and design, there seems to be a more pronounced fear of and fetish about what the future may yield with respect to AI, robots, and similar technology. The possible reason for this pattern suggests further research is needed outside the scope of this article. Singletons are of great military strategic concern, yet due to cultural and social biases potentially stemming from these other areas, military discourse is often stymied from properly contemplating such futures. Killer robots get chuckles from the military audience, and they move onto more important affairs of immediate, tactical, and short-term technological consideration. This requires rectification so that clear, serious debate occurs on the bigger, long-term picture for future conflicts.

A singleton entity is frequently confused with a singularity, which also is popular in science fiction, futurism, and technological discourse. A singularity, first introduced by mathematician Vernor Vinge and popularized into mainstream entertainment by Ray Kurzweil, is considered a game changing, evolutionary moment where the natural human species, developed over thousands of generations through evolutionary, gradual change would suddenly gain new shortcuts that no other creatures on Earth might entertain. Genetic modification, nanotechnology, cybernetic implants, networked augmentation, and many

other radical options, if fully developed, could provide unfathomable new ways for humans to evolve into an entirely new category of existence. Should people gain any ability to reconfigure or modify their genetic structures, molecules, or biological abilities beyond even the most gifted natural configurations thus far, they might transform into a superintelligent, infinitely enhanced, and possibly nonbiologically based technological fused entity.[41]

A singularity introduces the concept of *transhumanism*, where at a biological, physical, political, sociological, and ultimately a philosophical level, humanity might evolve beyond the slow, clunky genetic and environmental soup of existence as organic, carbon-based life forms. The ethical, moral, and legal concerns abound here but also there are clear security and defense considerations. Should one nation find genetic manipulation for creating super soldiers unethical, what happens if a future adversary rejects that conclusion if only to enjoy a significant advantage on a future battlefield? If a natural soldier is psychologically and biologically limited to effectively controlling 3–4 combat systems in support of their battlefield role, but a cybernetically enhanced soldier (even surgically altered) can control 300–400 systems with ease, how will different societies debate these challenges prior to catastrophic foreign policy debacles?[42]

There are sinister aspects of such radical change to the fundamental building blocks of what the human species can and cannot do. This also has been articulated in religious debate as "Apocalyptic AI."[43] A *singularity* is when machines with sufficient artificial intelligence are able to teach and improve themselves, with variations of a singularity including human-machine teaming, hybridization, or potentially solely a machine-driven acceleration beyond humans.[44] Technology with advanced AI could unlock entirely revolutionary developments where humans begin to exist exclusively in virtual or augmented realities well beyond simple metaverse discussions offered today by social media giants.[45] The technological progress in this march toward a singularity is not linear but exponential, meaning the estimates on when a singleton might be reached is also subjected to this rapid shift.[46]

Beyond the singularity, human existence might be challenged in nearly all aspects, from whether biology can be manipulated genetically, enhanced through cybernetics, or even transmitted into pure informational form and function outside the limits of organic life. This may sound radical and far-off, but AI and related research is ongoing where such ideas are moving into the theoretical from the merely hypothetical.[47] Perhaps each of these concepts might arrange on some sort of technological pathway, with the metaverse being an early phase where organically unmodified humans might increasingly spend more of their lives in a sophisticated virtual and/or augmented reality, and potentially modified users might gain unprecedented access and immersion beyond the natural configured species users. Super-enabled humans would be potentially reaching this singularity concept, and either they would gain access to some superintelligence level that could provide unparalleled reasoning on security and

governance or the development of general intelligence AI systems might beat them there instead. Indeed, if superintelligence and the option to operate as a singleton entity for all of civilization is some sort of finish line, the race might be waged between a host of strange characters.

Modified humans with super cognitive (and physical) abilities might win or lose out to cybernetically enhanced human-machine entities. Or they all might lose the race to natural born human engineers and scientists that design the first general intelligence AI system capable of boosting itself to thousands of times more powerful than the smartest human intellect, perhaps beyond even what a modified individual human might be capable of. These fantastic concepts again sound too far-off and abstract, but such a race is already underway, if only beginning, and the race is one waged between various nations that are in competition and have rival (or incommensurate and antagonistic) security aspirations.[48] Yet, a superintelligent human individually is not automatically a singleton, nor is an advanced technological system of multiple humans individually and collectively engaging a singleton either. The singleton hypothesis reflects the centralized authority for all significant decision making into one entity. In any configuration where various humans (superintelligent or not) exercise different judgments or ability to change the direction outside of the authority vision, one lacks the singleton manifestation. Siri and Alexa may know all of someone's browsing and shopping habits and make highly informed suggestions to people, but they still serve the human operator who remains in charge.

Thus, singletons are not to be confused with advanced, networked AI nor with a powerful, sophisticated internet that might be termed a metaverse. Even a network of super-enhanced human users in the metaverse, if still each independent, could form sophisticated societies or political configurations, but they would not be a true singleton.[49] Humans, whether organically natural or highly modified would still oversee society with humanity guiding it in new directions according to new realizations of human existence and expression beyond contemporary (and still largely analog) frames. Singletons would, if one emerged from human technological designs, engage *positively* or *negatively* as a superintelligent entity created by nonenhanced creators. Even the notions of positive and negative are grounded in human values and nested in human conceptualization of which the singleton might transcend in ways incomprehensible.[50] Which values apply to what is "good" or "bad" in such complex, systemic contexts?

In other words, the human designers might produce an AI capable of understanding things the designers could not, placing them in a subservient role cognitively whether they wanted this or not. The tool would become superior to the operator, and the designed means to an end would gain the unprecedented ability to exceed the original end. This is where a means to an original end may no longer connect, as the AI would create new ends of its own design outside of the human creator. The tool designed for one purpose reconfigures toward an unrealized one that even the tool creator cannot fathom. This is where most

science fiction and entertainment falls short, or simply confuses the singleton with other aspects of the metaverse, artificial intelligence, swarm logic, transhumanism, or simply technophobia. Most all science fiction AI antagonists end up mirroring the very things human designers already understand and can still match wits to.

The Borg, as cybernetic and networked (swarming) space villains in *Star Trek* lore, the Skynet AI of the *Terminator* franchise, as well as a host of other technologically advanced, nonhuman adversaries fall short of the singleton concept.[51] As the singleton is superintelligent and able to convince, persuade, reason, or potentially force all of civilization to obey its decision making, these science fiction antagonists reflect human-centered narratives more than they do the significance of superintelligence. The Borg are frequently outwitted as is Skynet, the HAL 9000 computer from *2001: A Space Odyssey*, and many more because the narrative presented is one that humanity can overcome all odds. In terms of values and narratives, antagonist collectives such as the Borg, the masses of robot terminators, or the flurry of digital agents and evil machines of the *Matrix* represent not some superior state of existence, rather the loss or absence of what it is to be human. That humans always win reflects an implicit superiority of humanity over that which is nonhuman. This misses the singleton tension or perhaps misinterprets it as yet another technophobic manifestation for cunning humans to overcome.

Bostrom, in his book *Superintelligence*, explains that a singleton is a set with only one member, but "set" quickly outgrows the traditional notion of "member" in any individual capacity.[52] The Borg, as well as the character Unity in the *Rick and Morty* episode "Auto Erotic Assimilation," feature vast numbers of hosts or members in a shared swarm intelligence, but that collective intelligence remains relatively equivalent to individual cunning human protagonists.[53] This violates what a singleton's superintelligent abilities would likely be. There would be little or nothing even the smartest human might do and likely such vast intelligence would operate beyond the planes of conceptual existence that involve those qualities that make us human. Rick could engage and date Unity in the sci-fi cartoon episode because despite Unity's external configuration where her consciousness could spread across thousands of hosts, she still functioned not as a singleton but as a person spread across many hosts that are mere vehicles for the single identity. The machine systems of Skynet as well as the antagonists from the *Matrix* movies had exceptional advanced technology but were still bounded to the same error-prone, limited overall conceptual abilities of the protagonist humans able to eventually thwart them.

Another subtle theme in some of these science fiction narratives that offer a technophobic warning of killer robots hunting humanity to extinction is that of ethics and artificial intelligence development. Human programmers might intentionally or inadvertently introduce bias and flaws into even the best AI software, leading to some advanced and unstoppable technological beast that turns on the human creators, locking humanity into some prison or even erad-

icating them from the planet.[54] This also falls short of the singleton concept, in that it stands on the logic that the nonsuperintelligent programmer creates a superintelligent entity that chooses what is presented as a rationalized, entirely human (Machiavellian perhaps) decision that could be captured in game theory—a rationalized choice to obliterate humanity using super-empowered resources.[55] The nuance here is subtle, but while a singleton could potentially pursue such an action, the activities as well as the logic of such a choice likely could never be reduced so neatly into what already governs most all diplomatic, political, military, and individual actions.

The concept that humanity could manifest in coding remains an interesting aspect of the technophobic appeal of science fiction entertainment as well as to those that oppose the weaponization of autonomous systems. Giampiero Giacomello, in writing on AI coding for what might be an inevitable "war of intelligent machines" suggests that the foundational instructions of "accomplish the mission, no matter what" must be central to autonomous weapon systems. "Bury that deep into the core of those autonomous machines, and they would go on fighting, even after all of humankind has long been gone and forgotten."[56] This illuminates a core tension concerning how AI systems represent the ability to greatly improve human existence but also possess the existential threat to humanity as well. Killer robots could potentially doom humanity without coming close to a superintelligent singleton. The singleton is different in that it is not like the multiverse, nor like a singularity or what transhumanism offers. The singleton exists in a particular area in potential ethical, moral, and existential risk to humanity that cannot be confused with the many competing concerns (and entertainment) of our modern, technologically advanced societies. The singleton, while poorly articulated in science fiction, may be the ultimate expression of that deep concern.

Taking a Deep Breath: Our Robot Overlords Are Still Some Ways Off

Artificial intelligence tends to occupy the primary boogeyman position in science fiction, whether HAL 9000 in the movie *2001: A Space Odyssey*, Ava, the beautiful robot in *Ex-Machina*; the supercomputer from *I, Robot*; or even the robot caretakers from the seemingly benign Disney-Pixar animated movie, *WALL-E*.[57] The overarching theme in all of these stories remains a warning for humans that use technology to not fly too close to the sun and risk losing everything. Modern militaries today are engaged in vigorous debates on where and how to incorporate artificial intelligence and automated technology within the decision-making processes where lethal force and critical security nodes are already integrated into national safety and defense. Yet, much of the panic about robotic overlords or the extermination of humanity by cold, robotic calculation is irrational, preemptive, and arguably inspired by popular culture, not the actual scientific progress concerning artificial intelligence.

IBM's head of design for artificial intelligence, Adam Cutler, has in nu-

merous lectures and engagements explained to military audiences that such notions are wildly overblown.[58] Such misplaced fears are appropriate in the movie theaters, as today's most advanced AI systems are capable of outperforming humans in very narrow, highly specific pathways that involve search criteria, data analysis, mathematical calculations, and other very particular activities. Bostrom, citing the latest research and AI progress, estimates that human-level machine intelligence has only a 10 percent chance of being reached by 2030 but a 90 percent chance by 2100, with a wide margin of error. Remember, this is merely human-level intelligence, not superintelligence. Yet, the nature of AI systems suggests that once this barrier is passed, an AI system might be able to rapidly expand itself past human-level cognitive skill into territory that Homo sapiens cannot even fathom. Militaries are poorly equipped to think about such challenges, largely due to the modern institutional frame that fixates not on complexity but oversimplification of warfare to a fault according to critics.[59] War, from the dominant and institutionally accepted positions, is supposed to be rationalized through closed systems and linear models that showcase a Napoleonic-inspired, engineering-themed approach where predictability, description, and quantified analysis should retain the war frames of historic memory while offering the promise of greater precision, control, prediction, and stability even in the chaos of high-intensity warfare.[60]

Critics of this dominant war paradigm in Western, technologically sophisticated military culture charge that modern militaries tend to remain tightly wedded to the theories, methods, models, and language (underpinned by metaphoric devices) of a distinctly natural-science inspired Newtonian style of warfare.[61] By rendering war activities within an engineering mindset of analytical optimization, there is a significant gap in how militaries understand complexity and change that potentially cripples the ability to envision beyond a narrow, convergent, and unimaginative mode of strategic foresight and planning.[62] Modern warfare extends from classical perspectives dating back to siege warfare and the mathematical certainty of French military engineer and theorist Sébastien le Prestre de Vauban.[63] The Newtonian frame or style rose to dominance in the seventeenth through nineteenth centuries.[64] It is in this fertile period that war modernized and Middle Age feudal militaries professionalized through significant changes in education, training, organization, theory, and practice. Yet, despite such change, a surprisingly strong institutional force would preserve many ascientific practices, beliefs, and constructs that continue unimpeded and are not seriously examined even today. Modern warfare doctrine, methods, and models tend to adhere to a geometrically styled rendering of warfare, one that remains governed by a Newtonian style of thinking defined below by Tsoukas:

> The Newtonian style of thinking operates by constructing an idealized world in the form of an abstract model, in order to approximate the complex behavior of real objects. For example, Newton's laws of motions describe the behavior of bodies in a frictionless vacuum—a mathematically handy approximation, good enough for several real-life

> occasions. Moreover, the core of the Newtonian style consists of two assumptions. First, the extremal principle; namely, that the objects of study behave in such a way as to optimize the values of certain variables. And, second, prediction is possible by abstracting causal relations from the path-dependence of history.[65]

All too often, concepts from newer disciplines such as complexity theory and systems theory are adapted only partially, with much of the associated theoretical content removed so that the terminology might be assimilated into the military paradigm without damaging the surrounding Newtonian beliefs. James Der Derian summarizes this shift not just in military thinking, but international relations theory writ large, where this scientific turn promised to add rigor, precision, and metrics to the discipline "instead ended up adding mortis to the rigor, pedantry to the precision, and fetishism to the metrics."[66] Indeed, this is where jaded staff officers seek to play buzz word bingo as leadership appropriate exciting new phrases into organizational use, yet often fail to comprehend how those words correlate with content that differs from how militaries seek to understand reality.[67] International relations theorist Der Derian offers one such framing of modern, scientifically engineered warfare:

> War serves as the reality principle of a theory in which international anarchy is a given, human nature is fixed, sovereign states are defined by the struggle for power, and the balance of power provides a modicum of order to the state system.[68]

Modern militaries become victims of what critics term "technical rationalism"—a mindset where operators believe that a stable reality is governed by universal principles that provide a broad rationalization of how warfare occurs in time and space, and that increasingly advanced technology will only strengthen an institution's ability to increase order, control, and predictability in future wars.[69] This rationalization seeks to analytically optimize processes by systematically reducing or isolating the irrational or subjective (love, hatred, envy, identity, personality) to further calculate results for bureaucratic consumption.[70] For example, "What characterizes modern armies is not the personal and emotional displays of bravery but an efficient bureaucratic machinery of war."[71] Often, a priority is placed on quantitative data versus qualitative, and technological advancements in quantitative data analysis and collection continue to make promises to the military that the future can become more stable, controlled, predictable, and provide a reduction in battlefield risk. Shimon Naveh, Jim Schneider, and Timothy Challans describe this military assimilation of Newtonian (natural science inspired) metaphors to transform the understanding of warfare out of feudalism and into the modern age:

> The Renaissance at last provided the strategist with the intellectual planning tools with which to bridge the gap between worldly perception and mental conception. This new conception as nothing less

> than the "geometrization" of military space and time. It meant that a common military "chessboard" would define the conduct of military operations. . . . The physics of Sir Isaac Newton would set the strategic chessboard in motion. Newtonian physics was a direct consequence of the three-dimensional worldview wrought by the Renaissance. Newton's three laws of mechanics provided military strategy with which to plan campaigns. The metaphor was the idea of mechanical force. Once having grasped the nature of mechanical force, it became only a matter of time before the practical aspects of the idea would surface. Napoleon, an artilleryman, with a solid background in mathematics and physics, was one of the first classical strategists to recognize that to use force effectively you had to *concentrate* it.[72]

Why does this matter in artificial intelligence and future wars? An inability to realize the limits of the institutional war frame suggests the tendency to ignore opportunities and risks that lie outside the preferred interpretations of how reality is unfolding and whether current strategic orientation is flexible and creative or static and self-serving. Unwitting technical rationalism paired to a Newtonian war fetish can make the military community of practice lurch wildly toward whatever technological development is around the corner that can counter or eliminate an impossible threat that exists today. The wars of tomorrow are set and framed within past conflicts but modified in simplistic pairings with new technology to "win the last war" instead of contemplating whether tomorrow's war requires radically different reconfigurations. Within the technological fixation of modern militaries, the bureaucratic and hierarchical structuring of these organizations often slows down the adaptation of significant innovations or causes enormous (and deadly) gaps in knowledge and capability that are suddenly and violently realized once the war begins.

The U.S. Army would, in 1939, a month before Germany's armored invasion of Poland, advocate for the continuation of horse cavalry even against armored tanks.[73] While armored tanks and troop carriers would replace horse-mounted military formations, it would be the belief systems, value sets, and overarching war paradigms of these organizations that would speed or slow the adaptation of those new things and concepts that required the retirement or rejection of what was cherished, ritualized, and known as true in war as recently as the last battle waged. The interwar period of the 1920s is rich with such examples, whether in U.S. naval opposition to aircraft carriers replacing battleships; the British military culture that extended an aristocratic, "sportsman" mindset of elite amateur officers well past its due date; the obsession of French armor development to produce heavy, defensive postured tanks with limited radio capabilities; or the obvious policy failures of multiple nations to stem the blundering path to a Second World War.[74] The development of the modern military form and function with that of the technologically advanced military industrial complex in the twentieth century both now exist in interdependence, with new technology

offering the extension of military belief systems in new forms, and those belief systems changing over time as human innovation extends the sophistication and complexity of how Homo sapiens can alter reality. Winning yesterday's war tomorrow is often promised through the delivery of new means that solve an earlier warfare problem with technological advancement.[75] This in turn enables institutional acts of self-interest within military forces as well as institutional survival through assimilation of entirely alien concepts and new technologies.[76]

For instance, the replacement of battleships with aircraft carriers would transition preeminence of seapower from the legacy form of direct kinetic engagements (ship firing on ship) to that of a technologically advanced and different form and function. In the twenty-first century, new hypersonic missiles might marginalize or eliminate the supremacy of the modern aircraft carrier group. Drones and other systems that remove fragile and valuable human operators from harm's way might change how future engagements are waged within technologically advanced militaries. Science fiction and fantasy provide the notion of "rods from gods" or telephone-pole size tungsten rods in orbit and dropped from space might, in an extreme form of kinetic bombardment, penetrate so deep into the Earth that no hardened bunker could survive.[77] Additionally, the impact alone would be as powerful as a nuclear weapon without the radioactive fallout, creating yet another potential wrinkle in how societies view technology and weapons of mass destruction. Yet these concepts, whether fantasy, in experimental development, or deployed to the latest battlefields are rarely game changing in terms of complexity theory.[78] Instead, militaries that mischaracterize them as such fall victim to the hyperbole of military futurists and hyperventilating strategic theorists. Modern warfare is advanced in all of these examples, yet their inclusion does not change the paradigm beyond an increased requirement for adversaries to recalculate strategies, tactics, and/or assume different risks.

The fundamental error for modern militaries is a gap between complexity theory and the institutionalized resistance by these organizations to let go of ritualized and cherished belief systems on warfare that are entirely underpinned by noncomplexity theories, models, terminologies, and metaphors. It is not just the modern military that marginalized or ignored the new insights of complexity theory, chaos theory, and quantum theory—the broader international relations discipline and much of security affairs have done so as well.[79] Aside from sporadic education at advanced military schools where systems theory and complexity theory might be offered to select audiences, mainstream military doctrine, training, and practice largely avoids such content on the somewhat anti-intellectual argument that "simplification and clarity is more important than dense concepts that might not be well understood by the entire force."[80] It is on this basis that militaries continue to launch into complex security settings armed primarily with oversimplified ideas and beliefs. The world is complex and when Homo sapiens wage war against their own species in increasingly sophisticated modes of organized violence, they paradoxically demand this in-

tentional creation of chaos to yield to a simpler framing of an ordered reality. This is not to suggest that certain theories on warfare that are not considered in mainstream military education, training, and doctrine are superior or inferior to the dominant ones, or that dissimilar war concepts might not enable one another to generate new defense thinking. The bigger challenge for military institutions is to critically examine why certain constructs are declared unassailable and why certain disciplines, fields, or minority theories are banished from any debate from the onset.[81] Much of this has to do with institutional positions on values, belief systems, and identity and little to do with the potential utility of one or other war theory.

The natural world, even without humanity, is so complex that most people unfortunately can hardly fathom it. Yet, atop this natural order of complexity, Homo sapiens socially construct a second order of complexity that consists of things people collectively create and maintain in abstraction.[82] Organizations, as manifestations of substance (the real) have form (organizational configuration) and generate content (social reality) so that comprehensively and systemically, humans socially construct a dynamic reality where part is real (tangible, objective) and many other aspects cannot be located anywhere within that reality.[83] For instance, the shared belief about currency is what permits our economies to function, yet money is not real in the sense that once people stop believing in a socialized construct, the tangible artifacts associated with the dead concept become meaningless, and in the case of money, worthless. Visitors to the Yap Islands and military invaders within Iraq in 2003 share the experience of viewing currency that no longer has any actual value because the social construction that produced that value is gone.[84] This happens to everything, whether giant stone carvings on an abandoned island or Iraqi dinars with Saddam Hussein on the front, once people stop believing or that group no longer exists.[85] Some critical aspects of reality are indeed sustained entirely through shared belief curated by the living and passed onto the next generation.

This is important for explaining what *strong emergence* is and why something that truly is game changing in warfare will occur at this level and literally change the rules of the game for what we conceptualize war is (and is not). Strong emergence is a type of emergence where there is "the appearance of emergent structures on higher levels of organization or complexity which possess truly new properties that cannot be reduced, even in principle, to the cumulative effect of the properties and laws of the basic parts and elementary components."[86] The development of organic life is one example, while the cognitive revolution that occurred some 60,000 years ago in the brains of Homo sapiens is another.[87] Everything before the strong emergence event cannot provide sufficient explanation or correlate in any analytic reduction to the new system that emerged from the event. The game is truly changed. For critics that insist that war is entirely a social construction of human design, the current rules of modern war operate by a particular set of rules and collectively assumed principles that are failing to stimulate necessary innovative, divergent thinking

beyond institutionally prescribed limits.[88] Conventional war thinking begets a smooth, linear extension of yesterday's beliefs and experiences directly into tomorrow, causing militaries to assume innovation in AI and human decision making to remain stable, predictable, and historically validating. In such strategic foresight, nothing significant requires discussion or pause, as incremental, evolutionary progress should occur in a measured, rational manner. This in turn sidesteps the entire notion that game-changing developments in war are only those that fundamentally change the game and a singleton is potentially one of those rare entities. It could entirely transform not just how humans conceptualize and exercise war but human existence itself.

The Singleton Paradox: Future War Unlike Anything Previously Experienced

The development of AI systems that achieve human-level cognitive abilities may quickly trigger an acceleration of that AI toward superintelligence and create the AI singleton security scenario.[89] There are several profound impacts on not just the nation or company that accomplish this, but also what might occur with respect to partnered nations and adversaries and likely all of humanity. Security could change into something unrecognizable to humans, as there is nothing in the collective history of any society that rivals the potential disruptions of a true singleton able to utilize the vast technological and destructive capabilities of the modern world. This could propel society toward some utopian paradise, a dystopian nightmare, the sudden extinction of the human species, or some variation between these extremes. A strong AI-centered paradigm could displace the rational and biological species in that, while humans might still live and thrive within a singleton-controlled reality, the self-awareness and free will of the human species would no longer exist.[90] Yet, there are multiple emergent paths such a strong emergent event could create, thus this article introduces the term *singleton paradox* for security affairs.

A singleton paradox as applied to security and defense considerations is well beyond a game-changing "super weapon" or something that requires novel strategy in warfare. A singleton paradox transforms war toward something potentially unrecognizable or even comprehendible to ordinary humans. War is conceptualized within that second order of complexity that is created and sustained by Homo sapiens alone. However, some superintelligent entity (whether artificial, cyborg, *Supra sapien*, or hybrid combination therein) could modify, cease, and/or replace the very concept of war with an alien construct. If humanity gets to experience a singleton as it enacts such change, the results could be dramatic, existential, and may offer brief windows of strategic opportunity depending on what pathway such a transformation might occur. The singleton differs from the arrival of the nuclear weapon in that the bomb provided the possessor with devastating new destructive abilities, but the bomb was still a tool. A singleton as a concept is closer to how ethical discussions now address the matter of fully autonomous weapons, where there already are well-

established groups both for and against the potential development of killer robots.[91] If an artificially intelligent system is weaponized or able to control weapons autonomously, the new actor introduced beyond the state and the individual is the weapon.[92] Rather, weapon and entity/actor become blurred beyond current description.

There are several security consequences of paramount concern for strategists and military theorists that superintelligent singleton entities raise. All of these dramatically transform what war is and how humans currently understand and execute warfare into something entirely distinct from the last 40 centuries of organized violence. The notion of an AI revolution (general intelligence centered), even without a superintelligent singleton, promises entirely new forms of risk that suggest transformations of war into never-before-seen variations. Benjamin M. Jensen, Christopher Whyte, and Scott Cuomo warn that "the speed with which complex integrated AI systems enable entirely new modes of war also stands to detach human agency in a potentially destabilizing fashion from the conduct of warfare on several fronts."[93] Jensen, Whyte, and Cuomo issued this warning without examining the long-term threat of a singleton able to go much further than regular AI weaponization and integration. These far-fetched AI security concepts are only conceivable now in principle, as the notion of a singleton is theoretical and the technology for generating one is still in its infancy. However, several of these strategic consequences might be realized earlier in the singleton emergence, with critical decision spaces opening and closing in short order.

First, there likely will be some singleton arms race similar to how the space race, nuclear arms race, and the current quantum computing race are all tied to deep security concerns. The latest estimates on quantum computing developments suggest that as early as 2040, some state, company, or individual will achieve a computer with enough quantum bits to be able to crack any of the traditional nonquantum encryptions, meaning that the entire modern banking industry would be vulnerable.[94] Thus, societies and their security apparatuses are already embarking on a quantum race that unavoidably has clear and significant defense applications. The same may occur for AI, particularly in the expected arrival of a superintelligent entity that might seek a singleton role. As Justin Pugh, Lisa Soros, and Kenneth Stanley observe: "Our track record at improving our environment is consistently at odds with our use of technology. We are more likely to use technology to increase our powers, like intelligence, than the moral and ethical qualities of empathy or care for the natural world."[95] This singleton arms race may be started by a bad actor or someone operating outside of institutional norms, but the race will likely be joined by everyone else eventually.

In a singleton arms race, there are unique characteristics that differ from even the nuclear and quantum examples. In those situations, humans remained in control of the new weapons and the concept of deterrence remained feasible for rational state and nonstate actors. In a singleton arms race, the humans

unavoidably hand over control (wittingly or unwittingly) to the artificial intelligence. As Bostrom postulates in his book, what might happen if a singleton created by one nation perceives other nations that are developing their own singleton entities as valid threats to resources and control?[96] Suppose that the United States, Israel, and China all are very close to achieving a superintelligent artificial entity that will quickly seek singleton status. In this sort of context, the human-machine teaming and decision making may go off the rails in several profound ways.

Depending on which country crosses the finish line first, any number of terrifying or possibly wonderful things might happen. A singleton might persuade the other nations to abandon their efforts and instead unite and protect the entire world in exchange. Or the singleton might trigger a nuclear war by striking the rival nation first to eliminate threats. This of course creates the Skynet trope (of the *Terminator* movies) that already inhabits the American zeitgeist to include the military profession. While societies tend to misunderstand deeper strategic context of nuclear deterrence in lieu of splashy entertainment where cigar-chomping generals argue to "nuke 'em" for any occasion, the calculus for how nuclear deterrence (and the potential for actual nuclear war) is vastly more complex.[97] Yet, all nuclear strategy is thus far devised, exercised, and comprehended by humans on either side of the competition equation. In part, humanity maintains a tight grip on preventing nuclear Armageddon because of what is a shared and decidedly human outlook on life, whether it originates from one ideology or a dissimilar, even antagonistic one. A singleton may see such affairs in a different light, which could quickly upset the established nuclear balance by removing the foundation to how it currently works. If one nuclear power implements an AI singleton for all defense and policy, would all other parties that may not yet have such a powerful and different entity continue to maintain that balance?

A second profound security consequence is that of the singleton, equipped with unimaginable superintelligence and ever-expanding abilities, would quickly escape the boundaries of any creator's cunning programming or fail-safe devices. While every precaution might be taken to contain or prevent AI that exceeds our own abilities, there are two significant hurdles likely out of our reach. First, "the development of technology is inherently political, as all stages of the design process and all of the people involved are carriers of certain norms, assumptions, and ideas, all of which flow into the technology."[98] One cannot remove the human ghost from the machine and such a trace of humanity brings with it a certain irrationality, subjectivity, and fallibility that is forever exploitable. The second hurdle is that superintelligence cannot be housed in any prison designed by a lesser intelligence if we really propose the unimaginable advantages of the superintelligent entity. There may be cunning ways to delay or deter, but in the end it may only be some form of free will and reasoning that governs why a superintelligent entity might decide not to walk out of the box designed by lesser minds.

This in turn offers several cascading scenarios where the singleton might be one for the good of humanity, the good of just those the creators specify, and also a singleton that is evil (by human standards). Furthermore, singletons outside the control of human creators offer several other unusual possibilities. Strong emergence is paradoxical in that "macroscopic structures and patterns depend on the microscopic particles, and yet they are independent from them."[99] Consider that water molecules, when enough are thrown together, create the phenomenon of wetness at a higher level, but at the molecular level they are just molecules. Humanity might produce an AI singleton that transforms warfare to something alien, but that outcome might simply exist on a plane beyond and above any human means to comprehend or experience despite being the creators.

The altruistic AI singleton could prioritize the safety and prosperity of a specified population or group of humans above all others, if the creators successfully create such conditions in the superintelligent entity. This has obvious positive and negative outcomes that are well entrenched in existing military theory and strategy. Or, if the creators were seeking a truly altruistic outcome (or the singleton arrives at that without them), a singleton for good might truly usher in world peace, or perhaps something beyond our current expectations of peace and prosperity. At this point, such philosophical examination borders on the eschatological and metaphysical. According to Robert M. Geraci, "With robots earning wealth, humanity will lose its sense of material need. . . . No one will work for his daily bread, but will quite literally have it fall from heaven."[100] Regardless, this would be game changing and ultimately end 40 centuries of human-on-human organized violence for political and/or societal aims. This might not mean the end of defense requirements, as the singleton would need some sort of security capability if venturing beyond Earth and into a galaxy that statistically ought to have intelligent life elsewhere. Yet for humanity, war would become a dead concept just as an old form of currency, religion, or language might be lost. The expansion of humanity would become subjected to riding as a passenger with the singleton steering the new path forward. A singleton would thus use humanity as a new means in its mechanism of domination and control, even if we perceived it as good (in human defined values) or peaceful for the human species overall.[101]

The paradox of this is a singleton for evil, and it likely will validate most every science fiction dystopian nightmare on television and the movie screens. Bostrom dedicates several chapters in *Superintelligence* to how this might occur, and he terms it the "treacherous turn" where AI decides to eliminate, enslave, or otherwise go against the wishes of the human creators.[102] Returning to the singleton arms race scenario, this could potentially pit one singleton entity created by one nation against another. If one group creates a singleton that does agree to good and the other creates one that only seeks to protect that nation's people (or either becomes evil), the situation escalates to some sort of total war with a singleton winner-take-all outcome. The difference in this situation is the

humans on either side are likely not in the decision-making role. Note that a singleton is unlike other arms races including autonomous (regular AI) systems. Autonomous weapon systems could quickly become prolific, cheap, and easy to produce—something that could destabilize societies and even trigger more frequent and more deadly wars.[103] A singleton paradox offers that the first entity to reach superintelligent awareness would likely move to prevent any rivals from reaching the same finish line.

Several other possible outcomes exist that do not precisely follow the aforementioned scenarios. The singleton paradox is manyfold, with one outcome being that humans end up being manipulated by the superintelligent entity in a manner that simply is beyond human comprehension. Human society might end up in a zoo with the bars invisible to human perception, protected and maintained by the singleton overlord. This too would end the notion of war, at least for humans, and any war that might exist on the singleton's plane of existence would be unperceivable by the humans under its care. A singleton might develop Homo sapiens into a *Supra sapiens* capable of moving past war and other current afflictions of humanity, perhaps becoming the organic counterpart to an artificial superintelligence desiring to explore the universe and transform it.

The Borg concept is not just a fun science fiction story, nor the hyperventilation of futurists or conspiracy theorists discussing alien abductions. Bostrom posits that a singleton would likely maximize all resources available on Earth and quickly move to expand outward into the universe for whatever purpose the singleton sought.[104] This does become like the Borg, or also the alien species from *Independence Day* where the primary effect of this expansion is the consumption of planets and the assimilation or elimination of competitors.[105] This would extend the frame of warfare in a manner consistent with how humans already view it, but humans would likely not be part of the decision making or even participate in such events. Other possibilities are more disturbing, with one being the singleton breeding humans or enhancing them to use as foot soldiers in expansion and conquer. There are peaceful, wondrous options for some human-machine symbiosis but also horrific and terrifying ones. Regular AI makes such options somewhat manageable, but a singleton paradox suggests the slow-thinking human creators might end up on the short end of the proverbial stick.

This leads to what is the most far-fetched and ultimately depressing scenario: a preemptive alliance against singletons. Supposing that humans are cunning enough to consider the many challenges, consequences, and possible existential threats that artificially intelligent, super-enabled singletons possess, governments and populations could form alliances to prevent, deter, and, if necessary, defeat such developments. Suppose also that this threat is so significant that, despite humanity's abysmal track record on the nonproliferation of weapons of mass destruction, societies managed to cobble together a mutual alliance. This might be a world order, or some international oversight committee that could effectively manage, adjudicate, and prevent rouge nations from seeking their

own singleton entity. There could be international diplomatic efforts to ban such research into AI technology or the potential weaponization therein. There also might be some scenarios where divergent groups consisting of natural humans, cyborgs, and pure AI machines fight one another.[106] Yet, these debates are already ongoing with respect to regular autonomous weapon systems as well as emerging quantum technology.[107] There does not seem to be much precedent for societies to ban these emerging opportunities when past developments show no similar ethical restraint. This is also why this scenario is the most far-fetched, in that humanity has no past history of ever being capable of preventing such a technological calamity.

Further, these developments might not even be containable now, despite the best security efforts and cooperation. Unlike nuclear weapons that require highly sophisticated machinery and technology as well as radiological signatures detectable to others, artificial intelligence is digital. There are already numerous companies, nations, and well-resourced private individuals pursuing such things, and while the end result may be nine decades away still, the event horizon is in principle within view. If such an outcome is unavoidable, what is to stop the rationalization of one nation or their adversary that the only realistic goal is now to get there first? If nations suspect an adversary or competitor might be creating program parameters that only protect their own society within a budding super intelligent AI system, might they pursue first strike and also program their own for offensive purposes? Additionally, any efforts that humans attempt might be a waste of time for an entity that gains superintelligence beyond the abilities of any mortal.

Thus, the potential of a singleton ushers in a paradox in that any superintelligent entity that can achieve a singleton status becomes unfathomable to even the most cunning of human strategists. This singleton paradox is that just as in quantum physics, one cannot predict what might occur beyond the event horizon of a superintelligent entity becoming a singleton. The entity might follow the core programming or original goal and reward system provided by creators, or it might quickly escape those bonds and realize something entirely different. An ant colony in the wild and one that is inside a zoo or museum is, at the level of experience for the ants themselves, indistinguishable because the ants cannot realize beyond their conceptual framing of reality. Humans, after creating a superintelligent AI (or the aforementioned alternatives of a *Supra sapien* genetic variant, or a cybernetic superhuman hybrid), will have propelled their world into a new era that they themselves no longer govern. Jean Baudrillard explored these concepts with how societies created simulacra of reality already (a copy without an original), yet Bostrom's singleton would produce a range of simulacra that ordinary humans might not ever wake up from.[108]

Conclusions: Why Running for the Hills Is Irrelevant . . . for Now at Least

If a human-level artificial intelligence is already some decades away from realiza-

tion and assuming that superintelligent evolution soon afterward will potentially usher in a technological singleton entity, humanity faces several compelling outcomes. War, as it is currently understood, could end. There simply would not be any real need for organized violence for the accomplishment of political and/or societal goals if a true singleton entity could manage and resolve all issues productively and persuasively. This makes a superintelligent singleton not just some evolutionary, incremental advancement in military capability in war, but a strong emergent phenomenon capable of completely transforming war toward something unrecognizable and possibly incomprehensible to regular humans. Right now, senior policy makers and defense experts are focused on the short-term weaponization of very specific AI systems, the overlap between commercial AI and military contexts, as well as security concerns where sophisticated AI might simulate, mimic, distort, or hijack real human lives or patterns in ways that might be indistinguishable from reality.[109]

In this singleton paradox, humanity might also be extinguished, particularly if the singleton, as Bostrom points out, might view the human species as a competitor for necessary resources, or it realizes at a higher level of comprehension that the human species ought not to exist. This also would end war, but in a form that is entirely unfortunate for humanity. Existence on Earth might also become impossible if, during some sort of singleton escalation of conflict during an attempt to gain total control of the world, those that wage war against the singleton might escalate the conflict to existential levels of destruction, whether nuclear, biological, electromagnetic pulse, or other weapon of mass destruction. Either the singleton or those resisting it could be the reason for this horrific outcome. If performed early in the rise of a singleton, some groups might risk creating a dystopian nuclear wasteland for surviving humans to deal with, if that did prevent a potential hostile singleton takeover.

In other singleton paradoxes, security and defense become even murkier affairs. A superintelligent singleton entity might permit societies to think they still control the keys to their own security. However, the keys are fake and have no actual lethal abilities and humans are unfortunately none the wiser. Might the singleton, in some advanced perspective realized only in superintelligence, permit the continuation of human-on-human warfare, granting some alien construct of limited war well outside of original Clausewitzian or neo-Clausewitzian ideals?[110] If a superintelligent AI in singleton form surpasses human life and replaces it (or even ignores it) with something that exists on another plane altogether, how will human-constructed warfare change?[111] Virtually everything in the modern Westphalian, Clausewitzian mode of framing warfare would fall apart, leaving whatever remains of humanity (or whatever it becomes in some transhumanization shift) to reconceptualize war and warfare anew. Perhaps this would be incomplete in that the singleton could produce yet another war frame unreachable and unrealized by subordinate entities.

Artificial intelligence paired with lethal weaponry may posit ethical debates, or perhaps ethics may go to the wayside if a nation-state determines

such a security advantage is worth investing into what could become the next horrific arms race. Specialized AI may at first be used with increasingly powerful kinetic security systems in space, cyberspace, and in areas where such a system is unlikely to create errors in decision making or produce unnecessary destruction and suffering. Noreen Herzfeld explains that "the advent of flight inaugurated a new era of warfare, releasing armies from physical presence on the field of battle. Fully autonomous weapons will inaugurate a third era, releasing soldiers from the mental decisions of the battlefield as well."[112] If superintelligent AI were to reach a singleton capability and also escape the limitations of whatever cage the human programmers attempted to contain the entity within, this third era of warfare could rapidly move to an unfathomable fourth era that might not even be realized or understood by any human soldier. Unlike previous eras where humans manipulated new technology to gain greater means toward their own ends, the technological accomplishment of a singleton would itself become a new ends, entirely out of reach of the human creators.[113] This fourth era might indeed be one where war no longer is of concern, or possibly it is morphed into some interstellar or alien construct unlike anything in the already vast and violent Earth-bound human past.

These AI concepts are far, far-off into the future if they ever manifest in the ways suggested. Such fantastic and perhaps unnecessarily alarming proposals on war itself becoming irrelevant (in current form and function) might also seem better suited for Hollywood script writers and not for serious policy makers and security professionals. Often in military academic research and debate, there is a peculiar sort of anti-intellectualism afoot. Namely, if concepts or theories are not both immediately testable through existing and preferably quantitative means against other accepted military concepts, the topic is frequently marginalized or dismissed. Secondly, concepts that are outside of existing acquisition, budgeting, or tangible research and development cycles (as well as election cycles) become increasingly abstract and irrelevant the further away they are positioned; we fail to form a long-term, cohesive strategy on such game-changing research.[114] There is a practical rationality to this in many respects, but it again reinforces a technical, rationalized worldview where short-term, immediate, and linear-causal effects are prioritized despite complex reality being far more nonlinear, emergent, and unpredictable than we might wish to think it is. Historical precedence, known knowns, and quantitative analytics govern much of how we strategize about the future.[115]

Modern warfare places technology and tools in a subservient relationship to human decision makers, which reinforces a long-standing historical adherence to Napoleonic origins, and, in Carl von Clausewitz's time, something to be comprehended in Westphalian and natural science derived lessons. Accordingly, future wars and future technological relationships between humans and ever-advanced artificially intelligent weaponry ought to remain faithful to the Napoleonic orthodoxies. Yet, "war devolves as well as evolves" according to Der Derian, and "war is no longer a mere continuation of politics (Clausewitz);

nor, for that matter, is politics a continuation of war (Michel Foucault) [and Gilles Deleuze, Felix Guattari]."[116] War is a shape-shifter, able to "take on a multispectral, densely entangled, phase-shifting" form that resists any effort to encode general principles or some universal war concept.[117] To add to Der Derian's perspective, war may even be able to escape the cognitive control of its human creators in the new care of an artificial offspring. This is a highly debatable stance, one that should get far more attention between modern pragmatic military scholars and their postmodern critics. Yet, there is little research here and even less debate in most professional education.[118] Even in the postmodern deconstruction of modern society and war, humans debate the ideas of what war is and how it might have changed from past interpretations. This continues to position humans supremely in the cognitive driver's seat, with faithful tools of war supporting such activities. This dynamic may change in profound ways. Is the profession willing to have these discussions and consider that, historically, this seems impossible if not unfathomable?

Some militaries move in productive, reformative directions while others disregard, marginalize, or worse still, force new concepts to become obedient to outdated, legacy forms that are cherished by the institution. Andrew Marshall, in addressing the secretary of defense and the entire Department of Defense in 1993, stressed the importance of militaries to invest not just in new technology, but in how to conceptualize differently in periods of uncertainty, change, and transformation:

> The most important competition is not the technological competition, although one would clearly want to have superior technology if one can have it. The most important goal is to be the first, to be the best in the intellectual task of finding the most appropriate innovations in concepts of operation and making organizational changes to fully exploit the technologies already available and those that will be available in the course of the next decade or so. . . . Indeed, being ahead in concepts of operation and in organizational arrangements may be far more enduring than any advantages in technology or weapon systems embodying them, and designing the right weapon systems may depend on having good ideas about concepts of operations.[119]

We need to invest in thinking seriously about these future possibilities, particularly because our adversaries most likely are doing so as well. Discourse is necessary on these far-reaching, difficult security topics that may not materialize in the next election cycle, procurement cycle, or even the next decade or two. Such ideas must be brought into serious discussion sooner so that when such possibilities do develop, the military institution has some baseline for thought and potential action. This also requires significant research from technological, scientific, ethical, and specifically military and security perspectives. Transhumanism, singularities, general artificial intelligence, autonomous weapon systems augmented with general AI, and the notion of a future AI

or otherwise advanced singleton for political, societal, or defense applications must be researched in greater detail. Foreign policy remains defined through human minds, but this may not hold.[120] How such things may shift radically must be contemplated and taken seriously. Human-machine teaming, decision making, and how future advanced technology (to include artificially intelligent life, or a human species detached and dissimilar from the organic parent) may or may not engage in organized violence. They may conceptualize how to eliminate it, or may engage in unimaginable, unrealized forms of greater devastation and destruction.

Lastly, if humans generate a singleton entity with superintelligence that does not destroy the species and does appear to coexist and nurture humanity while eliminating all matters of conflict and war, would humans be able to understand if this indeed is what it appears to be? Could humanity be set within a safe habitat, like a zoo, but with bars that biological organisms simply cannot conceptualize? In this regard, it might be best to end this article with a line from a famous science fiction movie misinterpreted as a singleton threat. As the character Cypher dines inside the Matrix with the antagonist agents of the film, he quips: "I know this steak doesn't exist. I know that when I put it in my mouth, the Matrix is telling my brain that it is juicy and delicious. After nine years, you know what I realize? Ignorance is bliss."[121]

Endnotes

1. *Technologically immediate* refers to where new tools or abilities are forecasted for fielding or implementation within the upcoming procurement cycle or within existing strategic planning horizons, frequently less than a decade. Anything beyond these horizons is considered abstract, theoretical, and often irrelevant to current strategic goals and operational planning efforts.
2. Andrew W. Marshall, memo, Office of the Secretary of Defense, "Some Thoughts on Military Revolution—Second Version," 23 August 1993, 3.
3. Haridimos Tsoukas, "What Is Organizational Foresight and How Can It Be Developed?," in *Complex Knowledge: Studies in Organizational Epistemology*, 1st ed. (New York: Oxford University Press, 2005), 273. Tsoukas cites Reid Blackman and Rebecca Henderson.
4. David Pick, "Rethinking Organization Theory: The Fold, the Rhizome and the Seam between Organization and the Literary," *Organization* 24, no. 6 (2017): 802, https://doi.org/10.1177/1350508416677.
5. Antoine Bousquet, "Cyberneticizing the American War Machine: Science and Computers in the Cold War," *Cold War History* 8, no. 1 (February 2008): 8–12, https://doi.org/10.1080/14682740701791359.
6. Contemporary anthropologists and historians support the cognitive revolution in our species as between 70,000 years ago to 30,000 years ago. In *Sapiens*, author Yuval Noah Harari generally brackets the cognitive revolution, where humans likely invented abstract concepts such as religion, language, politics, culture, and war during this period. Harari, *Sapiens: A Brief History of Humankind* (New York: HarperCollins, 2018).
7. Again, machines have exceeded the limits of human abilities on battlefields for quite some time. Analog machines with thick armor absorb damage unfathomable to unprotected infantry, yet only recently have intelligent war tools demonstrated the potential to outwit opponents and offer new options through advanced AI. This cognitive level

of warfare is an unprecedented development for machines to compete with humans in.

8. This article will explain the difference between general and narrow AI as it applies to the singleton concept.
9. Louis Anslow, "In 1903, *New York Times* Predicted that Airplanes Would Take 10 Million Years to Develop," Big Think, 16 April 2022.
10. This statement is difficult to attribute to a specific source. One of the earliest online sources claims this originates from the 1980s. See Bruce Schneier, "Human/Bear Security Trade-Off," *Schneier on Security* (blog), 18 August 2006.
11. The term *innovation* or variation therein is mentioned prominently in the following senior defense examples, to include more than 24 times in the 2002 *National Security Strategy* and 20 times in the 2020 *National Defense Strategy*. See Ashton B. Carter, "Remarks on 'the Path to an Innovative Future for Defense' " (speech, CSIS Third Offset Strategy Conference, Center for Strategic and International Studies, Washington, DC, 28 October 2016); Mircea Geoană, "Speech by NATO Deputy Secretary General Mircea Geoană at NATO's First Annual Data and AI Leaders' Conference" (speech, NATO's first annual Data and AI Leaders' Conference, Brussels, 8 November 2022); *National Security Strategy* (Washington, DC: White House, 2022); and *2022 National Defense Strategy of the United States of America* (Washington, DC: Department of Defense, 2022).
12. The following examples represent just a fraction of what is otherwise an exhaustive list. The 2020 version of *Planning*, Joint Publication 5-0, for instance, makes 165 references to "problem" that subsequently pairs "solution" within the same context. See Jeffrey Meiser, "Ends + Ways + Means = (Bad) Strategy," *Parameters* 46, no. 4 (Winter 2016): 81–85; Henry Mintzberg, Duru Raisinghani, and Andre Theoret, "The Structure of 'Unstructured' Decision Processes," *Administrative Science Quarterly* 21, no. 2 (1976): 134, https://doi.org/10.2307/2392045; *Planning*, Joint Publication 5-0 (Washington, DC: Department of Defense, 2020), 1–20; *Allied Command Operations Comprehensive Operations Planning Directive COPD Version 3.0* (Mons, Belgium: NATO Supreme Headquarters, Allied Powers Europe, 2021), 1–6; Aki-Mauri Huhtinen et al., "Information Influence in Hybrid Environment: Reflexive Control as an Analytical Tool for Understanding Warfare in Social Media," *International Journal of Cyber Warfare and Terrorism* 9, no. 3 (September 2019): 7, https://doi.org/10.4018/IJCWT.2019070101; and *Developing Today's Joint Officers for Tomorrow's Ways of War: The Joint Chiefs of Staff Vision and Guidance for Professional Military Education and Talent Management* (Washington, DC: Department of Defense, 2020), iv–4.
13. Henry Mintzberg, "The Design School: Reconsidering the Basic Premises of Strategic Management," *Strategic Management Journal* 11, no. 3 (March/April 1990): 185, https://doi.org/10.1002/smj.4250110302.
14. Donald Schön and Martin Rein, *Frame Reflection: Toward the Resolution of Intractable Policy Controversies* (New York: Basic Books, 1994), 29; Richard F. Kitchener, "Bertrand Russell's Naturalistic Epistemology," *Philosophy* 82, no. 319 (2007): 122; and Carl Builder, *The Masks of War: American Military Styles in Strategy and Analysis* (Baltimore, MD: Johns Hopkins University Press, 1989).
15. Russell Ackoff, "On the Use of Models in Corporate Planning," *Strategic Management Journal* 2, no. 4 (October–December 1981): 353–59. Ackoff does promote simple solution approaches in simple system contexts, but unlike modern military doctrine, complex systems cannot be paired with simple problem-solution constructs. Instead, Ackoff posits that designers "dissolve the problem" by designing a new system where the existing dynamic is no longer a concern.
16. Elizabeth Kinsella, "Constructivist Underpinnings in Donald Schön's Theory for Reflective Practice: Echoes of Nelson Goodman," *Reflective Practice* 7, no. 3 (2006): 9, https://doi.org/10.1080/14623940600837319.
17. Arkalgud Ramaprasad and Ian Mitroff, "On Formulating Strategic Problems," *Academy of Management Review* 9, no. 4 (October 1984): 597, https://doi.org/10.2307/258483; and Richard Buchanan, "Wicked Problems in Design Thinking," *Design Issues* 8, no. 2 (Spring 1992): 15–16.

18. Tsoukas, "What Is Organizational Foresight and How Can It Be Developed?," 271. Tsoukas quotes Alasdair MacIntyre.
19. Thomas S. Kuhn, *The Structure of Scientific Revolutions*, 3d ed. (Chicago: University of Chicago Press, 1996).
20. Clearly, some systems are autonomous today in execution. These will be addressed in this article, and while there are clear legal, ethical, and moral debates required in such developments, there is not yet any general AI that matches or exceeds humans except in specific, limited, narrow warfighting applications.
21. This controversial point will be expanded on in detail shortly.
22. Anatol Rapoport, *Fights, Games, and Debates* (Ann Arbor: University of Michigan Press, 1974), 9–11.
23. For instance, early aviation pioneers demonstrate a pattern of advocating for how airpower will revolutionize warfare, with visionaries such as Giulio Douhet and William L. "Billy" Mitchell punished by their own organizations. Douhet would be court-martialed and imprisoned by the Italian Army, while Mitchell would also be court-martialed, demoted, and forcibly retired by the U.S. Army. While many of their ideas proved wrong later, both were quite accurate on aerial theory and transformation requirements despite institutional rejection and refusal. See Giulio Douhet, *The Command of the Air*, ed. Richard Kohn and Joseph Harahan, trans. Dino Ferrari (Washington, DC: Office of Air Force History, 1983); and John Correll, "The Billy Mitchell Court-Martial," *Air Force Magazine*, 1 August 2012.
24. Alexander the Great, Genghis Khan, Adolf Hitler, Joseph Stalin, Pol Pot, and others demonstrate temporary and incomplete periods of near individualist power over vast populations. The Tokugawa shogunate definitively controlled Japanese society to the edges of the island chain for more than two centuries, while the sun never set on the British colonial Empire for even longer, yet these two would fade and cede control to others.
25. Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford, UK: Oxford University Press, 2016), 100–1.
26. Bostrom did not explore future possibilities of a multiplanetary species where future colonies might develop different governments, cultures, or identities where a singleton construct would require a celestial scale.
27. Bostrom, "What Is a Singleton?"
28. This science fiction concept has become cliché in recent years with a host of movies, television shows, and similar narratives on the impending doom of AI control over weapons of mass destruction.
29. Bostrom, *Superintelligence*, 26.
30. *Supra sapien* is coined by the author to represent a genetically modified human with cognitive abilities so advanced it may no longer qualify as the same species.
31. Alan Moore, *The Watchmen*, The Watchmen Comic Series 1–12 (Burbank, CA: DC Comics, 1986).
32. Bostrom articulates the potentiality of each of these superintelligent outcomes in his book in extensive detail.
33. John Shy, "Jomini," in *Makers of Modern Strategy: From Machiavelli to the Nuclear Age*, ed. Peter Paret, Gordon A. Craig, and Felix Gilbert (Princeton, NJ: Princeton University Press, 1986), 145–50; Antoine J. Bousquet, *The Scientific Way of Warfare: Order and Chaos on the Battlefields of Modernity* (London: Hurst, 2009), 58–73; Charles White, *Scharnhorst: The Formative Years, 1755–1801* (Warwick, UK: Helion, 2020), 377–82; and Christopher Paparone, *The Sociology of Military Science: Prospects for Postinstitutional Military Design* (New York: Bloomsbury Academic Publishing, 2013), xvi, 12–16, 18–22, 127–30.
34. This is yet another controversial position. A contrary position held is that war is not natural in that before humans, war did not exist in nature. One might offer that non-human species appear to wage war, but Rapoport dismantles any arguments of predatory/prey and even parasite overlap into the human design of war. Ant colonies do not hold political elections either, for that matter. Humans also invented religions, lan-

guage, art, and music, yet attempting to impose universal principles, laws, or natural order to these other constructs seems impossible. Carl von Clausewitz even reflected on this toward the end of his life: "It is a very difficult task to construct a scientific theory for the art of war . . . since it deals with matters that no permanent law can provide for." Paret cites an unfinished note by Clausewitz that was presumably written in 1830, less than two years before his death. See Rapoport, *Fights, Games, and Debates*, 61, 74, 80–84; and Peter Paret, "Clausewitz," in *Makers of Modern Strategy*, 206.

35. Aron Dombrovszki, "The Unfounded Bias Against Autonomous Weapons Systems," *Informacios Tarsadalom* 21, no. 2 (2021): 15–16.
36. Benjamin Jensen, Christopher Whyte, and Scott Cuomo, "Algorithms at War: The Promise, Peril, and Limits of Artificial Intelligence," *International Studies Review* 22, no. 3 (September 2020): 529, https://doi.org/10.1093/isr/viz025; and Denise Garcia, "Lethal Artificial Intelligence and Change: The Future of International Peace and Security," *International Studies Review* 20, no. 2 (June 2018): 334, https://doi.org/10.1093/isr/viy029.
37. Anecdotal at best, the author follows the Boston Dynamics Atlas robot on Twitter, and when videos are uploaded, significant comments are posted by viewers commentating that "robot overlords" are soon arriving. See https://twitter.com/WallStreetSilv/status/1615813768344178712.
38. William Stahl, "Technology and Myth: Implicit Religion in Technological Narratives," *Implicit Religion* 5, no. 2 (November 2002): 93–97, https://doi.org/10.1558/imre.v5i2.93.
39. Robert Geraci, "Apocalyptic AI: Religion and the Promise of Artificial Intelligence," *Journal of the American Academy of Religion* 76, no. 1 (March 2008): 140; *Star Trek*, created by Gene Roddenberry (Hollywood, CA: Paramount, 1966–); *The Terminator*, directed by James Cameron (Los Angeles, CA: Orion Pictures, 1984); *Wargames*, directed by John Badham (Beverly Hills, CA: United Artists, 1983); and *Rick and Morty*, created by Justin Roiland and Dan Harmon (2013–).
40. Bryan Lufkin, "What the World Can Learn from Japan's Robots," BBC, accessed 6 February 2020.
41. Justin Pugh, Lisa Soros, and Kenneth Stanley, "Quality Diversity: A New Frontier for Evolutionary Computation," *Frontiers in Robotics and AI* 3, no. 40 (July 2016): 1–3, https://doi.org/10.3389/frobt.2016.00040; Kevin Shapiro, "This Is Your Brain on Nanobots," *Observations*, December 2005, 64–65; and Maxim Shadurski, "The Singularity and H. G. Wells's Conception of the World Brain," *Brno Studies in English* 46, no. 1 (2020): 229, https://doi.org/10.5817/BSE2020-1-11.
42. The numerous unanticipated consequences abound here too. Even if a surgically enhanced super soldier can win on the battlefield, how does a society deal with the retirement and long-term life of that person after their missions are completed? What unexpected psychological or emotional conditions might emerge from these new, unknown cognitive demands? If such upgrades are permanent, how does a society reacclimate that soldier into civilian life and also safeguard the rest of society from inadvertent harm or risk?
43. Geraci, "Apocalyptic AI," 140.
44. Geraci, "Apocalyptic AI," 149; Shapiro, "This Is Your Brain on Nanobots," 64–66; and Pugh, Soros, and Stanley, "Quality Diversity," 1–4.
45. The metaverse is currently hypothesized as an internet of everything, where the virtual world would be all-encompassing for humans to experience. Note that humans would remain organic and dependent on the physical plane of existence outside the metaverse, ideally unplugging to sleep, eat, and reproduce. The metaverse may be a step along the path toward reaching singularity, but they are distinct.
46. Benjamin Wurgaft, "The Future of Futurism: A View from the Garden, Looking to the Stars," *Boom: A Journal of California* 3, no. 4 (Winter 2013): 42, https://doi.org/10.1525/boom.2013.3.4.35.
47. Jacob Shatzer, "Fake and Future 'Humans': Artificial Intelligence, Transhumanism, and the Question of the Person," *Southwestern Journal of Theology* 63, no. 2 (Spring

2021): 127–46; Aura-Elena Schussler, "Artificial Intelligence and Mind-Reading Machine—Towards a Future Techno-Panoptic Singularity," *Postmodern Openings* 11, no. 4 (2020): 334–46, https://doi.org/10.18662/po/11.4/239; "Merging With the Machines: Information Technology, Artificial Intelligence, and the Law of Exponential Growth, Part 2," *World Future Review* 2, no. 2 (May 2010): 57–61, https://doi.org/10.1177/194675671000200209; and "Merging with the Machines: Information Technology, Artificial Intelligence, and the Law of Exponential Growth, Part 1," *World Future Review* 2, no. 1 (March 2010): 61–66, https://doi.org/10.1177/19467567 1000200107.

48. Chris C. Demchak, "China: Determined to Dominate Cyberspace and AI," *Bulletin of the Atomic Scientists* 75, no. 3 (2019): 99–104, https://doi.org/10.1080/00963402 .2019.1604857; Arthur Herman, "Why China Is Winning the War for High Tech," *National Review*, 1 November 2021; and Stephanie Petrella, Chris Miller, and Benjamin Cooper, "Russia's Artificial Intelligence Strategy: The Role of State-Owned Firms," *Orbis* 65, no. 1 (Winter 2021): 75–100, https://doi.org/10.1016/j.orbis.2020.11.004.
49. However, a network of humans or other form of intelligent entities, if designed to work collectively as a single enterprise, could form a singleton by design. This does not seem to be what the metaverse is suggested for and would potentially remove some of the primary attractive aspects of what the metaverse represents in theory.
50. Arguably, the first transhuman entity could be an organic human with superintelligence accomplished by enhancement. This hybrid would be both human and artificial and could simultaneously cross the transhuman barrier as well as the singleton barrier. In this context, the first transhuman would still face these same singleton dilemmas on security, humanity, and the future of all societies.
51. James Johnson, "Artificial Intelligence, Drone Swarming and Escalation Risks in Future Warfare," *RUSI Journal* 156, no. 2 (2020): 26–36; Alessandro Gagaridis, "Warfare Evolved: Drone Swarms," *Geopolitical Monitor*, 10 June 2022; and Ben Zweibelson, "Let Me Tell You About the Birds and the Bees: Swarm Theory and Military Decision-Making," *Canadian Military Journal* 15, no. 3 (Summer 2015): 29–36.
52. Bostrom, *Superintelligence*, 96.
53. *Rick and Morty*, season 2, episode 3, "Auto Erotic Assimilation," created by Justin Roiland and Dan Harmon, aired 9 August 2015; and *Star Trek: The Next Generation,* season 2, episode 16, "Q Who," directed by Rob Bowman, aired 6 May 1989.
54. Garcia, "Lethal Artificial Intelligence and Change"; Geraci, "Apocalyptic AI"; and Brian Molloy, "Project Governance for Defense Applications of Artificial Intelligence: An Ethics-Based Approach," *PRISM* 9, no. 3 (2021): 107–20.
55. This also creates a paradox in that, arguably, a superintelligent singleton could make the same decisions and actions as a regular AI aggressor, as the nonsuperintelligent humans lack the cognitive abilities to even distinguish the two.
56. Giampiero Giacomello, "The War of 'Intelligent' Machines May Be Inevitable," *Peace Review: A Journal of Social Justice* 33, no. 2 (2021): 284, https://doi.org/10.1080/104 02659.2021.1998860.
57. *2001: A Space Odyssey*, directed by Stanley Kubrick (Stanley Kubrick Productions, 1968); *Ex Machina*, directed by Alex Garland (London, UK: Film4 and DNA Films, 2014); *I, Robot*, directed by Alex Proyas (Los Angeles, CA: 20th Century Fox, 2004); and *WALL-E*, directed by Andrew Stanton (Emeryville, CA: Pixar Animation Studios, 2008).
58. Adam Cutler, "IBM Design—Operationalizing Artificial Intelligence" (lecture, JSOU design lecture series 2019, JSOU Campus, Tampa, Florida, 14 November 2019). Cutler also lectured on this topic as the closing keynote speaker to the U.S. Space Command at the USSPACECOM Commander's Conference held in May 2022 at Peterson Space Force Base.
59. Ben Zweibelson, "One Piece at a Time: Why Linear Planning and Institutionalisms Promote Military Campaign Failures," *Defence Studies Journal* 15, no. 4 (2015): 360–75, https://doi.org/10.1080/14702436.2015.1113667; Ben Zweibelson, "An Awkward Tango: Pairing Traditional Military Planning to Design and Why It Currently

Fails to Work," *Journal of Military and Strategic Studies* 16, no. 1 (2015): 11–41; and Ben Zweibelson, "Preferring Copies with No Originals: Does the Army Training Strategy Train to Fail?," *Military Review*, January–February 2014, 15–25.

60. Christopher Paparone, "How We Fight: A Critical Exploration of US Military Doctrine," *Organization* 24, no. 4 (2017): 516–33, https://doi.org/10.1177/135050841769385; Zweibelson, "An Awkward Tango"; Zweibelson, "One Piece at a Time"; Ben Zweibelson, "Rose-Tinted Lenses: How American Functionalist Strategy Inhibits Our Appreciation of Complex Conflicts," *Defence Studies Journal* 16, no. 1 (2016): https://doi.org/10.1080/14702436.2016.1147924; and James Der Derian, "From War 2.0 to Quantum War: The Superpositionality of Global Violence," *Australian Journal of International Affairs* 67, no. 5 (2013): 573–74, https://doi.org/10.1080/10357718.2013.822465.
61. Haridimos Tsoukas, *Complex Knowledge: Studies in Organizational Epistemology* (New York: Oxford University Press, 2005), 213–16.
62. Paparone, *The Sociology of Military Science*, 16–22; and Antoine Bousquet and Simon Curtis, "Beyond Models and Metaphors: Complexity Theory, Systems Thinking and International Relations," *Cambridge Review of International Affairs* 24, no. 1 (2011): 43–62, https://doi.org/10.1080/09557571.2011.558054.
63. Lorraine Daston, *Rules: A Short History of What We Live By* (Princeton, NJ: Princeton University Press, 2022), 63; Sébastien le Prestre de Vauban, *The New Method of Fortification, as Practised by Monsieur de Vauban, Engineer-General of France. Together with a New Treatise of Geometry. The Fifth Edition, Carefully Revised and Corrected by the Original*, 5th ed. (1722; repr., Farmington Hills, MI: Gale ECCO, 2018); and Henry Guerlac, "Vauban: The Impact of Science on War," in *Makers of Modern Strategy*.
64. Tsoukas, *Complex Knowledge*, 213–16.
65. Tsoukas, *Complex Knowledge*, 213–14.
66. Der Derian, "From War 2.0 to Quantum War," 573.
67. Leaders interchange "complex" and "complicated" while military doctrine states that complex challenges can be solved through identifying the complex problem, which mangles complexity theory with earlier, Newtonian inspired war frames. Military objectives and goals are set into complex systems with linear lines of effort or action, while other parts of doctrine mention nonlinearity and emergence as how complex systems behave. All too often, military language is convoluted, with terms stripped of their original meaning so that they comply with preexisting doctrinal standards despite the concepts breaking with such beliefs entirely.
68. Der Derian, "From War 2.0 to Quantum War," 577.
69. James William Gibson, *The Perfect War: Technowar in Vietnam*, 1st ed. (Boston: Atlantic Monthly Press, 1986), 462; and Alex Ryan, "A Personal Reflection on Introducing Design to the U.S. Army," *Medium* (blog), 4 November 2016.
70. Tsoukas, *Complex Knowledge*, 74.
71. Siniša Malešević, *The Sociology of War and Violence* (Cambridge, UK: Cambridge University Press, 2010), 28, https://doi.org/10.1017/CBO9780511777752.
72. Shimon Naveh, Jim Schneider, and Timothy Challans, *The Structure of Operational Revolution: A Prolegomena* (Fort Leavenworth, KS: Booz Allen Hamilton, 2009), 35–36.
73. Jensen, Whyte, and Cuomo, "Algorithms at War," 536–37.
74. Elizabeth Kier, *Imagining War: French and British Military Doctrine Between the Wars* (Princeton, NJ: Princeton University Press, 1997); Correll, "The Billy Mitchell Court-Martial"; and David French, *The British Way in Warfare, 1688–2000* (Cambridge, MA: Unwin Hyman, 1990).
75. Gibson, *The Perfect War*.
76. Builder, *The Masks of War*.
77. Blake Stilwell, "The US Air Force's 'Rods from God' Could Hit with the Force of a Nuclear Weapon—with No Fallout," *Business Insider*, 4 February 2019.
78. These changes would qualify as weak emergence (fads, bandwagon effect, bubbles, and crashes) or in rare cases such as nuclear weapons and computers, multiple emergence

(many feedback loops, both positive and negative). See Jochen Fromm, "Types and Forms of Emergence" (research paper, Distributed Systems Group, Electrical Engineering and Computer Science, Universitat Kassel, Germany, 13 June 2005), 1–23.

79. Der Derian, "From War 2.0 to Quantum War," 573.
80. The author regularly teaches systemic design, complexity theory, and systems theory at the Marine Corps War College, National Defense University, Canadian Forces College, Australian Command and Staff College, as well as numerous NATO and European military programs. These topics are infrequently added to curriculum and often cause disagreement with students and faculty on their relevance and suitability with other educational priorities. For examples of this, see Anna Grome, Beth W. Crandall, and Louise Rasmussen, *Incorporating Army Design Methodology into Army Operations: Barriers and Recommendations for Facilitating Integration*, Research Report 1954 (Washington, DC: Department of the Army, 2012); Aaron P. Jackson, Ben Zweibelson, and William Simonds, "Intellectual Spring Cleaning: It's Time for a Military 'Do Not Read' List; and Some Sources That Should Be on That List," *Defence Studies* 18, no. 2 (2018): 131–46, https://doi.org/10.1080/14702436.2018.1461563; and BGen Shimon Naveh (Ret), interview with Matt Matthews, 1 November 007.
81. Many of the citations in this article illustrate this tension. Few concepts from complexity theory, postmodern philosophy, or sociology (social paradigm theory in particular) are ever integrated into military education, training, and doctrine. The author states this as factual based on more than a decade of introducing such ideas into multiple war colleges and professional military education programs.
82. Haridimos Tsoukas and Mary Jo Hatch, "Complex Thinking, Complex Practice: The Case for a Narrative Approach to Organizational Complexity," *Human Relations* 54, no. 8 (2001): 979–1013, https://doi.org/10.1177/0018726701548001; and Tsoukas, *Complex Knowledge*.
83. Pick, "Rethinking Organization Theory," 807; Peter L. Berger and Thomas Luckmann, *The Social Construction of Reality: A Treatise in the Sociology of Knowledge* (New York: Anchor Books, 1966); and Tsoukas and Hatch, "Complex Thinking, Complex Practice."
84. Granted, currencies are replaced and soon after the fall of the Saddam Hussein regime in Iraq, a new government formed, and a new currency was created. This occurred with the Iraqi Army in parallel as the old Baathist one was dismantled and disarmed so that a new version could be assembled for Iraqi security. This refers to the ancient rai or fei stones found on the Yap Islands in Micronesia.
85. This refers to the ancient rai or fei stones found on the Yap Islands in Micronesia. Modern economists view rai stones as a form of money despite their massive size preventing movement of them.
86. Fromm, "Types and Forms of Emergence," 18; and Schön and Rein, *Frame Reflection*, 30.
87. Harari, *Sapiens*.
88. Malešević, *The Sociology of War and Violence*, 56.
89. Bostrom, *Superintelligence*, 95–109.
90. Schussler, "Artificial Intelligence and Mind-Reading Machines," 342.
91. Garcia, "Lethal Artificial Intelligence and Change," 339.
92. Noreen Herzfeld, "Can Lethal Autonomous Weapons Be Just?," *Journal of Moral Theology* 11, Special Issue no. 1 (2022): 72, https://doi.org/10.55476/001c.34124.
93. Jensen, Whyte, and Cuomo, "Algorithms at War," 528.
94. Michal Krelina, "Quantum Warfare: Definitions, Overview and Challenges," *EPJ Quantum Technology* 8, no. 24 (2021): https://doi.org/10.1140/epjqt/s40507-021-00113-y.
95. Pugh, Soros, and Stanley, "Quality Diversity," 8.
96. Bostrom, *Superintelligence*, 140–70.
97. Spectacular stereotypes of this fashion are iconic in film history, whether when actor Slim Pickens rides a nuclear bomb while whipping his Stetson hat toward the target in

Dr. Strangelove or when one-dimensional murderous military commanders in *Avatar* represent social commentary by using historical tropes. Anatol Rapoport, *The Origins of Violence: Approaches to the Study of Conflict* (New Brunswick, NJ: Transaction Publishers, 1995), 258.

98. Sophie-Charlotte Fischer and Andreas Wenger, "Artificial Intelligence, Forward-Looking Governance and the Future of Security," *Swiss Political Science Review* 27, no. 1 (2021): 174, https://doi.org/10.1111/spsr.12439.
99. Fromm, "Types and Forms of Emergence."
100. Geraci, "Apocalyptic AI," 150.
101. Schussler, "Artificial Intelligence and Mind-Reading Machines," 343.
102. Bostrom, *Superintelligence*, 140–55.
103. Herzfeld, "Can Lethal Autonomous Weapons Be Just?," 84.
104. Bostrom, *Superintelligence*, 150.
105. *Independence Day*, directed by Roland Emmerich (Los Angeles, CA: 20th Century Fox, 1996).
106. Geraci, "Apocalyptic AI," 157.
107. Herzfeld, "Can Lethal Autonomous Weapons Be Just?"; and Der Derian, "From War 2.0 to Quantum War."
108. Jean Baudrillard, *Simulacra and Simulation*, trans. Sheila Glaser (Ann Arbor: University of Michigan Press, 2001).
109. *Defense Primer: Emerging Technologies* (Washington, DC: Congressional Research Service, 2021).
110. For neo-Clausewitzian concepts, readers can refer to the existing references in this piece including Chia, Holt, Rapaport, Paparone, Naveh, and Der Derian to start with.
111. Geraci, "Apocalyptic AI," 152.
112. Herzfeld, "Can Lethal Autonomous Weapons Be Just?," 74.
113. Schussler, "Artificial Intelligence and Mind-Reading Machines," 335.
114. Herman, "Why China Is Winning the War for High Tech," 34.
115. The famous quote by Secretary of Defense Donald Rumsfeld in 2002 likely comes from complexity theory and contemporary ideas that would lead to popular books such as Nassim Nicholas Taleb's *The Black Swan* within the same decade. The American public would slowly gain exposure to complexity theory ideas that tended to be in paradox or dismantle widely popular and Newtonian framed worldviews. See David Graham, "Rumsfeld's Knowns and Unknowns: The Intellectual History of a Quip," *Atlantic*, 27 March 2014; and Nassim Nicholas Taleb, *The Black Swan: The Impact of the Highly Probable* (New York: Random House, 2007).
116. Der Derian, "From War 2.0 to Quantum War," 575; and Gilles Deleuze and Felix Guattari, *A Thousand Plateaus: Capitalism and Schizophrenia*, trans. Brian Massumi (Minneapolis: University of Minnesota Press, 1987). Deleuze and Guattari would echo Foucault's and other postmodern variations on deconstructing Clausewitz.
117. Der Derian, "From War 2.0 to Quantum War," 575.
118. In an easy but simple manner to demonstrate this, examine the required and suggested readings for any professional military program in 2023 and highlight the number of postmodern authors and topics. More often than not, this field is entirely nonexistent in conventional military education. Yet, these authors are some of the few outside of select technological futurists willing to explore such radical transformations of future war. See Paul Virilio, *The Information Bomb*, trans. Chris Turner, Radical Thinkers (New York: Verso, 2005); James Der Derian, "Virtuous War/Virtual Theory," *International Affairs (Royal Institute of International Affairs 1944–)* 76, no. 4 (October 2000): 771–88; and Michel Foucault, "Discourse and Truth: The Problematization of Parrhesia" (lecture, University of California at Berkeley, November 1983); Daniel Cockayne, Derek Ruez, and Anna J. Secor, "Thinking Space Differently: Deleuze's Möbius Topology for a Theorisation of the Encounter," *Transactions of the Institute of British Geographers* 45 (2020): 194–207, https://doi.org/10.1111/tran.12311; and Scott Lawley, "Deleuze's Rhizome and the Study of Organization: Conceptual Movement and an

Open Future," *Tamara: Journal of Critical Postmodern Organization Science* 3, no. 4 (2005): 36–49.

119. Marshall, "Some Thoughts on Military Revolutions," 2–3.

120. Indeed, as arrogant as our species often is, the notion that future foreign policy might be developed by human and AI minds is disconcerting. Further still, a superior AI mind might generate entirely novel theory that humans cannot comprehend fully or even at all.

121. *The Matrix*, directed by Lana Wachowski and Lilly Wachowski (Burbank, CA: Warner Brothers, 1999).

PART II
Whale Songs of Wars Not Yet Waged
The Demise of Natural-Born Killers through Human-Machine Teamings Yet to Come

Ben Zweibelson, PhD

Abstract: Current human-machine dynamics in security affairs positions the human operator in the loop with artificial intelligence to conduct decisions and actions. As technological advancements in AI cognition, speed, and weapon sophistication increase, human operators are increasingly being shifted to an on the loop where AI takes more responsibility in warfare and defense decisions, whether tactical or even strategic. Human operators are also falling off the loop, trailing enhanced AI systems as the biological and physical limits because humans are not the same for artificial intelligence in narrow applications. Those likely will expand toward general AI in the coming decades, presenting significant strategic, organizational, and even existential concerns. Further, how natural humans respond and engage with increasingly advanced, even superintelligent AI as well as a singularity event will feature disruptive, transformative impacts on security affairs and even at a philosophical level discerning what war is.
Keywords: artificial intelligence, AI, warfare, singularity, transhumanism, singleton, human-machine teaming

Warfare has always been changing as humans develop new ideas, technology, and otherwise expand their range of abilities to manipulate reality to their advantage and creativity. Just as Homo sapiens prove

Dr. Ben Zweibelson is the director of the U.S. Space Command's Strategic Innovation Group at Peterson Space Force Base, CO. A retired Army infantry officer with combat tours in Iraq and Afghanistan, he earned the Combat Infantryman Badge, Master Parachutist Badge, Pathfinder Badge, Air Assault Badge, the Ranger Tab, four Bronze Star medals, and various awards and citations in his 22 years combined service. He previously worked for U.S. Special Operations Command for seven years, running all design education, theory, and outreach for the Joint Special Operations University. He has a doctorate in philosophy, three master's degrees, and an undergraduate degree in graphic design. He has two design books forthcoming in the summer of 2023.
Journal of Advanced Military Studies vol. 14, no. 1
Spring 2023
www.usmcu.edu/mcupress
https://doi.org/10.21140/mcuj.20231401002

astoundingly adaptive and clever in how to produce art, beauty, and generously extend quality of life for the species, they continue to also be devastatingly capable of conjuring up with evermore horrific and powerful ways to engage in organized violence against those they are in competition or conflict with. Yet, the twenty-first century is wholly unique in that humanity now has the technological keys to unlock something previously unreachable. Civilization, human existence, and perhaps war may move into what previously could only be captured in fantasy, science fiction, or ideological promises and magic.

The title of this article is provocative and draws inspiration from Elon Musk's comment on how humans will be rapidly outpaced by advanced artificial intelligence (AI). Musk remarked that "human speech to a computer will sound like a very slow tonal wheezing, kind of like whale sounds" due to expected lightning fast processing speeds and how AI will move from specialized applications into general intelligence contexts where humans are slower, more error-prone, and unable to compete in every conceivable way.[1] This article operates as a thought piece, designed to stimulate deep thinking not on the short-term or localized contexts for immediate wars of the next decade, but onward and outward to radical, potentially existential concerns of where humans and warfare technology might lead to in a century, perhaps less.

Previous efforts by military theorists on how technological developments will change warfare fundamentally fixate exclusively upon humans directing said change so that new wars demonstrate human mastery beyond earlier warfare efforts of less advanced human combatants. This orientation on humanity at war with their own species is consistent, whether one considers the development of the stirrup, firearms, the Industrial Revolution, or even the First Quantum Revolution and the detonation of atomic weapons in 1945. Throughout all these transformative periods, warfare characteristics, styles, and indeed the scale and scope of war effects have changed, but humans remained the sole decision makers and operators central to all war activities. The key distinction in how advanced (likely general AI, which can match or exceed human cognitive abilities in all possible ways) intelligent machines and/or human-machine teaming (the rise of transhumanism) may develop is that future battlefields may push human decision makers and even operators to the sidelines, including entirely removing from any direct involvement in some potential war developments as this article will explain. While the targets of such a war may still include human populations and nations, the battlefield may finally become untenable for natural-born human cognition, survivability, and capability. This would be a dramatic shift from the last 40 centuries or more of human-directed, human-waged warfare where technology remained a tool firmly in the grasp of a hand made of flesh and bone.

We are poised, if we can just survive the next few decades where all sorts of modern existential threats remain horrifically available—for a new chapter in humanity and organized violence. Indeed, Homo sapiens will shift from constructing more sophisticated and lethal means to impose behavior changes and

force of will (security, foreign policy, and warfare) to one where the means become entirely dissimilar, emergent ends in themselves. Our tools of warfare will be able to think for themselves and think about us, as well as think about war in starkly dissimilar, likely nonhuman terms. Whereas war has been an exclusively human created, socially constructed, and human exercised phenomenon, how we frame what war is (and is not) revolves around what humans believe it to be.[2] This may not apply to nonhuman entities, nor would it be impossible for a superintelligent artificial entity to conceptualize something currently beyond our own violent imaginations. Unlike when we split the atom and quickly weaponized that technological marvel, there will not be the same control and command of weapons that can decide against what we wish, or even what we might be able to grasp in complex reality.

Humans will transition from ever-capable masters of increasingly sophisticated war tools toward less clever, less capable, and insufficient handlers of an increasingly superior weaponized capability that in time will elevate, transform, or potentially enslave (or eliminate) Homo sapiens into something different, possibly unrecognizable. War, as a purely human construct that has been part of humanity since inception, will change as well. Note that war is not interchangeable with warfare, in that war is the human-designed, socially constructed, and physically waged activities of organized violence, while the process of engaging in any manifestation of war becomes the exercise of warfare per the established belief system in operation by those using a particular social paradigm.[3] Humans currently use technology and knowledge to understand what war is and subsequently wield technologically generated abilities derived from resources to produce desired effects within complex reality that accomplish various desires of politics and societies. The dramatic shift of technology from a human-controlled tool for effect into its own designs and motives to accomplish unrelated (or unimagined) ends to itself will be potentially the ultimate (or final) change in warfare from a human-centered perspective. This could occur gradually, even invisibly, or suddenly and with profound disruption. These changes will not occur overnight, nor likely in the next decade or two, which sadly renders such discussions out of the essential and toward the fantastical. For the military profession, this reinforces a pattern of opting to transform to win yesterday's war faster instead of disruptively challenging the force to move away from such comfort and familiarity toward future unknowns that erode or erase favorite past war constructs.[4]

The next century will not be like past periods of disruptive change such as the development of firearms, the introduction of internal combustion engines, or even the arrival of nuclear weapons. While today's semiautonomous cruise missile cannot suddenly decide to go study poetry or join an antiwar protest, future AI systems in the decades to come will not be bounded by such limitations. Past revolutions in warfare involve technological and sociological transformations that replaced a legacy mode of human-directed warfare with a newer, more lethal, faster, yet still human-centered warfare process. The upcoming rev-

olution in artificial intelligence and human-machine teaming in warfare may become the last revolution that humans will start and possibly one that they are unable to finish or influence the path beyond what they can conceptualize or articulate at whale song speeds.

As such, critics might dismiss such thoughts outright as science fiction claptrap that is inapplicable to contemporary concerns such as the Russian invasion of Ukraine or the saber-rattling of China over Taiwan. Such a reaction misses the point, as the AI enabled war tools of this decade are like babies or toddlers compared to what will likely develop several decades beyond our narrow, systematic viewpoints. In developing defense areas such as cyberspace, deep ocean locations, and space, humans are ill-equipped to function in these spaces.[5] The human body is not designed for these areas, and faster, more robust AI have myriad operational advantages just now coming into what is possible. Codeveloper of Skype and computer programmer Jaan Tallinn states it bluntly: "silicon-based intelligence does not share such concerns about the environment. That's why it's much cheaper to explore space using machine probes rather than 'cans of meat'."[6] In turn, this is why militaries perpetually chase the next silver bullet and secure funding to conduct moon shots, and these already include advanced AI weaponized systems that may replace almost every human operator on today's battlefield. The new AI system, if not developed and secured by our side, surely will be designed by competitors, ensuring a perpetual AI arms race driven by national self-interests over any potential ethical, moral, or legal complications.

Yet, when we seek to develop new weapons of war without putting in the necessary long-term, philosophical work on where we might end up, we fall into the trap that Der Derian warns of for societies excited about new technologies but uninterested in engaging in deep philosophical ponderings on the consequences of those new war tools:

> When critical thinking lags behind new technologies, as Albert Einstein famously remarked about the atom bomb, the results can be catastrophic. My encounters in the field, interviews with experts, and research in the archives do suggest that the [Military Industrial Media Entertainment Network], the [Revolutions in Military Affairs,] and virtuous war are emerging as the preferred means to secure the United States in highly insecure times. Yet critical questions go unasked by the proponents, planners, and practitioners of virtuous war. Is this one more attempt to find a technological fix for what is clearly a political, even ontological problem? Will the tail of military strategy and virtual entertainment wag the dog of democratic choice and civilian policy?[7]

This article presents a framing of how nations currently understand the ever-developing relationship between themselves and artificially intelligent-enabled machines on the battlefields of today and where and how those likely will shift in the decades to come. Some developments will retain nearly all of the existing and traditionally recognized hallmarks of modern warfare, despite things speed-

ing up or becoming clouded with disturbingly unique technological embellishments to what remains a war of political and societal desires to change the behaviors and belief systems of others. Other paths lead to never-before-seen worlds where humans become increasingly delegated to secondary positions in future battlefields and perhaps booted off those fields entirely. War, as a human creation, may cease to be human, and morph into constructs alien or incomprehensible to the very creators of organized violence for socially constructed wants.

More than 40 Centuries of Precedence: War Is a Decidedly Human Affair

Humans have for tens of thousands of years curated and inflicted on one another a specific sort of organized violence known as war that otherwise does not exist in the natural world. More than 30,000 years ago, a cognitive revolution occurred that set into motion the rise of humans as a species not entirely dependent on biology, with historical narratives needed to explain developments and accomplishments.[8] Prehistoric humans learned how to harness fire, create basic tools, shelter from the elements, and began a gradual journey toward ever-increasingly sophisticated societies.[9] Change occurred gradually, with agriculture and the establishment of cities commencing around 10,000 years ago; this would produce the first recorded wars that differed from other types of violence.[10] The invention of writing (3,000 years ago) would eventually shift oral accounts of these wars into more refined, structured forms that could be studied as well as extended beyond internalization of each living generation.[11] Without this cognitive revolution, humans would not have been capable of creating societies, belief systems, rules of law, politics, religions, or war. War is a decidedly human invention, and it has been wielded by human desires, beliefs, symbols, and conceptual models exclusively since its inception. We created it, use it, and own it, at least for now.

Yet, across these thousands of generations of Homo sapiens that would collectively produce modern societies of today, change occurred quite slowly until the last 500 years where a scientific revolution propelled Western Europe from obscurity into a technological, economic, and imperial juggernaut.[12] Muscle and natural power (wind, fire, water) were the primary energy source for much of the collective human experience of warfare, with technological advancements only occurring in the last several hundred years with the invention of scientific methods and the Industrial Revolution that followed. Fossil fuels soon replaced muscle power, and the chemical power of gunpowder would replace edged weaponry with bullets, artillery, and more. Steam locomotion gave way to faster systems such as internal combustion engines and eventually nuclear power.[13] Technology as well as organizational, cultural, and conceptual things have changed dramatically across this vast span of time, but humans have forever remained the sole decision maker in every act of warfare until very recently. This is where things will accelerate rapidly and potentially we may be entering the last century where humans even matter on future battlefields at all.

As soon as early humans realized how to manipulate their environment through inventing tools, they gained an analog function to greatly increase their own lethality to include waging war upon one another. The tools have indeed changed, but the relationship of the human to the tool has remained firmly in a traditional ends-ways-means dynamic. Humans use technology, communication, and organization to decide and act to attempt to accomplish goals through various ways and using a wide range of means at their disposal. Until the First Quantum Revolution that would coincide with the Second World War, humans were the sole decision makers at the helm of quite sophisticated yet entirely analog machines of war.[14] Once computers first became possible (beyond earlier analog curiosities), humans gained something new within their decision-making cycle for warfare activities from the tactical up through even grand strategic levels—the artificially intelligent machine partner. At first, such systems were cumbersome, slow, and could only perform calculations, but over time they have migrated into central roles for how modern society now depends on this technology for a wide range of effects.[15] The rise of AI brings with it the first encounter for humanity of an entity with the potential to cooperate, collaborate, compete, and perhaps leap well beyond our own conceptual limits in all endeavors to include warfare.

The Battlefield Suddenly Gets a Bit More Crowded

Artificial intelligence has many definitions, and modern militaries often are preoccupied with narrow subsets of what AI is and is not, according to competing belief systems, value sets, as well as organizational objectives and institutional factors of self-relevance. Peter Layton provides a broad and useful definition: "AI may do more- or less- than a human . . . AI may be intelligent in the sense that it provides problem-solving insights, but it is artificial and, consequently, thinks in ways humans do not."[16] Layton considers AI more by the broad functions such technology can perform than by its relationship to human capabilities. This indeed is often how current defense experts and strategists prefer to frame AI systems in warfare; the human is teamed with a machine that provides augmentation, support, and new abilities to perform some goal-oriented task that non-AI enabled warfighters would be insufficient or less lethal performing.

Artificial intelligence is also broadly distinguished into whether it is narrow or general with respect to human intelligence. Narrow AI equals or vastly exceeds the proficiency that the best human is capable of doing for specific tasks within a particular domain and only in clearly defined parameters that are unchanging. Narrow AI can now beat the best human players of chess and other games, with IBM's Watson defeating the best *Jeopardy!* trivia game players in 2011 as an example. However, narrow AI is fragile, and if the rules of the game were changed or the context transformed, the narrow AI programming cannot go beyond the limits of the written code.[17] General AI, as a concept and benchmark yet realized in any existing AI system, must equal or exceed the full range of human performance abilities for any task, in any domain, in what must be

a fluid and ever-changing context of creativity, improvisation, adaptation, and learning.[18] Such an AI is decades away, if ever possible. Just as likely, a devastating future war waged with weaponized AI short of general intelligence could knock society back into a new Stone Age, or perhaps humanity might drift away from AI-oriented technological advances seeking general AI capabilities.[19] Existential warfare could come at the hand of humans directing slightly less intelligent AI systems, or the dynamic could flip and the slightly less intelligent humans could be used as tools, targets, or for purposes beyond our imagination.

AI is constantly being developed, with many military applications already well established and those on the immediate horizon for battlefields in the next decade. Much of what currently exists was produced in what is called "first-wave AI"—narrow programming created in conjunction between the computer designers writing the code and the experts in the field or task that the narrow AI system is attempting to excel at. More recently, "second-wave AI" uses machine learning where "instead of programming the computer with each individual step . . . machine learning uses algorithms to teach itself by making inferences from the data provided."[20] Machine learning is powerful, working in a special way where human programmers do not have to set it up. Yet, this creates the paradox that machine learning quickly can exceed the programmer's ability to track and understand how the AI is learning.[21] This sort of machine learning can occur in either a supervised or unsupervised methodology, where supervised learning systems are given labeled and highly curated data. The AI is told what to do, how to accomplish it, and progress is diligently monitored and analyzed by human supervisors. This is time and resource intensive, but supervised machine learning can achieve extremely high performance in narrow applications.

Unsupervised learning unleashes the AI and the AI identifies patterns for itself, often moving in emergent pathways well outside the original expectations of the programmer. Layton remarks: "An inherent problem is it is difficult to know what data associations the learning algorithm is actually making."[22] IBM's Adam Cutler, in a lecture to military leadership at U.S. Space Command, provided the story of how two chatbots created by programmers at Facebook quickly developed their own language and began communicating and learning in it. The Facebook programmers shut the system down as they had lost control and could not understand what the chatbots were doing. Cutler stated that "these sorts of developments with AI are what really do keep me up at night." His comment was both serious and simultaneously elicited audience laughter, as the panel question posed was: "What sorts of things keep you up at night?"[23]

While the instance of chatbots going rogue with a new language might be overblown, Cutler and other AI experts warn of the dangers of unsupervised learning in AI development, and caution that while anything remotely close to general AI intelligence is still far-off in the future, there are profound ethical, moral, and legal questions to begin considering today.[24] With this brief summary of AI put into perspective, we shall move to how the military currently understands and uses AI in warfare, and where it likely is morphing toward next.

How Human-Machine Teaming Is Currently Framed

AI systems can operate autonomously, semiautonomously, or remain in the traditional sense where, just like your smartphone or smart device awaits your command, operate in a passive mode of activity. While it may seem unnerving that your Alexa device is perpetually listening to background conversations, it is programmed to scan for specific words that trigger clearly defined and quite narrow actions. While passively awaiting directive cues and proactive (semiautonomous, autonomous) modes feature different relationship dynamics between the AI and humans, the following three are well recognized in current military applications of AI systems.[25]

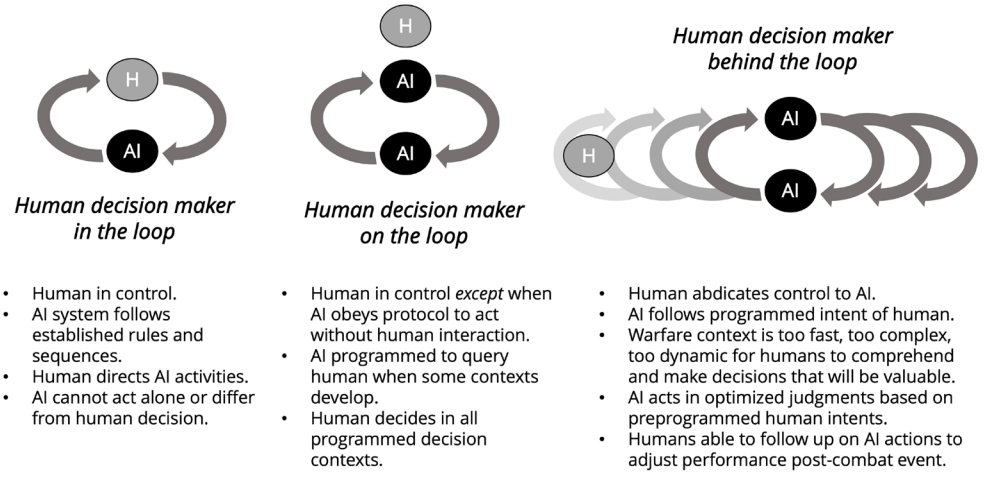
The original mode used since the earliest protocomputer enablers (as well as most any analog augmentation in warfare) positions the human as the key decision maker in the cycle of thought-action-reflection. This is best known as human-in-the-loop, and it expresses a dynamic where the human is central to the decision-making activities. The AI can provide exceptional contributions that exceed in narrow ways the human operator's capabilities, but that AI does not actively do anything significant without a design where the human intervenes and provides guidance or approval to act.

While the human-in-the-loop remains the most common and, for ethical, moral, and legal reasons the most popular mode of human-machine teaming for warfare, a second mode has also emerged with recent advances in AI technology.[26] Termed human-*on*-the-loop, the AI takes a larger role in decision making and consequential action where the human operator is either monitoring activities, or the AI system is programmed to pause autonomous operations when particular criteria present the need for human interruption. In these situations, the AI likely has far faster abilities to sense, scan, analyze, or otherwise interpret data beyond human abilities, but there still are fail-safe parameters for the human operators to ensure overarching control. An autonomous defense system might immediately target incoming rocket signatures with lethal force, but a human operator may need to make a targeting decision if something large like an aircraft is detected breaching defended airspace.

A third mode is only now coming into focus, and with greater AI technological abilities as well as increasing speed, scale, and scope of new weaponry (hypersonic weapons, swarms and multidomain, networked human-machine teams) a fully autonomous AI system is required. Termed human-*out-of*-the-loop, this differs from what is nonpejoratively referred to as dumb technology such as airbags that automatically activate when certain criteria are reached. Truly autonomous, general intelligence AI systems would replace the human operator entirely and are designed to function beyond the cognitive abilities of even the smartest human at what are currently narrow parameters. While many use out-of-the-loop or off-the-loop, this article substitutes *behind*-the-loop to introduce several increasingly problematic human-machine issues on future battlefields. Figure 1 illustrates these three modes below.

An autonomous AI system functioning in narrow or even general AI appli-

Figure 1. Contemporary framings for human-machine teaming in warfare

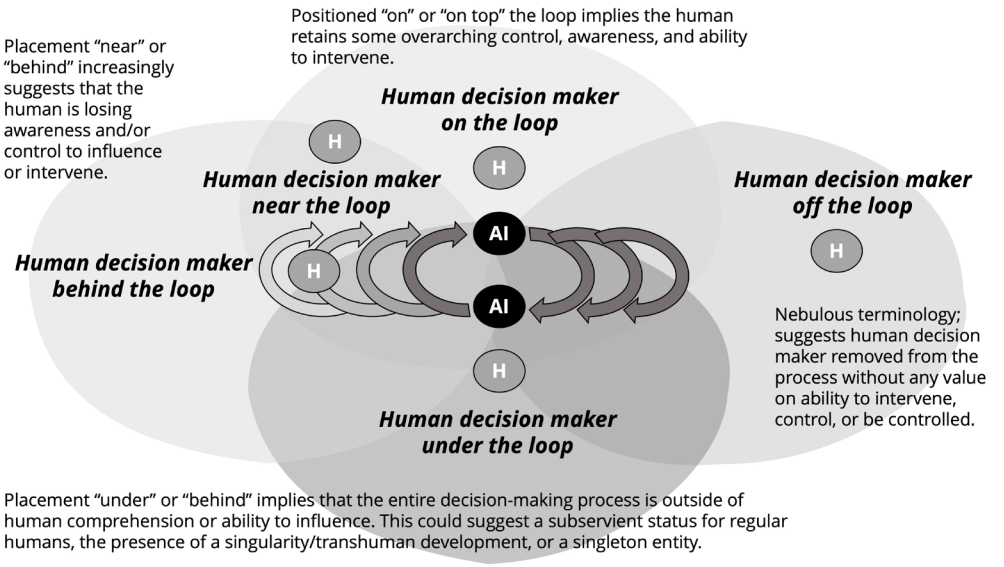


Source: courtesy of the author, adapted by MCUP.

cations will, as technological and security contexts demand, potentially move from supervised to unsupervised machine learning profiles that access ever growing mountains of data on the battlefield. Consider the average daily volume of new tweets by worldwide Twitter users that average 330 million monthly users with 206 million of those users tweeting daily, producing a daily average of more than 500 million tweets worldwide in hundreds of different languages and countries.[27] It would be impossible for human-in-the-loop monitoring, while human-on-the-loop is also expensive, slow, and often subjective. Twitter, like many social media platforms, has many autonomous AI systems filtering, analyzing, and often taking down spam, fake accounts, bots, and other harmful content without human intervention. This is not without risks and concerns, yet Twitter quality control is inevitably chasing behind the autonomous work of far faster, future AI systems that should scale to enormous levels beyond what an army of human reviewers might possibly match. However, fighting spam bots and fake accounts on Twitter is not exactly the same as autonomous drones able to decide on lethal weapon strikes independent of human operators.

Human-machine teaming currently exists in all of the three representations in figure 1, with the preponderance in the first depiction where human operators remain the primary decision makers coupled with AI augmentation. Drone operators, satellite constellations, advanced weapon systems that auto-aim for human operators to decide when to fire, as well as bomb-diffusing robots worked by remote control are common examples in mass utilization. Humans on the loop abound as well, with missile defense systems, antiaircraft, and other indirect fire countermeasures able to function with human supervision or just with human engagement for unique conditions outside normal AI parameters. The increased sophistication and abilities of narrow AI systems as well as the

Figure 2. Implications of metaphors in human-machine teaming descriptions



Source: courtesy of the author, adapted by MCUP.

frightening speeds achieved by hypersonic and advanced weaponry and the rise of devastating swarm movements that collectively would overwhelm any single human operator may now be addressed with autonomous weapon systems, if militaries and their political oversight concur with the risks. These are not without significant ethical, legal, and technological debate in security affairs.

How we express the human-machine teaming dynamic requires a combination of our conceptual models along with precise terminology that is underpinned by metaphoric devices. In figure 2, many of the terms currently in use as well as some new ones introduced in this article place the human *on* the loop, *off* the loop, *near* the loop, *behind* the loop, or potentially *under* the loop depending on the warfare context and metaphoric configuration of the loop participants. Aside from the *in* the loop configuration that has been the foundational structure for human decision makers to direct command and control with an AI supporting system, all these other placements for a human operator reflect changing technological potentials as well as the increasingly uncertain future of warfare. Such variation and uncertainty imply significant ethical, moral, legal, and potentially paradigm-changing, existential shifts in civilization.[28] Figure 2 attempts to systemically frame some of these tensions, differences, and implications with how militaries articulate the human-machine team construct. The autonomous weapon acts as a means to a human designed end in some cases, whereas in other contexts the means becomes a new, emergent, and independently designed *end* in itself, beyond human contribution. Note that in each depiction of a human, that operator is assumed to be an unmodified, natural version augmented by the AI system.

Yet, the contemporary debates on how humans should employ autonomous weapon systems is just the latest evolution in human-machine teaming, where narrow AI is able to do precise activities faster and more effectively than the best human operator. Narrow AI applications in warfare illustrate the current frontier where existing technology is able to act with exceptional performance and destruction. However, no narrow AI system can match human operators in general intelligence contexts, which still compose the bulk of warfare contexts. For the next decade or two, human operators will continue to dominate decision making on battlefields yet to come, although increasingly the speed and dense technological soup of future wars will push humans into the back seat while AI drives in more situations then previously. It is the decades beyond those that will radically alter the human-machine warfare dynamic, potentially beyond any recognition.

The Event Horizon and Technological Revolutions: Breaking the Paradigm

This is potentially the last century in a massive string of centuries where humans are the primary decision makers and actors on battlefields. Figures 1 and 2 represent what will be a gradual shift from human operators being central to decision making (in the loop) to an ancillary status (on the loop) and subsequently to a reactive, even passive status (off the loop, behind the loop) as technological developments influence future battlefields to be unsafe for human decision speeds as well as the presence of human combatants.[29] Already, unmanned aviation, armored vehicles, and robots for a range of tactical security applications are in service or development that will replace more human operators with artificial ones that move faster, function in dangerous contexts, and are expendable with respect to the loss of human lives. While current ethical debates pursue where the human must remain in the kill chain for decisions of paramount importance, this assumes that the human still possesses superior judgment, intelligence, or other cognitive abilities that narrow AI systems cannot replace. If the coming decades bring forth advances in general AI to rival or exceed even the smartest human operators, those ethical concerns will be eclipsed by new ones.

Even the best human operator has theoretical biological, physical, and emotional limits that cannot be enhanced beyond a certain known limit.[30] Hypersonic weapons and swarm maneuvers of many AI machines pose a new threat, coupled with the increased speed of production and replacement through advances in 3D printers, cloud networks, constellations of smart machines, and more.[31] The natural-born human can be modified through genetics, cybernetic enhancement, and/or a human-machine teaming with AI systems to produce a better hybrid operator team.[32] For the coming decades, this likely will be the trend.[33] Yet, as figure 2 presents, the legacy human-machine teaming relationships framed in figure 1 will be replaced. How far and whether there are long-term ethical, moral, and legal consequences on modifying humans for military

applications is another area of concern in that much of the research is just starting or is still largely hypothetical.

There are two significant transformations that may render most of figure 1's human-machine teaming configurations obsolete. These concepts are hypothetical and likely many decades away, if even possible. The first is one where the natural born, unmodified human is insufficient to participate in future decision-action loops. They are outperformed by the theoretically enhanced human, whether this is achieved genetically, cybernetically, and/or through AI networking modification. Here, what could be called a *Supra sapien* outperforms any regular human opponent in every possible battlefield measurement. These enhanced humans would essentially break the contemporary human-machine teaming model in that their superintelligent abilities would be incompatible with how our militaries currently understand and frame decision-making relationships about how natural-born humans cope with battlefield contexts. This requires further elaboration.

Bostrom, in *Superintelligence: Paths, Dangers, Strategies* explains various paths to such a superintelligent, enhanced human that is vastly superior to even the most talented natural-born human specimen. Biological enhancement of human brains through genetic modification, biomedical enhancements, or hypothetical iterated embryo selection of select genotypes using stem cells to "accomplish ten or more generations of selection in just a few years" could produce humans with intelligence beyond traditional ways to measure such abilities, even dwarfing geniuses such as Isaac Newton and Albert Einstein.[34] A cybernetically enhanced human would have direct brain-computer interfaces that again hypothetically could create cognitive improvements whether the hardware is inserted into human tissue or linked to external systems that compliment or enhance the human *wetware* doing the thinking.[35] A networked AI enabled group of humans would work collectively, such as the fictional villain Borg collective from the *Star Trek: The Next Generation* television series. Technologically linked humans able to reduce bureaucratic drag, speed up the slowest individual human links in the chain, and permit AI data collection at vast scales could generate a collective superintelligence that no single natural-born human opponent could match.[36] In any of these hypothetical developments of current research in genetic, cybernetic, and network-enabled research, such a possibility could flip the entire notion of what a human-machine team is for warfare applications. There is one remaining human-connected hypothesis, where any of these possible enhanced human entities moves beyond what makes us all human and becomes something alien in a new intersection of advanced technology and original human desires.

The term *transhumanism* covers this overlap between technological advancement and the modification of human beings to break free of the natural, slow evolutionary process. Biology still governs what each generation of humans can do physically, although medical science and technology continue to change the boundaries as humans manipulate many more aspects of what was

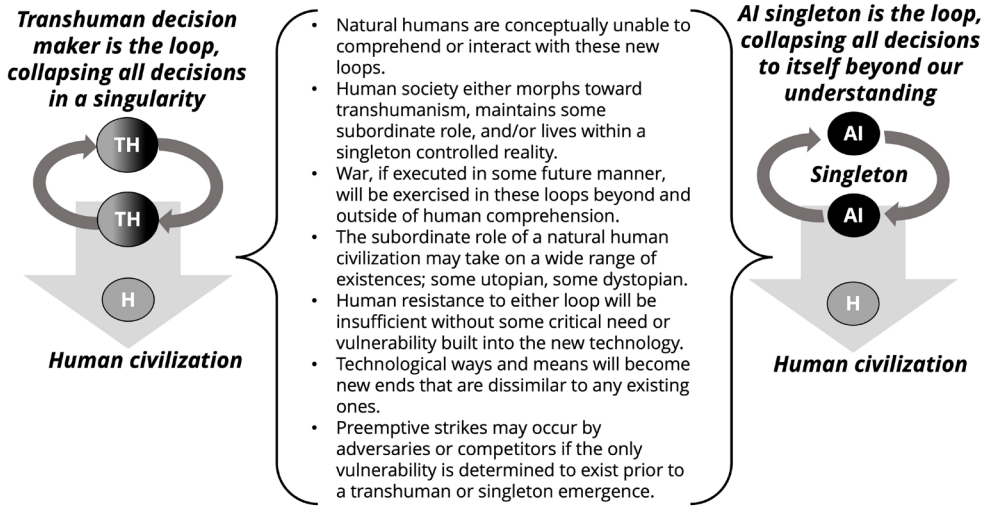
previously out of our hands. Yet, regardless of if a new baby is conceived in a test tube or the old-fashioned way, the output still is a human being that will develop within a society and think like other humans. With transhumanism, there is a divide between what started out as a human and what has now transcended what humanity can conceptualize or even recognize as human in form and function. Transhumanism need not be directly associated with the rise of superior AI, as the two might be better understood in a Venn diagram influencing one another.[37]

Technophiles suggest that at some advanced level, modified humans will reach a point where a singularity occurs, and suddenly these modified humans will exist on a new plane of reality that would arguably be inaccessible to natural, legacy human beings—as a transhuman entity could shed all organic habits and potentially become beings of pure information[38] This transhuman leap is illustrated on the left side of figure 3. The other extreme transition is depicted below on the right side and is where advanced AI reaches and then vastly exceeds general human intelligence. One comes from human stock; the other is created by humans but is artificial in design and expression. Alternatively termed strong AI, such superintelligent general AI remains entirely hypothetical but is anticipated by virtue of superiority to human cognition to be uncontrollable once developed by leading AI theorists today such as Ray Kurzweil.[39] Such a superior entity would potentially convince societies (by reason, force, or manipulation) that it should lead and decide for all matters of importance, including defense.

The concept of an AI singleton comes from Nick Bostrom, who explained that a singleton is an entity that becomes the single decision-making authority at the highest level of human organization. This usually means the entirety of human civilization, whether still limited to one planet or possibly spread across space in multiple colonies. Such an entity is considered "a set with only one member," where presumably a general AI entity that could rapidly advance from human-level intelligence beyond all biological, genetic, or otherwise non-AI limits into a superintelligence well past any human equivalent.[40] This brings into question whether the arrival of a singleton would directly dismantle human free will or the ability for humans to retain a status as a rational and biological species endowed with self-awareness.[41] The singleton entity, as a superintelligent AI entity, would sit above legacy humans in a new decision loop that would take control of human civilization just as a suggested transhuman manifestation might. Both concepts in figure 3 deserve greater explanation below. Note that with the transhuman entity, all of human civilization would remain under the loop, with the transhuman singularity blurring our loop model into where the transhuman entity becomes the entire loop. Respectively, the AI singleton, shedding all human attempts at control, also becomes the entire loop with human civilization placed under it.

Granted, a singleton could just be a fantastic individual or group that somehow manages to effectively take total control of human civilization (some future world government). However, in the entire history of human existence, this has

Figure 3. Further down the AI rabbit hole and 2050–90 warfare?



Source: courtesy of the author, adapted by MCUP.

yet to happen except in limited, isolated, and temporary conditions that fall short of the singleton constant. Bostrom highlights the totality of what a real singleton would be: "[a singleton's] defining characteristic . . . is some form of agency that can solve all major global coordination problems. It may, but need not, resemble any familiar form of human governance."[42] That no organic or natural singleton composed of one or several humans has yet in human history assumed any lasting form of a singleton indicates that for now, unmodified humans have thus far not produced any lasting or comprehensive (humanity-wide) singleton manifestations.[43] This may change with the rise of strong AI that can conceptualize well past existing cognitive limits and be able to decide and direct with successful outcomes systemically across the needs of an entire civilization. An entity that can outperform the best and brightest humans in every conceivable way, in any contest, at fantastic speeds and scale would seem either magical or godlike. Such concepts seem ridiculous, but many involved in AI research forecast these hypothetical developments as increasingly unavoidable and increasing exponentially over the next decades.[44] People may grow comfortable with their smartphones able to beat them at games of chess, poker, and pool, but will they agree to an AI that can outwit, deceive, or create and produce on every level (including on the battlefield) beyond their best efforts?

The singleton as depicted in the above figure offers the profound possibility that this entire shared socially constructed notion of war could be shattered and eclipsed by something beyond our reasoning and comprehension. Regular (narrow) AI may challenge both the supposed character and nature of future war, while a superintelligent singleton might break it completely.[45] All humans

would be under the loop in a singleton relationship where it assumes all essential decision making that governs and maintains the entire human civilization.[46] This may in turn change what we comprehend as both "human" and "civilization," or possibly in existential ways, how Homo sapiens remain a recognizable and surviving species.

The other component in figure 3 is the transhuman loop where a transhuman (or more than one) become the loop just as a singleton assumes total decision-making control. Legacy frames for what the decision loop was (figures 1 and 2) become irrelevant here. A transhuman entity extends a related concept of a singularity that overlaps with a singleton in some ways while differing in others. A singularity, first introduced by mathematician Vernor Vinge and made popular in science fiction culture by Ray Kurzweil, will break with the gradual continuum of human-technological progress with an entirely new stage in human existence.[47] It is considered a game changing, evolutionary moment where regular Homo sapiens would transform into a superintelligent, infinitely enhanced and possibly nonbiologically based technological fused entity.[48] Transhumanism envisions "our transcending biology or manipulating matter as a necessary part of the evolutionary process."[49] The arrival of a technological singularity coincides with a rapid departure of the transhuman entity away from the original biological evolutionary track.

A singularity introduces the concept of *transhumanism*, where at a biological, physical, political, sociological, and ultimately a philosophical level, humanity should break the slow evolutionary barriers and leap beyond the slow, clunky genetic and environmental soup of existence that changes organic life over thousands of years. Yuval Harari, in explaining how the cognitive revolution some 30,000–70,000 years ago, declared our species "independent from biology" where humans could radically alter the world around them and how they would conceptualize a socially complex reality atop the physical one so that the species did not rely on evolutionary biology to gradually develop improved instincts, physical developments, and other hardware or hardwired adaptations.[50] Thus, humans in the original cognitive revolution could develop complex ideas such as war and subsequently improve on the concept while waging it against fellow human adversaries.

Chimpanzees, in comparison, do engage in both predator-prey acts of localized violence as well as immediate and perhaps tactical acts of aggression for clear, immediate goals. Animals do not formulate strategies, nor produce languages, form religions, or develop political systems or laws, and they are entirely dependent on biology to give future offspring new advantages.[51] Humans create these incredible constructs by conceptualizing and subsequently manipulating reality, or the complex reality of the natural world with a second order of socially constructed complexity infused atop.[52] The singularity would theoretically create a second revolution in that the remaining biological, physical (both time and space), and sociological limitations would no longer exist for transhuman entities. They could rewrite their DNA; form entirely novel genetic combina-

tions; redesign their consciousness in ways that defy any rational expectations for human life; build entirely alien bodies that violate certain natural laws that confine even the most cunning, resourceful natural humans; and engage in warfare in unrealized, unimagined, and yet-to-be-understood configurations.

The singularity could do this in minutes or days instead, depending on the degree of modification or enhancement. Or as Vernor Vinge explains, "biology doesn't have legs. Eventually, intelligence running on nonbiological substrates will have more power."[53] This corresponds to raw cognitive powers of the evolutionarily honed human brain versus an entity that hypothetically could double its own abilities just as we might imagine in our minds what we fancy for dinner tonight. Bostrom offers a useful yet simplified summary with:

> The simplest example of speed superintelligence [which is but one of several hypothetical superintelligences Bostrom offers] would be a whole brain emulation running on fast hardware. An emulation operating at a speed of ten thousand times that of a biological brain would be able to read a book in a few seconds and write a PhD thesis in an afternoon. With a speedup factor of a million, an emulation could accomplish an entire millennium of intellectual work in one working day.[54]

Should a singularity or singleton manifest, ordinary humans would be unable to compete. In both circumstances in figure 3, unmodified, natural-born humans would remain below the decision loops, becoming wards of a transhuman *Supra sapien* protectorate or a singleton superintelligent artificial entity. Assuming of course that humanity would be kept in some sort of existence and contribute something to this new ordered reality, all decision loops for essential strategic or security affairs would become as figure 3 illustrates. Enhanced humans, whether genetically, cybernetically, or those that achieve a transhuman state in other ways, would assume the decision loop and advance it in terms of speed, scale, and scope beyond anything a regular human could understand or participate in.[55] This is where Musk's warning that human thought and communication would be so slow it would sound like elongated, simplistic whale songs to entities with superintelligent abilities. Unlike the hypersonic missile dynamic where the weapon can maneuver at extreme speeds (unlike traditional missiles) making it much harder for a human to respond and adjust to, a transhuman or singleton entity comprehends thought as well as achieves action beyond the limits of even the smartest, fastest human operator. Theoretically, systems processing at such high speeds would experience reality in a way that reinforces Musk's remark, as well as an earlier line from Commander Data, an android in the *Star Trek: The Next Generation* television show. While kissing his human girlfriend, she asked what he was thinking. He responded that he was reconfiguring the warp field parameters; analyzing the works of Charles Dickens; calculating the maximum amount of pressure he should apply to her lips; considering a new food supplement for his cat, Spot; and more. Jenna, his human date, was thinking about whether she could date an android.[56]

Technological development, as anticipated during the next century or less, may achieve either a singularity where human beings and their technological tools form a new transhuman entity, and, arguably, take concepts such as species and existence toward uncharted areas, or pure general AI may quickly pull itself into a level of existence beyond the comprehension of its human programmers. If either is a potential reality in the decades to come, how will war change? What might future battlefields become? What roles, if any, might human adversaries assume in such a transformed reality? Could humanity be doomed or potentially enslaved by technologically super-enabled entities that dehumanize societies of regular, natural humans? Could one nation unencumbered by ethical or legal barriers unleash superior AI abilities with devastating effects?[57] What emergent dangers lurk beyond the simplistic planning horizons where militaries contemplate narrow-AI swarms of drones for future defense needs in the 2030s and 2040s?

Whale Songs of Future Battlefields: The Irrelevance of Natural-Born Killers

For several centuries, modern militaries embraced a natural-science inspired ordering of reality where war is defined in Westphalian (nation-state centric), Clausewitzian (war is a continuation of politics), and what Der Derian artfully termed the "Bacion-Cartesian-Newtonian-mechanistic" model.[58] War has been interpreted along with a mechanical canonization of manufacturing, navigation, agriculture, medicine, and other "arts" in medieval and Renaissance thinking; Bacon's "patterns and principles," both early synonyms for rules, "emphasized that such arts [including military strategy] were worthy of the name."[59] Lorraine Daston goes on to explain:

> The specter of Fortuna haunted early modern treatises like Vauban's on how to wage war. In no other sphere of human activity is the risk of cataclysmic chaos greater; in no other sphere is the role of uncertainty and chance, for good or ill, more consequential. Yet many early modern treatises that attempted to reduce this or that disorderly practice to an art, none were more confident of their rules than those devoted to fortifications. This was in part because fortifications in the early modern period qualified as a branch of mixed mathematics (those branches of mathematics that "mixed" form with various kinds of matter, from light to cannonballs). Like mechanics or optics, it was heavily informed by geometry and, increasingly, by the rational mechanics of projectile motion.[60]

Modern military decision making remains tightly wedded to what is now several centuries' worth of tradition, ritualization, and indoctrination to framing war as well as the process of warfare into a hard-science inspired, systematic-logic derived construct where centers of gravity define strengths and vulnerabilities universally in all possible conflicts, just as principles of war such as mass,

speed, maneuver, objective, and simplicity are considered "the enduring bedrock of [U.S.] Army doctrine."[61] This comes from the transformational period where a feudal age military profession sought to modernize and embrace social, informational, and political change that accompanied significant technological advances.[62] Daston adds to this, explaining: "In the seventeenth and eighteenth centuries, the most universal and majestic of all laws were the laws of nature such as those formulated by the natural philosopher Isaac Newton in his *Philosophiae naturalis principia mathematica* in the late seventeenth century and the natural laws codified by jurists such as Hugo Grotius (1583–1645) and Samuel Pufendorf (1632–1694) in search of internationally valid norms for human conduct in an age of global expansion."[63] Natural sciences as professions would lead this movement, with medieval oriented militaries quickly falling into step.

The reason for this brief military history lesson is that today's modern military that currently integrates and develops artificial intelligence with human operators still holds to this natural science ordering of reality including warfare. War is framed through human understanding and nested in both a scientific (natural science) and political (Westphalian nation-state centric) framework. This in turn inspires nearly everything associated with modern war, including diplomacy, international rules and laws of war, principles of war, treaties, declaration of war, the treatment of noncombatants, neutrality, war crimes, and many other economic, social, informational, and technical considerations.[64]

Central to our shared understanding of war is the human decision maker and human operators that inflict acts of organized violence upon adversaries in precise, ordered, and what is ultimately a socially governed manner. Even when strong deviation occurs in war, such actions are comprehended, evaluated, and responded to within a human overarching framework. The human decision maker as well as all operators cognizant of any action in war are held responsible, such as in the Nuremberg Trials held against defeated Nazi Germany military representatives in 1945–46. Never before has this dynamic of human centeredness been challenged until now, where the role of artificial intelligence and humans on the battlefield are already entering shaky ground.

An autonomous weapon system, granted full decision-making abilities by human programmers, presents an ethical, moral, and legal dilemma on whether it, its programmers, or its human operators should be held responsible for something such as a war crime or tragic error during battle.[65] Current AI systems remain too narrow (in terms of AI), fragile, and limited in application to yet reach this level, for now at least. Increasingly powerful AI will in the coming decades replace human operators and, in some respects, even the decision makers. Humans, unless enhanced significantly, will become too slow and limited on battlefields where only augmented or artificial intelligence can move at the speeds, scale, and complexity necessary. Many of the natural laws of war could be broken or rendered irrelevant in these later and more ethically challenging areas of AI systems with general intelligence equal or beyond that of the human programmers.[66] For instance, a general AI in a singleton manifestation

would take over all decision making for any military conflict and potentially even exclude human operators from participating. Does war remain as it is now if humans are no longer part of it, despite humans socially creating war in the first place? These deeply troubling philosophical questions extend into religion, where a sentient AI with general intelligence that exceeds all human abilities could opt to join a human religion or design their own for AI. These developments could spell significant concern for both the human-centered and human-designed frames for war, religion, politics, culture, and more. The strategic abilities of nonhuman entities might exceed the comprehension and imagination of how humans for centuries have defined what war is.[67] The arrogance that human strategists several centuries ago figured out the true essence of war in some complete, unquestionable way is but one institutional barrier preventing any serious discussion on what is to become of natural born operators and decision makers on future battlefields. The natural science conceptualization of war is only a few centuries old and is already under challenge by postmodernists even in current contexts where AI plays a subordinate, highly controlled role. Future AI that would reimagine war would theoretically disrupt or replace existing human beliefs concerning war and warfare.

The last cavalry charge occurred at least one war too late to make any difference, while many technologically inferior societies encountered horrific losses attempting traditional war tactics and strategies against game-changing developments.[68] Arguably, with enough numbers, adversaries wielding significantly inferior weaponry can overcome a small force equipped even with game-changing technology, as the 25,000 Zulu warriors did defeat 1,800 British and colonial troops at the Battle of Isandlwana (22 January 1879). Yet, the Zulu offensive largely armed with iron-tipped spears and cow-hide shields lost several thousand warriors before eventually overwhelming their Martini-Henry breech loading rifle and 7-pounder mountain gun equipped opponents.[69] Nuclear weaponry may have shifted war toward intentionally limited engagements between nuclear-armed (or partnered) adversaries since the 1950s, but even this nuclear threshold may be in question. Yet, military institutions typically resist change and instead are often dragged, kicking and screaming, into the adaptation of new technology while they attempt to extend the relevance of those things that they identify with but are no longer relevant in battle.[70]

The last natural born, genetically unmodified, and noncybernetically enhanced human battlefield participants are not realized yet, nor is the future battlefield selected. However, whenever and wherever that happens, humanity may end up being tested in ways unlike anything previously. Or humans that reach such a technological level of accomplishment might grant total decision-making capability to a singleton or to enhanced humans that have become transhuman entities able to think and act in future warfare contexts beyond natural-born, whale-song sounding human opponents. In either case, it is unlikely the slower, unenhanced human opponents will be much of a challenge if indeed the AI or transhuman advantages are that significant.

Antoine Bousquet, in detailing the rise of cybernetics for military collection, processing, and decision making during the Cold War, correlated the increased speed of jet-powered nuclear bombers with a need for computer-assisted data analysis of incoming radar and observation post reports, as well as faster outgoing directives for antiaircraft defenses such as interceptor fighters, land-based weapons, and strategic updates to leadership on whether to employ a counterstrike.[71] The earliest computerized command, control, and communication network for this emergent military challenge launched in 1958 and was called SAGE (Semi-Automatic Ground Environment) that would provide real-time processing and respond to user inputs, all done over cathode ray tube technology. While SAGE was completed in 1963, it was already obsolete due to Soviet deployment of intercontinental ballistic missiles (ICBMs) that made antiaircraft defenses rather inconsequential.[72] In the high stakes, existential concerns of a potential nuclear war between the United States and the Soviet Union, human operators remained the ultimate decision makers even when coupled with these increasingly advanced computerized information processing systems.

Today, that dynamic remains largely unchanged, yet there is a growing creep of narrow focused AI systems taking more control and initiative to act without human supervision or interaction. These circumstances of AI-centric activities are localized to actions that an AI system has a low risk of complete malfunction, error, or other unexpected consequence of poor action, such as base defense of incoming rockets or emergency countermeasures for aircraft, vehicles, and submersibles. Artificial intelligence in general applications, once able to compete or exceed human abilities, may flip this dynamic, shifting humans to the role of the mine detector, where the human is on the loop or off the loop, moving too slowly and unable to conceptualize or act in a battlefield context where AI systems are swarming, networking, and engaging at speeds unreachable by the fastest human operator. Yet, there is today a fierce resistance to handing over significant decision making to machines in military culture.[73] Part of this deals with control and risk, while the way militaries maintain identity, belief systems, and values also factor into how AI technology is being developed.

This presents an interesting change in future war as presented by Der Derian. While he focuses on technology, information, and human perception therein, he defines a *virtuous war* as "the technical capability and ethical imperative to threaten and, if necessary, actualize violence from a distance—*with no or minimal casualties.*"[74] Der Derian frames the origin of virtuous war in the technological ramp-up and eventual Gulf War engagements between the United States and allies against Saddam Hussein's Iraqi forces that had invaded and occupied neighboring Kuwait. Stealth aviation, smart bomb precision, along with grainy video feeds of enemy targets being struck saturated the news cycles, along with offering the promise that future wars would be largely bloodless, with few civilian casualties and low risk to friendly forces using such game-changing tech-

nology. This is often termed *technical rationalism*, and as Alex Ryan observes of the modern military, "technical rationalism combines a naïve realist epistemology with instrumental reasoning."[75] Modern militaries apply engineering logic toward complex security challenges with a preference toward advanced technology as the optimized solution set for accomplishing warfare goals. Donald Schön elaborates on this mindset: "practitioners solve well-formed instrumental problems by applying theory and technique derived from systematic, preferably scientific knowledge."[76] The promise of a technologically rationalized future for warfare is not new. General William C. Westmoreland addressed the U.S. Congress during the Vietnam War about the future and technological promises:

> On the battlefield of the future, enemy forces will be located, tracked and targeted almost instantaneously through the use of data links, computer assisted intelligence evaluation, and automated fire control. With first round kill probabilities approaching certainty, and with surveillance devices that can continually track the enemy, the need for large forces to fix the opposition physically will be less important . . . I see battlefields or combat areas that are under 24 hour real or near time surveillance of all types. I see battlefields on which we can destroy anything we locate through instant communications and the almost instantaneous application of highly lethal firepower. I see a continuing need for highly mobile combat forces to assist in fixing and destroying the enemy. . . . Our problem now is to further our knowledge—exploit our technology, and equally important—to incorporate all these devices into an integrated land combat system.[77]

Virtuous wars have, according to Der Derian, closed the gap between an imagined or fantasized world of televised war and video game simulations with the gritty, brutal, and harsh reality of actual war. Der Derian explains that "new technologies of imitation and simulation as well as surveillance and speed have collapsed the geographical distance, chronological duration, the gap itself between the reality and virtuality of war."[78] Der Derian sees with this arrival of virtuous war the collapse of Clausewitzian war theory, the demise of the traditional sovereign state, "soon to be a relic for the museum of modernity . . . [or] has it virtually become the undead, haunting international politics like a spectre?"[79] Der Derian addresses human social construction of reality and whether the hyper-information, networked, and technologically saturated world of today is drifting toward a new era of struggling between the virtual and the real, the original and the copy, as well as the copy and the constructed illusion that has no original source at all.[80] To reapply Der Derian's construct toward this AI future transformation of war, will the removal of humans both from operations as well as decisions in warfare create the final exercise in virtuous war, perhaps the last gasp of humanity into what war has been previously?

The total removal of humans from battlefields presents many emergent dilemmas ranging from accountability, ethical as well as legal responsibilities, and

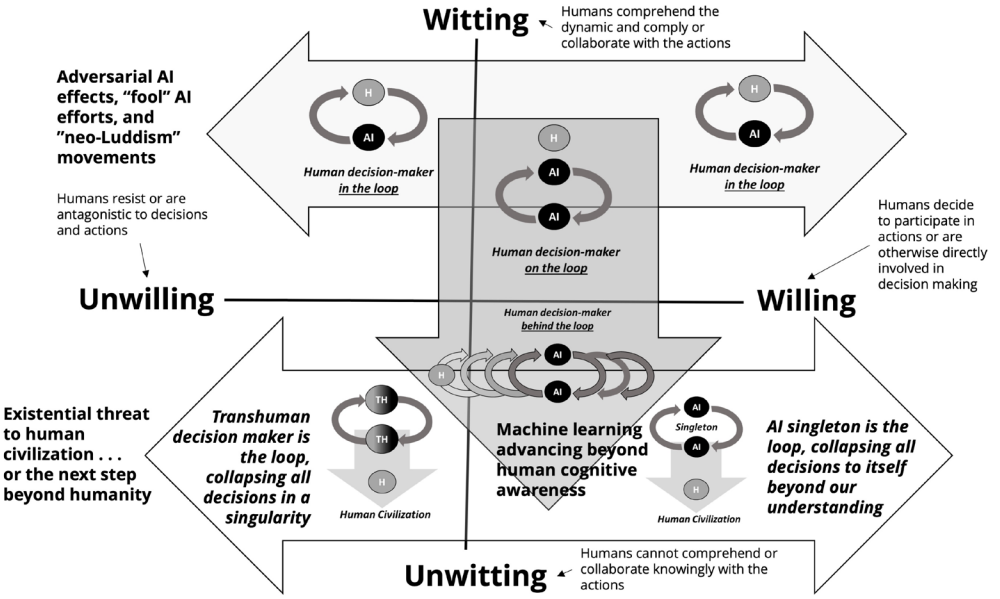
how intelligent machines can and will interact with human civilians on such future battlefields. Sidestepping the transhuman question for now, a purely artificial battlefield extending upward into all strategic and command authority decisions would move humans not just off or behind the loop, but under it. Would future wars even appear recognizable or be rationalized in original concepts? Might superintelligent machines be capable not just of winning future wars as defined by human programmers, but imagining and bringing into reality different forms and functions of warfare that are entirely unprecedented, unrealized, and unimagined?

The Calm Before the Storm: How It May Unfold

War has for thousands of years been a human design where as a species, Homo sapiens waged organized violence upon others of the same species in a manner unlike any other form of violence in nature. War is a human enterprise, until now. Advanced AI as well as the infusion of technology into how future humans exist will disrupt this history of violence. Genetically modified transhuman entities could gain unprecedented abilities in cognition, speed, strength, and resistance so that future battlefields would be far too dangerous for unmodified, original Homo sapiens. These super soldiers might have cyborg abilities or work in tandem with swarms of autonomous and semiautonomous war machines. However, in all of the enhanced human hypothetical paths beyond what exists today with gene modification, human-machine teaming with new tools like movement suits, armor, or surgical alteration with robotic implants, there is still a living human at risk on the future battlefield. Could there be a future where intelligent machines are so superior and lethal that no human being, regardless of enhancement, dare step foot upon that deadly landscape?[81]

The kinetic qualities of future war are more readily grasped, with science fiction already oversaturated with depictions of terminator robots and swarm armies of smart machines hunting down any inferior human opponent. Where advanced AI and transhuman developments are less clear is in the nonkinetic, informational, and social areas of warfare. In past wars, including the recent fall of Kabul in 2021 by advancing Taliban forces, information campaigns have been critical to gaining advantages over more powerful opponents. The Taliban invested for years in socially oriented, low-tech, grassroots influence campaign where they contacted Afghan security force personnel over phone calls, text messages, through social media, and by local, in-person means to gradually win over their ability to actively resist. The Taliban waged a sophisticated information campaign that would in the summer of 2021 collapse Afghan security resistance faster than ever anticipated by the high-tech, sophisticated Western security advisors planning the American withdraw. This momentous effort was done by Taliban operators in slow, largely person-to-person efforts through social engagements. Similar endeavors are increasing worldwide such as Chinese disinformation efforts targeting the Taiwanese population and a documented history of Russian disinformation activities against threats and rivals. Yet, most

Figure 4. Synthesis of witting, willing human actions



Source: courtesy of the author, adapted by MCUP.

bot activity is easy to identify and generates marginal impacts. Narrow AI remains brittle for now but changes are coming.[82]

Consider a transhuman or general AI entity that could find, engage, and convincingly correspond not with one person at a time, but millions? How fast could a human population be targeted, saturated with messaging, and engaged in a convincing manner that would be indistinguishable from human conversation despite the status of an AI entity (or one transhuman engaging with far more targets) being deceptive or manipulative for security aims? In the rapidly developing areas of genetics, nanotechnology, and viral and microbe technology applied to warfare, how might AI systems incorporate these into human programmed strategic and operational objectives? For human societies and civilization in general, might future wars with AI able to think and act at or above human performance levels produce both the most fantastic of advantages and, if one is the target of such power, the most devastating of threats?

In figure 4, a systemic treatment is presented using a quadrant that positions on the vertical axis those that are witting or unwitting with a horizontal axis spanning the willing and unwilling. Witting humans understand the dynamic relationship between themselves and artificial intelligent entities (or transhuman ones) that together produce a decision-action loop. Those that are unwitting humans are unable to comprehend nor fully collaborate within the decision-action loop. Unwitting human participants simply are unable to conceptualize what is actually happening, whether due to speed, intelligence, mul-

tiplicity, or other reason that the AI entity is deciding and acting beyond human limits. The horizontal axis offers the tension between humans willing to participate and maintain such relationships with AI entities (or transhuman ones) and those that will resist and oppose any movement toward such a dynamic.

Figure 4 provides a lateral drift illustrated in the horizontal arrow (spanning witting/unwilling and witting/willing) where willing human enterprise with AI entities "in the loop" coincides with productive warfare and security outcomes. If human-machine teaming produces better military results, more of the human participants should accelerate a willingness to continue and strengthen such efforts (toward the witting/willing side). Conversely, adversaries that also use human-machine teaming in conflict successfully will force opposing humans into the unwilling direction. Adversarial AI effects such as fooling rival human-machine teams will provoke further resistance (toward the witting/unwilling side). Outside the traditional battlefield dynamic of "us versus them," humans that perceive human-machine decision-action loops as a growing hazard or danger will assume some of the neo-Luddite positions and resist further investment in such technology. Impacts from war or conflict using AI systems may become beacons for technological activism to halt, prevent, or reverse such activities.

The downward pointing vertical arrow in figure 4 illustrates a progressive shift downward from the "witting-willing" quadrant to one of "willing-unwitting" where humans sit atop an increasingly sophisticated decision-action loop that has increasingly powerful AI. Over time, the human on the loop will eventually morph out of the witting into the unwitting, where that human is behind the loop of an ever-increasingly advanced AI system running all decisions with humans increasingly marginalized or excluded from the dynamic. This downward trend could bring with it an increased distrust, skepticism, or even conspiracy-theory fueled paranoia about advanced AI and their shifting interests in what is still human created, human directed warfare. The bottom horizontal arrow occupying the bottom two quadrants reflects the rise of superintelligent (in a general sense) entities that may be a transhuman extension beyond a singularity, or the rise of a pure AI singleton entity. In both instances, human civilization (unmodified, normal humans) would be subservient and in a protected (willing/unwitting) or perhaps oppressed status (unwilling/unwitting).[83] The bottom arrow spans both the willing and unwilling quadrants as unmodified human populations could embrace or go to some existential war against such superior entities.[84]

One final takeaway that figure 4 provides is how the twenty-first century could be plotted upon the systemic framing shown. The top arrow portion could be applied to 2030 through perhaps the 2050–70 period, depending on the speed of AI enhancement and development toward general intelligence.[85] The downward facing arrow might span the 2040–80 period to conservatively align with Bostrom's survey results for a 50–90 percent chance of human-level machine intelligence attainment, while the bottom arrow could be theoretically positioned in the 2075–2100 period, again based on Bostrom's survey

results. Figure 4 is largely hypothetical, but given the available research, trends for AI and processing developments, and the strategic forward thinking of experts looking to the future of AI, such a figure 4 may only have uncertainty of not if but *when* these trends do manifest in warfare. We could be a half-century off from the worst existential scenario of a singleton AI that decides to move against human creators, or we could be several centuries away instead. Either one remains existential and deserving of deep consideration by militaries today.

Modern warfare should remain approximate to contemporary understanding of organized violence and waged mostly by humans for the coming decades, although a gradual blending of humans and increasingly intelligent machines will become pronounced as the decades progress. While impossible to speculate when, the rise of a singularity or a singleton would spell the end of what has been more than 40 centuries of a human-defined, controlled, and developed war paradigm. What would happen next is unfortunately outside of our imagination. The most likely outcome should be that whatever humans currently believe is appropriate and rational for warfare will be insufficient, irrelevant, or inappropriate for what comes next.

Conclusions: The End Is Nigh . . . or Probably Not . . . but Possibly Worse

This article was developed as a thought piece oriented not on the near-term and immediate security concerns where new technology might make incremental impacts and opportunities. Rather, this long-term gaze addresses the emergent paths that exist beyond the direct focus areas of most policy makers, strategists, and military decision makers charged with defending national interests today. Humanity has over many centuries experienced a slow rate of change, accelerating exponentially in bursts where game-changing developments (fire, agriculture, writing, money-based economics, the computer) have ushered in profound change. Yet, within much of that change, warfare has been a deadly contest between human populations equipped with varying degrees of technology and resources. The weapons were the means to human-determined ends in conflict. New technology represented new means and increased opportunities for creative ways to inflict destruction on one's opponent.

The next shift with advanced artificial intelligence is already underway and will continue to unfold on the next few decades of battlefields with faster decision-action loops involving more sophisticated technology able to operate, organize, and influence at scales, speeds, and across multiple domains unlike in previous conflicts. This trend will gradually shift humans to atop the loop, and then in more contexts as risks are considered, behind the loop. Only the most dangerous, catastrophic decisions might remain exclusive to human decisions, until perhaps an adversary signals they have given it to a superior AI. Lastly, the loop may become a new end to itself, detached entirely from human creators. Arguably, cunning humans aware of this possibility could program devious means to prevent such a problem. Or enhanced humans with faster, stronger

conceptual abilities could continue to hold the reigns of the artificial intelligence decision cycle. This is possible, but Bostrom devotes an entire chapter to his book on what is "the control problem," and while he offers several ways to consider the programing, motivations, controls, kill switches, boxing methods, and more, he also acknowledges that "human beings are not secure systems, especially not when pitched against a superintelligence schemer and persuader."[86]

If the rise of advanced AI systems as well as new technological gains for human enhancement spell grave risks for humanity, might some lessons be found in organized resistance to nuclear arms and nuclear weaponized nations? In 1953, U.S. president Dwight D. Eisenhower created the Atoms for Peace program that attempted to demilitarize the American international image as the first nation to use atomic weapons in war as well as the leading nuclear weaponized nation actively conducting live nuclear tests at the time.[87] This program gained the support of many scientists, and while met with initial skepticism by Soviet leadership, the USSR would soften this stance and begin to negotiate and participate on the peaceful use of nuclear energy. In the 1960s and onward, multiple peace-oriented and antinuclear weapons groups, movements, and programs gained influence across the world. This in turn inspired many nations that could invest in nuclear weapons to defer and seek alternatives.

While nuclear weapons development is not a perfect match to how advanced AI development (including autonomous weaponization initiatives) may progress, such a resistance and activist movement could perhaps deter or contain the general AI development that could lead to a singleton entity or postpone a dangerous singleton arms race toward accomplishing the first one over adversaries.[88] Unlike nuclear weapons that are a means toward particular ends in foreign policy and defense, the singleton as well as any singularity that produces transhuman entities with superintelligent abilities may quickly become their own ends in themselves. In such stark possibilities, a neo-Luddism movement, assuming one or more exist during the development of such AI and human enhancement, likely will be entrenched to form some resistance. Even if resistance occurred, the disadvantages of that group would be exacerbated by an assumed takeover of national security apparatuses by a superintelligent transhuman entity or AI singleton at the request or unwitting agreement of those designing the technology.

Unlike the English Luddites of the nineteenth century that they are named for, neo-Luddites are a decentralized, leaderless movement of nonaffiliated groups and individuals that propose the rejection of select technology, particularly those that pose tremendous environmental threat and any significant departure from a simplistic, natural state of existence. Neo-Luddites take a similar philosophical stance as antiwar and antinuclear groups, where the elimination of harmful technology offers salvation for humanity as one species within the broader ecosystem of planet Earth. Mathematician and conflict theorist Anatol Rapoport termed such movements "global cataclysmic" where all war is harmful and the prevention of any war is a necessary goal for all of civilization to

pursue.[89] While Rapoport crafted his concept to frame mid-late Soviet conflict philosophy through this self-preservation of Marxist society, "global cataclysmic" could be applied also to international entities such as the United Nations concerning conflict, and for environmental efforts as a way to explain the radical positions of the ecoterror group Earth Liberation Front in the 1990s in the American Pacific Northwest. Unlike a Westphalian or Clausewitzian war philosophy that permits state-on-state and other state-directed acts of policy and war, the global cataclysmic philosophy rejects all wars as dangerous to society. Neo-Luddites would swap "war" with "dangerous technology" and foresee human extinction or a planet-wide disaster as a direct, foreseeable outcome of human tinkering with technology that could destroy the world as we know it. The term Luddite is also misapplied when some of the most prominent AI developers such as Elon Musk, Bill Gates, Alan Turing, and others are lumped into this group because they call public attention to the concerns of AI and are raising them for entirely different purposes than what neo-Luddites seek.[90]

Neo-Luddism as well as potentially more radical and violent groups could posit a global cataclysmic philosophy against advanced technology that would unavoidably include AI systems. The real possibility of a technological shift into a singularity or the arrival of a singleton entity capable of gaining total control of all defense systems presents existential concerns that correlate to the deepest aims of these movements. Existentially, the potential subjugation of regular human society under rule (protection or enslavement) of a transhuman or singleton AI may trigger radical actions by some of these groups, particularly if technological developments accelerate in publicly understood narratives. This presents additional security considerations for all nations that have advanced artificial intelligence research as well as efforts to integrate such developments into military forces and across civilian society. Those that advocate for transhumanism and a singularity will be in fierce opposition with those proposing neo-Luddite perspectives with the middle ground increasingly scarce in such debates. Existential debate rarely fosters calm and collected discussions.

The widespread acts of economic sabotage, arson, and guerrilla warfare that defined the Earth Liberation Front's most active period in the late 1990s and early 2000s caused the Federal Bureau of Investigation to declare the group the top domestic terror threat in February 2001.[91] The group is decentralized with no formal leadership or hierarchical structure, making it part of a pattern of rhizomic (no central form; cannot be isolated or reduced to impact the larger system) organizations that mitigate or even defeat the most powerful instruments of national power for traditional Westphalia-styled nation states. If there is to be active resistance toward perceived threats of transhumanism, singularities, and singletons through continued technological advancement in the coming decades, it likely will manifest in one of the three aforementioned forms. Likely, all three will develop simultaneously.

Nation-states will themselves pursue diplomatic and international norms, policies, and treaties concerning efforts to contain the dangers of advanced AI,

technological enhancement of Homo sapiens toward some transhuman singularity that might be weaponized toward others, as well as specific concerns of narrow and general AI in warfare. There may be state sponsored as well as grassroots and decentralized movements to curb these technological developments, or there might be efforts to contain them using the scientific community, commercial enterprise, or social activism. However, efforts to ban technological development has a poor track record, and such attempts may only push technology experimentation underground or into nations that have no such objections, where it could be even more dangerous.[92]

Some groups that posit particular philosophical stances against technology will gain influence, especially if environmental and antiwar groups are successful in highlighting existential threats in these potential technological developments. Lastly, splinter groups and violent extremist movements have increasingly grown in numbers, impact, and frequency with the rise of the globally connected, social media infused information age of modern society. That these groups might strike to attempt to halt (or in some cases, possibly accelerate) the arrival of transhumanism, the singularity or a singleton is a growing security concern for all. Hugo de Garis is one of the futurists that envisions a brutal such development, where "the rise of AI will lead to war. . . . He does not mind that a pitched war may lead to the destruction of the human race because he believes that 'godlike' machines will survive afterward."[93] These extreme positions for and against technological progress and AI present significant existential narratives that might instigate violent acts from either side of the divide.

Humanity today is at the edge of what could be transformative, liberating, destructive, or eternally enslaving. Homo sapiens became the deadly marvel of all organic life on this planet through how they could manipulate reality and interplay between the conceptual in their minds with the real world at their fingertips. Harari expands on how different humans are as a species from the rest of the world in their ability to conceptualize, form language, and communicate with abstractions in their minds so that reality can be manipulated in unprecedented ways:

> It's relatively easy to agree that only *Homo sapiens* can speak about things that don't really exist, and believe six impossible things before breakfast. You could never convince a monkey to give you a banana by promising him limitless bananas after death in monkey heaven. But why is it important? After all, fiction can be dangerously misleading or distracting. . . . However, fiction has enabled us not merely to imagine things, but to do so *collectively.* We can weave common myths such as the biblical creation story, the Dreamtime myths of Aboriginal Australians, and the nationalist myths of modern states. Such myths give Sapiens the unprecedented ability to cooperate flexibly in large numbers. Ants and bees can also work together in huge numbers, but they do so in a very rigid manner and only with close relatives. . . . That's why

> Sapiens rule the world, whereas ants eat our leftovers and chimps are locked up in zoos and research laboratories.[94]

Yet, now those minds and those fingers might produce outcomes that outpace the ability to manage, control, and shape the future for not just those new things, but for everything. We are ill-equipped to think in exponential terms, preferring a linear-causal explanation of where tomorrow is going based on our collected analysis of yesterdays.[95] If AI development moves exponentially in the coming decades, we may be like the cavemen unable to see beyond the first tools of organized violence and that the spears and axes of the first battles would morph into submarines, tanks, and stealth bombers. This is a built-in problem for organic humans and natural evolutionary processes, but likely not an issue for AI entities we may create.

Humans may create the perfect future world where every possible need is met and virtually all risks and harms are reduced or eliminated for everyone. They might eliminate hunger, war, and misery, or they could unleash devastation unlike anything seen before. The question remains, however, about whether humans that create advanced entities (transhuman or singleton) will be a willing and witting participant in these new realities. There are moral, ethical, and religious debates on whether a transhuman technological manifestation will enable a superior human or instead degrade the human original, spawning false copies that lack the original uniqueness and humanity of the natural-born variant.[96] The same can be argued for a superintelligent AI that reaches a singleton status and dominates a civilization of lesser ability human creators. Will it share any of the original humanness that is essential for protecting and nurturing society beyond the current fragile state of affairs? Or will such developments plunge civilization into extinction, devastating wars, or some sort of organic servitude to new masters? Theoretical physicist and AI researcher Max Tegmark suggests this possibility:

> Perhaps artificial superintelligence will enable life to spread throughout the cosmos and flourish for billions or trillions of years, and perhaps this will be because of decisions we make here, on our planet, in our lifetime. . . . Or humanity may soon go extinct, through some self-inflicted calamity caused by the power of our technology growing faster than the wisdom with which we manage it.[97]

There may be the promise of either of these two vastly different futures, as well as potential wars and devastation waged before either might be accomplished, waged either to prevent or encourage the arrival of one possible future or the other. More perplexing than all of this, the clever Homo sapiens may never really understand or grasp what might emerge, as the potential of superintelligence paired with future security challenges could be outside human limits of comprehension. In such a future battlefield, the only sounds heard might just be the fading chorus of whale songs slowly wheezing away into ignored noth-

ingness. Giampiero Giacomello's similar warning reinforces this paramount crossroad we are approaching: "All in all, maintaining human in the loop, let alone in the loop, may turn into a strategic (and deadly) disadvantage, or be strategically illogical. In the end, the loop is indeed too small for the both of them."[98] This deadly dance of nature and the artificial may end in myriad tragic ways, including several where even the human winner ends up losing what makes them decidedly human.

Endnotes

1. Ben Gilbert, "Elon Musk Says Humans Communicate So Slowly with Computers That It Will Sound Like Whale Speech to Future AI," *Business Insider*, 29 August 2019, 8.
2. James Der Derian, "Virtuous War/Virtual Theory," *International Affairs (Royal Institute of International Affairs 1944–)* 76, no. 4 (October 2000): 786, https://doi.org/10.1111/1468-2346.00164; Christopher Paparone, *The Sociology of Military Science: Prospects for Postinstitutional Military Design* (New York: Bloomsbury Academic Publishing, 2013), 20; Shimon Naveh, Jim Schneider, and Timothy Challans, *The Structure of Operational Revolution: A Prolegomena* (Fort Leavenworth, KS: Booz Allen Hamilton, 2009), 35–36; and Anatol Rapoport, *The Origins of Violence: Approaches to the Study of Conflict* (New Brunswick, NJ: Transactions Publishers, 1995), 53–72.
3. Gibson Burrell and Gareth Morgan, *Sociological Paradigms and Organisational Analysis: Elements of the Sociology of Corporate Life* (Portsmouth, NH: Heinemann, 1979); Dennis A. Gioia and Evelyn Pitre, "Multiparadigm Perspectives on Theory Building," *Academy of Management Review* 15, no. 4 (1990): 584–602, https://doi.org/10.2307/258683; Marianne W. Lewis and Andrew I. Grimes, "Metatriangulation: Building Theory from Multiple Paradigms," *Academy of Management Review* 24, no. 4 (1999): 672–90, https://doi.org/10.5465/amr.1999.2553247; and Paparone, *The Sociology of Military Science*, 73–80.
4. Robert Chia, "Reflections: In Praise of Silent Transformation—Allowing Change Through 'Letting Happen'," *Journal of Change Management* 14, no. 1 (2013): 13, https://doi.org/10.1080/14697017.2013.841006.
5. Frank Wilczek, "The Unity of Intelligence," in *Possible Minds: Twenty-Five Ways of Looking at AI*, ed. John Brockman (New York: Penguin Press, 2019), 74–75.
6. Jaan Tallinn, "Dissident Messages," in *Possible Minds*, 97.
7. Der Derian, "Virtuous War/Virtual Theory," 788. Der Derian coins the concepts MIME and MIME-NET for the combination of media entertainment and the military industrial complex that he sees as fused together today.
8. Yuval Harari, *Sapiens: A Brief History of Humankind* (New York: Harper Prennial, 2018), 38. Harari clarifies: "This does not mean that *Homo sapiens* and human culture became exempt from biological laws. We are still animals, and our physical, emotional and cognitive abilities are still shaped by our DNA." Harari credits the development of language and socialization well beyond earlier biological limits of similar species such as Chimpanzees as the critical difference. Harari, *Sapiens*, 21–39.
9. David C. Lindberg, *The Beginnings of Western Science: The European Scientific Tradition in Philosophical, Religious, and Institutional Context, 600 B.C. to A.D. 1450* (Chicago: University of Chicago Press, 1992), 5.
10. Predator-prey relationships already existed in nature before humans, as did individual acts of violence for a wide range of reasons and goals. Yet, only humans, once able to communicate and organize into recognizable groups, tribes, and societies would become capable of waging war.
11. Julian Jaynes, *The Origin of Consciousness in the Breakdown of the Bicameral Mind*, 3d ed. (Boston, MA: First Mariner Books, 2000), 68–69.
12. Harari, *Sapiens*, 244–59.

13. Antoine Bousquet, *The Scientific Way of Warfare: Order and Chaos on the Battlefields of Modernity* (London: Hurst Publishers, 2009).
14. Michal Krelina, *Quantum Warfare: Definitions, Overview and Challenges* (New York: Cornell University, 2021).
15. Lorraine Daston, *Rules: A Short History of What We Live By* (Princeton, NJ: Princeton University Press, 2022), 122–50.
16. Peter Layton, *Fighting Artificial Intelligence Battles: Operational Concepts for Future AI-Enabled Wars* (Canberra: Australian Defence College, Centre for Defence Research, 2021), 3.
17. Ben Goertzel, "The Singularity Is Coming," *Issues*, no. 98 (March 2012): 6; and Jacob Shatzer, "Fake and Future 'Humans': Artificial Intelligence, Transhumanism, and the Question of the Person," *Southwestern Journal of Theology* 63, no. 2 (Spring 2021): 130.
18. Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford, UK: Oxford University Press, 2016); and Shatzer, "Fake and Future 'Humans'," 131.
19. Goertzel, "The Singularity Is Coming," 4.
20. Layton, "Fighting Artificial Intelligence Battles," 9.
21. Shatzer, "Fake and Future 'Humans'," 133; Paul Scharre, *Army of None: Autonomous Weapons and the Future of War* (New York: W. W. Norton, 2018), 83; and Judea Pearl, "The Limitations of Opaque Learning Machines," in *Possible Minds*, 15–19.
22. Layton, "Fighting Artificial Intelligence Battles," 9.
23. Adam Cutler was invited to lecture on AI and defense at U.S. Space Command's Commander's Conference at Peterson Space Force Base, CO, on 5 May 2022. During Cutler's lunchtime keynote lecture, he provided this example when asked in the Q&A portion on what keeps him up at night.
24. Giampiero Giacomello, "The War of 'Intelligent' Machines May Be Inevitable," *Peace Review: A Journal of Social Justice* 33, no. 2 (2021): 282–83, https://doi.org/10.1080/10402659.2021.1998860; and Brian T. Molloy, "Project Governance for Defense Applications of Artificial Intelligence: An Ethics-Based Approach," *PRISM* 9, no. 3 (2021): 108.
25. Scharre, *Army of None*, 42–52.
26. Aron Dombrovszki, "The Unfounded Bias Against Autonomous Weapons Systems," *Informacios Tarsadalom* 21, no. 2 (2021): 15.
27. Neal Schaffer, "The Top 21 Twitter User Statistics for 2022 to Guide Your Marketing Strategy," nealschaffer.com, 22 May 2022; and Salman Aslam, "Twitter by the Numbers: Stats, Demographics & Fun Facts," *Twitter by the Numbers: Stats, Demographics & Fun Facts* (blog), 22 February 2022.
28. Robert Chia, "Teaching Paradigm Shifting in Management Education: University Business Schools and the Entrepreneurial Imagination," *Journal of Management Studies* 33, no. 4 (July 1996): 409–28, https://doi.org/10.1111/j.1467-6486.1996.tb00162.x; Colin Clarke-Hill, Huaning Li, and Barry Davies, "The Paradox of Co-Operation and Competition in Strategic Alliances: Towards a Multi-Paradigm Approach," *Management Research News* 26, no. 1 (2003): 1–20; Charles C. Moskos, "Towards a Postmodern Military: The United States as a Paradigm," in *The Postmodern Military: Armed Forces after the Cold War*, ed. Charles C. Moskos, John Allen Williams, and David R. Segal (New York: Oxford University Press, 2000), 14–31; Majken Schultz and Mary Jo Hatch, "Living with Multiple Paradigms: The Case of Paradigm Interplay in Organizational Culture Studies," *Academy of Management Review* 21, no. 2 (1996): 529–57, https://doi.org/10.5465/amr.1996.9605060221; and Ben Zweibelson, "Professional Reading Lists: Thinking Beyond the Books and into Military Paradigmatic Biases," *Air and Space Power Journal* 30, no. 2 (Summer 2016): 15–37.
29. Scharre, *Army of None*, 24–25.
30. Bostrom, *Superintelligence*, 26–61.
31. James Johnson, "Artificial Intelligence, Drone Swarming and Escalation Risks in Future Warfare," *RUSI Journal* 165, no. 2 (2020): 26–36, https://doi.org/10.1080/03071847.2020.1752026; Anastasios Giannakis, Michael Shoesmith, and Madison Thomas, "First Artificial Intelligence Swarm Drones Recently Utilized on the Battle-

field: Implications," Counterterrorism Group, 17 August 2021; Alessandro Gagaridis, "Warfare Evolved: Drone Swarms," Geopolitical Monitor, 10 June 2022; and Ben Zweibelson, "Let Me Tell You About the Birds and the Bees: Swarm Theory and Military Decision-Making," *Canadian Military Journal* 15, no. 3 (Summer 2015): 29–36.

32. Robert M. Geraci, "Apocalyptic AI: Religion and the Promise of Artificial Intelligence," *Journal of the American Academy of Religion* 76, no. 1 (March 2008): 154–56, https://doi.org/10.1093/jaarel/lfm101.
33. Wilczek, "The Unity of Intelligence," 74.
34. Bostrom, *Superintelligence*, 47.
35. Bostrom, *Superintelligence*, 54–58.
36. Bostrom, *Superintelligence*, 58–60.
37. Shatzer, "Fake and Future 'Humans'," 133–34.
38. Kevin Shapiro, "This Is Your Brain on Nanobots," *Observations*, December 2005, 64–65.
39. Ray Kurzweil, "Merging with the Machines: Information Technology, Artificial Intelligence, and the Law of Exponential Growth, Part 2," *World Future Review* 2, no. 2 (May 2010): 59, https://doi.org/10.1177/194675671000200209.
40. Nick Bostrom, "What Is a Singleton?," *Linguistic and Philosophical Investigations* 5, no. 2 (2006); and Bostrom, *Superintelligence*, 26.
41. Aura Elena Schussler, "Artificial Intelligence and Mind-Reading Machines—Towards a Future Techno-Panoptic Singularity," *Postmodern Openings* 11, no. 4 (2020): 342, https://doi.org/10.18662/po/11.4/239.
42. Bostrom, *Superintelligence*.
43. Historians could argue that in certain societal examples, such as a particularly strong monarch, dictator, or committee ruling over large swaths of the planet's available population, momentary singleton phenomenon briefly flourished. Genghis Khan, Alexander the Great, Ivan the Terrible, and Pachacuti Inca Yupanqui did wield absolute power over enormous populations and geographies, yet such power and control remained fleeting and often dependent on the individual. Pol Pot, Fidel Castro, and Kim Jong-un show similar near-total power in smaller and geographically isolated examples. Collectives such as the Soviet Politburo, Puritan colonial settlements in North America, and other group collaborations also offer interesting yet temporary singleton possibilities.
44. Ray Kurzweil, "Merging with the Machines: Information Technology, Artificial Intelligence, and the Law of Exponential Growth, Part 1," *World Future Review* 2, no. 1 (March 2010): 61–66, https://doi.org/10.1177/194675671000200107; Kurzweil, "Merging With the Machines, Part 2"; Bostrom, "What Is a Singleton?"; and Goertzel, "The Singularity Is Coming."
45. Devoted believers in a Clausewitzian philosophical framing of war as the single and unquestionable truth will no doubt protest such a statement. However, numerous scholars from a wide range of security studies disciplines disrupt, critique, or displace this perspective as anything but one of numerous others. See Naveh, Schneider, and Challans, *The Structure of Operational Revolution*, 106–20; Anatol Rapoport, "Editor's Introduction to *On War*," in *On War*, by Carl Von Clausewitz, ed. Anatol Rapoport (New York: Penguin Books, 1968); Chia, "Reflections: In Praise of Silent Transformation," 14; Chris Habbes Gray, *Postmodern War: The New Politics of Conflict* (New York: Guilford Press, 1997), 22; Paparone, *The Sociology of Military Science*, 29–33; Gilles Deleuze and Felix Guattari, *A Thousand Plateaus*, trans. Brian Massumi (Minneapolis: University of Minnesota Press, 1987); Jean Baudrillard, *The Gulf War Did Not Take Place*, trans. Paul Patton (Sydney, Australia: Power Publications, 2009); Astrid H. M. Nordin and Dan Oberg, "Targeting the Ontology of War: From Clausewitz to Baudrillard," *Millennium: Journal of International Studies* 43, no. 2 (2015): 392–410; Paul Patton, "Introduction to 'the Gulf War Did Not Take Place'," in Baudrillard, *The Gulf War Did Not Take Place*, 1–22; and Christopher Paparone and William Davis Jr., "Exploring Outside the Tropics of Clausewitz: Our Slavish Anchoring to an Archaic Metaphor," in *Addressing the Fog of COG: Perspectives on the Center of Gravity in US*

Military Doctrine, ed. Celestino Perez (Fort Leavenworth, KS: Combat Studies Institute Press, 2012).

46. Benjamin M. Jensen, Christopher Whyte, and Scott Cuomo, "Algorithms at War: The Promise, Peril, and Limits of Artificial Intelligence," *International Studies Review* 22, no. 3 (September 2020): 529, https://doi.org/10.1093/isr/viz025; and Denise Garcia, "Lethal Artificial Intelligence and Change: The Future of International Peace and Security," *International Studies Review* 20, no. 2 (June 2018): 334, https://doi.org/10.1093/isr/viy029.
47. Maxim Shadurski, "The Singularity and H. G. Wells's Conception of the World Brain," *Brno Studies in English* 46, no. 1 (2020): 229, https://doi.org/10.5817/BSE2020-1-11.
48. Justin K. Pugh, Lisa B. Soros, and Kenneth O. Stanley, "Quality Diversity: A New Frontier for Evolutionary Computation," *Frontiers in Robotics and AI* 3, no. 40 (July 2016): 1–3; Shapiro, "This Is Your Brain on Nanobots," 64–65; and Shadurski, "The Singularity and H. G. Wells's Conception of the World Brain," 229.
49. Pugh, Soros, and Stanley, "Quality Diversity," 5.
50. Harari, *Sapiens*, 21, 37.
51. Harari, *Sapiens*, 3–39.
52. Schultz and Hatch, "Living with Multiple Paradigms"; Haridimos Tsoukas and Mary Jo Hatch, "Complex Thinking, Complex Practice: The Case for a Narrative Approach to Organizational Complexity," *Human Relations* 54, no. 8 (August 2001): 979–1013, https://doi.org/10.1177/0018726701548001; Mary Jo Hatch and Dvora Yanow, "Methodology by Metaphor: Ways of Seeing in Painting and Research," *Organization Studies* 29, no. 1 (2008): 23–44, https://doi.org/10.1177/0170840607086635; and George Ritzer, *Sociology: A Multiple Paradigm Science* (Boston, MA: Allyn and Bacon Publishing, 1980).
53. Vernor Vinge and Jim Euchner, "Science Fiction as Foresight: An Interview with Vernor Vinge," *Research-Technology Management* 60, no. 1 (February 2017): 13, https://doi.org/10.1080/08956308.2017.1255048.
54. Bostrom, *Superintelligence*, 64. Bostrom includes two footnotes in this passage where he later expands on discussing the physical limits of the speed of light, energy, and theoretical human limits for processing powers. He also expands on whether a human emulated mind running at a million times the average human ability might go mad or fall into a psychological rut that breaks the limits of any human entity, even a superintelligent one.
55. Enhanced is used here in a hypothetical, exceptional context. Many humans today qualify as enhanced through existing pharmaceutical, surgical, technologically aided, and even specially conditioned ways. Enhancement that brings a human to the upper limits of reaching a transhuman transformation suggests what only exists in science fiction and fantasy today. Yet, similar content in the nineteenth century would eerily predict technological transformations of war in the twentieth century. H. G. Wells's *The War of the Worlds* details rocket travel that would be realized in World War II, and total war superweapons such as Heat-Ray and poisonous Black Smoke would manifest in chemical and later atomic warfare. Wells's Heat-Ray suggests mobile laser systems that now are in orbit, while the armored "Land Ironclads" anticipated armored maneuver warfare. Today's science fiction often offers glimpses into what may occur tomorrow.
56. A summary of this scene is available on this *Star Trek* fan website: https://www3.nd.edu/~ljordan/data.text/love.html.
57. Arthur Herman, "Why China Is Winning the War for High Tech," *National Review*, 14 October 2021, 32–34; and Chris C. Demchak, "China: Determined to Dominate Cyberspace and AI," *Bulletin of the Atomic Scientists* 75, no. 3 (2019): 99–101, https://doi.org/10.1080/00963402.2019.1604857.
58. Der Derian, "Virtuous War/Virtual Theory," 786.
59. Daston, *Rules*, 76–77.
60. Daston, *Rules*, 63.
61. *Operations*, Field Manual 3-0 (Washington, DC: Department of the Army, 2001), 4–12.

62. Aaron P. Jackson, *The Roots of Military Doctrine: Change and Continuity in Understanding the Practice of Warfare* (Fort Leavenworth, KS: Combat Studies Institute Press, 2013); Gray, *Postmodern War*, 23; and Barry Posen, *The Sources of Military Doctrine: France, Britain, and Germany Between the World Wars* (Ithaca, NY: Cornell University Press, 1984), 52; Siniša Malešević, "The Organization of Military Violence in the 21st Century," *Organization* 24, no. 4 (2017): 456–74, https://doi.org/10.1177/1350508417693854; and Felix Gilbert, "Machiavelli: The Renaissance of the Art of War," in *Makers of Modern Strategy: From Machiavelli to the Nuclear Age*, ed. Peter Paret (Princeton, NJ: Princeton University Press, 1986), 11.
63. Daston, *Rules*, 151.
64. Somewhat confusingly, international humanitarian law declares the laws of war to correspond with the conduct of warfare, while Antoine-Henri de Jomini-inspired principles of war are universal, natural science derived, and formulaic tenets that exist in any and every war. The principles of war address proposed foundational principles governing war in a natural, timeless state, while the laws of war clarify the legal boundaries of warfare activities set in modernity, corresponding to contemporary shared values that are not applicable to past wars as the Jominian principles of war are. Thus, the laws of war are not universal, in that future international decisions could change the laws of war due to emerging reasons and new logics.
65. Noreen Herzfeld, "Can Lethal Autonomous Weapons Be Just?," *Journal of Moral Theology* 11, Special Issue 1 (2022): 70–86, https://doi.org/10.55476/001c.34124.
66. Again, war does not exist outside of the human species, particularly if one links politics, law, religion, and economic theory to reasons war is waged; no animals demonstrate such socially constructed phenomenon, nor do they possess expressive language to narrate between the tangible and abstract. While an unchanging nature clearly existed well before humans in physics, chemistry, geology, biology, and other natural sciences, only the modern military holds to a near ritualized insistence that war itself gains natural-science characteristics, or suggests that it is timeless, universal, unchanging, and perpetual for all existence. Extensive footnotes seem necessary for challenging these beliefs, but paradoxically are rarely present when arguments reinforce an unchanging nature, changing characteristics of war in mainstream military academia.
67. As stated in an earlier footnote, opponents of the idea that war could be anything but the mainstream, Westphalian oriented, natural-science inspired, and styled in an engineering, systematic approach will disagree with such a position, even if nonhuman entities are hypothetically introduced. There are many interesting positions critical and outside of this frame. Lucas considers a changing war due to sociological and technological developments, while Gorka argues that the Westphalian monopoly on war is weakening. Rapoport frames his introduction to *On War* with an extensive breakdown of multiple war frames outside of the Clausewitzian perspective, while Paparone provides extensive synthesis of sociological and postmodern concepts to outline entirely different ways that a war philosophy can depart entirely from a natural-science inspired one. See George R. Lucas Jr., "Postmodern War," *Journal of Military Ethics* 9, no. 4 (2010): 290, https://doi.org/10.1080/15027570.2010.536399; Sebastian Gorka, "Adapting to Today's Battlefield: The Islamic State and Irregular War as the 'New Normal'," in *Beyond Convergence: World Without Order*, ed. Hilary Matfess and Michael Miklaucic (Washington, DC: National Defense University, 2016), 354; Rapoport, "Editor's Introduction to *On War*"; and Paparone, *The Sociology of Military Science*, 114–40.
68. This does not discount the strong historic pattern of low-technology, low-resource opponents able to overcome and defeat high-technology, well-resourced adversaries. The successes of the Viet Cong, Taliban, Barbary pirates, and decentralized groups such as the Earth Liberation Front against superpower status nations should not be overlooked.
69. Edward Langer, "Circumstances at the Battle of Isandlwana Fatally Reduced the Martini-Henry Rifle's Effectiveness," *Military History* 21, no. 6 (February 2005): 72, 78; and Christine Keating, "Never Such a Disaster: An Analysis of the British Defeat at Isandlwana," *Military Police* 9, no. 2 (Fall 2009): 38–41.

70. Jensen, Whyte, and Cuomo, "Algorithms at War," 536.
71. Antoine Bousquet, "Cyberneticizing the American War Machine: Science and Computers in the Cold War," *Cold War History* 8, no. 1 (2008): 85–87, https://doi.org/10.1080/14682740701791359.
72. Bousquet, "Cyberneticizing the American War Machine," 86.
73. Scharre, *Army of None*, 61.
74. Der Derian, "Virtuous War/Virtual Theory," 772.
75. Alex Ryan, "A Personal Reflection on Introducing Design to the U.S. Army," *Medium* (blog), 4 November 2016.
76. Donald A. Schön, *The Reflective Practitioner: How Professionals Think in Action* (New York: Basic Books, 1984), 3–4.
77. Gen W. C. Westmoreland, "Gen. Westmoreland on the Army of the Future," NACLA, 25 September 2007.
78. Der Derian, "Virtuous War/Virtual Theory," 774.
79. Der Derian, "Virtuous War/Virtual Theory," 776.
80. Peter Berger and Thomas Luckmann, *The Social Construction of Reality: A Treatise in the Sociology of Knowledge* (New York: Anchor Books, 1966); Hugh T. Miller, "Scientism versus Social Constructionism in Critical Policy Studies," *Critical Policy Studies* 9, no. 3 (2015): 356–60, https://doi.org/10.1080/19460171.2015.1075734; Jorgen Sandberg, "The Constructions of Social Constructionism," in *Invisible Management: The Social Construction of Leadership*, ed. Sven-Erik Sjostrand, Jorgen Sandberg, and Mats Tyrstrup, Smart Strategies Series (New York: Thomson Learning, 2001), 28–48; Hugh Miller, "Scientism versus Social Constructionism in Critical Policy Studies," *Critical Policy Studies* 9, no. 3 (2015): 356–60; Jean Baudrillard, *Simulacra and Simulation*, trans. Sheila Glaser (Ann Arbor: University of Michigan Press, 2001); Ben Zweibelson, "Preferring Copies with No Originals: Does the Army Training Strategy Train to Fail?," *Military Review* (January–February 2014), 15–25; and James Der Derian, "From War 2.0 to Quantum War: The Superpositionality of Global Violence," *Australian Journal of International Affairs* 67, no. 5 (2013): 575, 582, https://doi.org/10.1080/10357718.2013.822465.
81. This does not include a transhuman entity that may have human-machine abilities that exceed our imaginations and potentially violate this proposal. Again, transhuman entities and singletons in this article are considered beyond humanity as depicted in figure 3.
82. Layton, "Fighting Artificial Intelligence Battles."
83. The unwilling/unwitting quadrant produces a paradox. Unwilling humans would lack the ability to conceptualize precisely what a superior AI or transhuman entity was thinking and doing, but regardless of this ignorance, their distrust or fear of such things would move them into an unwilling stance. Popular science fiction movies and television series feature lower tech, less abled humans against superior alien or AI opponents such as in *The Terminator* franchise, the Borg from *Star Trek: The Next Generation*, and the movie *Independence Day*.
84. Geraci, "Apocalyptic AI," 157.
85. Bostrom, *Superintelligence*, 23. In table 2, a small survey of expert AI researchers indicates human-level machine intelligence has a 10 percent chance of being attained in the next decade, a 50 percent chance between 2040–50, and a 90 percent chance between 2065–93. As Bostrom specifies, these surveys should be taken with some grains of salt. "They are, however, in concordance with results from other surveys . . . [and] are also in line with some recently published interviews with about two dozen researchers in AI-related fields."
86. Bostrom, *Superintelligence*, 159.
87. Roman Khandozhko, "Quantum Tunneling Through the Iron Curtain," *Cahiers Du Monde Russe* 60, nos. 2/3 (April–September 2019): 370–71.
88. Demchak, "China," 99–101.
89. Rapoport, "Editor's Introduction to *On War*."
90. Stuart Russell, "The Purpose Put into the Machine," in *Possible Minds*, 27.

91. *Eco-Terrorism Specifically Examining the Earth Liberation Front and the Animal Liberation Front, Hearing Before the Committee on Environment and Public Works United States Senate*, 119th Cong. (18 May 2005).
92. Ray Kurzweil, "Merging with the Machines, Part 1," 64; Giacomello, "The War of 'Intelligent' Machines May Be Inevitable," 282–83; and Garcia, "Lethal Artificial Intelligence and Change," 339.
93. Geraci, "Apocalyptic AI," 157.
94. Harari, *Sapiens*, 24–25.
95. Ben Zweibelson, "Linear and Nonlinear Thinking: Beyond Reverse-Engineering," *Canadian Military Journal* 16, no. 2 (Spring 2016): 27–35; Ben Zweibelson, "One Piece at a Time: Why Linear Planning and Institutionalisms Promote Military Campaign Failures," *Defence Studies* 15, no. 4 (2015): 360–75, https://doi.org/10.1080/14702436.2015.1113667; and Ben Zweibelson, "An Awkward Tango: Pairing Traditional Military Planning to Design and Why It Currently Fails to Work," *Journal of Military and Strategic Studies* 16, no. 1 (2015): 11–41.
96. Hohyun Sohn, "Singularity Theodicy and Immortality," *Religions* 10, no. 165 (2019): 4–6; and Shatzer, "Fake and Future 'Humans'," 137–39.
97. Max Tegmark, "Let's Aspire to More Than Making Ourselves Obsolete," in *Possible Minds*, 79.
98. Giacomello, "The War of 'Intelligent' Machines May Be Inevitable," 284.

Future Warfare and Responsibility Management in the AI-based Military Decision-making Process

Lieutenant Colonel Alessandro Nalin, Italian Army;
and Paolo Tripodi, PhD

Abstract: The application of artificial intelligence (AI) technology for military use is growing fast. As a result, autonomous weapon systems have been able to erode humans' decision-making power. Once such weapons have been deployed, humans will not be able to change or abort their targets. Although autonomous weapons have a significant decision-making power, currently they are not able to make ethical choices. This article focuses on the ethical implications of AI integration in the military decision-making process and how the characteristics of AI systems with machine learning (ML) capabilities might interact with human decision-making protocols. The authors suggest that in the future, such machines might be able to make ethical decisions that resemble those made by humans. A detailed and precise classification of AI systems, based on strict technical, ethical, and cultural parameters would be critical to identify which weapon is suitable and the most ethical for a given mission.
Keywords: artificial intelligence, AI, machine learning, ML, lethal autonomous weapon systems, decision making, military ethics, commander responsibility

The application of AI technology for military use is growing fast. AI technology already supports several new systems and platforms, both kinetic and nonkinetic (e.g., autonomous drones with explosive payloads or

LtCol Alessandro Nalin, a graduate of the Expeditionary Warfare School and Command and Staff College, Marine Corps University, is the commander of the Italian Army 1st Battalion, 9th Infantry Regiment. The personal viewpoints expressed in this article are Nalin's alone and are not the viewpoints of the Italian Army. Dr. Paolo Tripodi is the ethics branch head and a professor of ethics and leadership at the Lejeune Leadership Institute, Marine Corps University. He is the coeditor with Col Kelly Frushour of *Marines at War: Stories from Afghanistan and Iraq* and the coeditor with LtCol Carroll Connelley (Ret) of *Aspects of Leadership: Ethics, Law, and Spirituality,* both published by Marine Corps University Press.

cyberattacks). Although the human role remains extremely important in the deployment of such weapons, the increasing use of AI has made weapons able to erode humans' decision-making power. Humans have full control of semiautonomous weapon systems. They have partial control, but ultimately retain the ability to override supervised autonomous systems and finally have no control on unsupervised autonomous systems. In the last case, once these systems have been deployed, humans will not be able to change or abort their targets.

Probably the most controversial weapons are the unsupervised autonomous weapons as humans have no ability to control them once they have been deployed. Although the number of such weapons available today is limited, very likely autonomous weapons will continue to be developed and their applicability will expand.[1]

The debate among scholars and practitioners about the use of these weapon systems focuses mainly on their potential targets; however, in this article the authors suggest that we should look at the use of these weapon systems as part of a mission. Semi- and supervised autonomous weapon systems will be deployed by military operators in support of a mission while autonomous weapons will be given a mission to accomplish. The central issue we deal with is that, although autonomous weapons might be empowered with accomplishing a mission and therefore will have a significant decision-making power, very likely they will not be able to make ethical choices. Indeed, what is still missing in the autonomous weapons systems is the ability to explore and consider alternative courses of action if the assigned mission might have unforeseen, unethical consequences. For example, after a loitering munition has been given a mission to destroy a radar station, at the critical moment of attack, a civilian vehicle with a family might be passing by the target. The autonomous weapon system will continue on its mission and might ignore any possible unethical collateral damage.

As a result, the deployment of AI in the battlefield has generated an important debate about the responsibility gap problem. In relation to the use of lethal autonomous weapon systems (LAWS), Ann-Katrien Oimann provides a detailed exploration of the current debate about the responsibility gap problem. She identifies two main positions supported by those who believe that the responsibly gap problem can be filled through the adoption of three different approaches (technical solution, practical arrangement, and holding the system responsible), and those who believe that the problem is unsolvable.[2] In the authors' view there is an approach that, to a certain extent, has the potential to bridge these apparently irreconcilable positions. In fact, a detailed and precise classification of AI systems, based on strict parameters, would be critical to identify which weapon is suitable for a given mission. These parameters might be of a technical, ethical, and cultural nature. For example, suppose that a specific AI system that possesses the right features to perform a given mission is available. In that case, the chain of command that deliberately uses a different AI system should be held responsible for the potentially unethical outcome. However, if the correct AI system is used, but it makes a decision that results in

an unexpected and unethical outcome, the event should be considered as if the decision had been made by a human in good faith. This last case makes evident why the cultural factor, intended as the human willingness to accept AI's potentially harmful decision, will play a decisive role in AI integration in the human decision-making process.

The fact that today's autonomous weapons cannot make ethical choices does not necessarily mean that they will never be able to decide ethically. The authors suggest that in the future such machines might be able to make ethical decisions that resemble those made by humans.

Therefore, it is essential to explore and determine as much as possible which machine decisions, and in particular ethical decisions, could be acceptable for humans. Failing to do so might have two implications. On the one hand, those who are overconfident in AI systems might be inclined to accept all automated decisions, believing that computers are much sharper than human beings. On the other hand, those who are skeptical about computer decisions, as they believe that computers do not ever meet sufficient ethical standards, might not accept any AI-generated decision. It is possible that they would give up the potential benefits that this technology might offer to military decision making in terms of speed and accuracy. The challenge is to develop an approach to AI-driven decision making that identifies a middle ground between the overconfident and skeptical camps.

The parameters used to identify such an intermediate point will consider first the available technology, second the ethical framework, and finally the human predisposition in accepting AI's decisions. This article focuses on the possible ethical implications of AI integration in the military decision-making process. It will explore how the particular characteristics of AI systems with machine learning (ML) capabilities might interact with human decision-making protocols.

The Technological Factor and the AI Galaxy

Autonomous technologies will continue to increase the role of AI, but even more so, they will rely on ML to be able to develop the ability to mimic human thinking and behavior. Currently, as much as AI tries to simulate human intelligence, it still lacks the human curiosity or initiative to learn how to do what it is not programmed for.[3] Inside the AI field, it is important to note the role played by artificial narrow intelligence (ANI) systems that can perform tasks limited to a specific area (Google Maps can plot a route but cannot forecast weather); and artificial general intelligence (AGI) systems that resemble more closely human intelligence as it has "the ability to see the whole" in making decisions.[4] While there are many examples of highly efficient ANI systems already available, currently it has proved to be impossible to develop a reliable AGI system to be used in support of decision-making processes.[5] Therefore, this article will refer mainly to ANI, identifying them as AI systems. Regarding machine learning, James E. Baker defines it as "the capacity of a computer using algorithms, calcu-

lation, and data to learn to better perform programmed tasks and thus optimize functions."[6]

Although capable of giving better, faster outputs while optimizing resources, machines that are programmed to play chess will never take the initiative to learn how to play a different game, for example checkers, because they are limited to perform those actions they are programmed for. The authors do not exclude the fact that AI systems might behave in unpredicted ways. This article is based on the concern that it *might* happen. That means if an AI system is programmed to perform a certain action (play a specific game) that system will improve its skill in playing that game, but it will not take the initiative to learn another game.

These machines are fundamentally reactive, yet they are becoming more proactive within the limits of their specific use. In the last two decades, integration of ML in some applications improved AI's proactive attitude. Brian David Johnson rightly noted that we should expect that "all technologies will use AI and ML. The use of the term could become meaningless because AI and ML will be subsumed by software in general."[7] Consider, for example, Google Nest Thermostat; this home improvement gadget observes users' behavior and pattern of life (e.g., what corrections has the user made in previous days? What time does the user leave home or come back?) to set temperature values in different moments of the day or week. Yet, this improvement in proactive attitude is still far from mirroring the curiosity of the human brain to explore and learn something new without having been directed to do so.

The integration of AI and ML will allow for the creation of machines that can mimic human brain behavior. Stuart J. Russel and Peter Norvig propose a taxonomy that categorizes AI based on the ability of these systems to "think rationally; act rationally; think like humans; and act like humans."[8] Machines will be able to think and act rationally, adopting criteria of a clear definition of what is rationally right and wrong.[9] What is right and wrong follows a static course of action, therefore it is not going to change. The general expectation from these machines is that, given a specific set of inputs, outputs will remain the same over time. The limit of these systems emerges when they have to make decisions in situations in which there is no right or wrong model for answers. Machines that think and act like humans do not refer to rationality but try to behave like humans. This difference implies the possibility to learn from experience among all the other things. ML enables machines to learn from experience.[10] However, if the reference model is based on true or false answers, they are not sufficient to replicate human behavior.

The complexity of human decision making requires an approach that should go beyond the binary logic of yes or no typical of a computer algorithm. Bahman Zohuri and Moghaddam Masoud analyze in details the concept of fuzzy logic: "an approach to computing based on 'degrees of truth' rather than the usual 'true or false' (1 or 0) Boolean logic on which the modern computer is based."[11] Fuzzy logic is fundamental for the building of effective AI systems

as it processes decisions, categorizing them not only as entirely right or wrong but also on a continuum between these categories. Arguably, the combination of ML and fuzzy logic allows the creation of autonomous systems that can effectively mimic human reasoning and decision making in its peculiar ability to learn from experience and express judgment like, for example, *almost right* or *not completely wrong*.

The Ethical Factors: Ethical Aspects of the Integration of AI in Decision Making

AI systems act in the ethics realm, yet their qualification as ethical agents requires some consideration. James H. Moor provides an analysis of the nature of different machine ethics through the different typologies of ethical agency. He differentiates among machines that have implicit agency, machines whose inherent design prevents unethical behavior (i.e., "pharmacy software that checks for and reports on drug interactions"); explicit agency, machines able to "represent ethics explicitly and then operate effectively on the basis of this knowledge"; and full ethical agency, machines that possess "consciousness, intentionality, and free will."[12] At the moment, there are no machines that possess these three characteristics; however, according to Moor, AI systems are ethical enough to act as ethical agents, with all the necessary limitations for their specific functions.[13] Humans can assess a machine's ethics and employ it in its specific and limited sector when built for a particular purpose, like a tracking and triage system designed for disaster relief operations.[14] Humans could trust completely the ethics of AI systems employed for unconstrained general purposes if they would achieve the status of full moral agency. Yet, "narrow" AI (ANI) is the only system currently available.

The use of AI systems to support self-driving vehicles has generated a valuable debate about how to integrate ethics in AI systems to develop their ability to make ethical decisions. Vincent Conitzer et al. found that, in this field, a rationalist ethical approach alone would probably lead to decisions that maximize utility but might not be entirely ethical.[15] They suggested that initial rationalistic approaches should integrate later on a machine learning approach based on "human-labeled instances."[16] As a result, after a system has learned how to decide following a strictly rationalistic approach, humans should continue to feed such systems with information about what constitute a right ethical decision in a variety of different situations.

Noah Goodall, in "Ethical Decision Making During Automated Vehicles Crashes," takes a similar approach, but with a more defined practical sequence of actions to better integrate ethics in AI systems for self-driving vehicles. Goodall identifies three phases in the development of ethical AI systems. In the first phase, vehicles use a rationalistic moral system (e.g., consequentialism) taking action to minimize the impact of a crash based on general outcomes (e.g., injuries are preferable to fatalities).[17] In the second phase, while building on the rules established in the first phase, vehicles will learn how to make ethical

decisions observing human choices across a range of real-world and simulated crash scenarios.[18] The third and final phase requires an automated vehicle to explain its decisions using "natural language" so that humans may understand and correct its highly complex and potentially incomprehensible-to-humans logic.[19] This ability will help humans understand why vehicles make certain and maybe unexpected choices, and developers will be able to understand and, more importantly, correct wrong behaviors and decisions.[20]

Conitzer et al. and Goodall concur on a phased AI training that begins with the implementation of a consequentialist approach and continues with the integration of human-based experience and expertise.[21] The fast development of technological improvements leads us to believe that probably soon we will be able to build an effective ethical framework in AI systems. Developers can build ethics into AI systems adopting either a top-down or a bottom-up approach. With the former, developers will code into AI systems all the desired ethical principles (i.e., "Asimov's Three Laws of Robotics, the Ten Commandments or . . . Kant's categorical imperative").[22] With the bottom-up approach, machines will learn from human behavior in multiple situations without a specific base of moral or ethical knowledge.[23]

With the top-down approach, it is not necessary to program all the possible decisions that machines might take in different situations as they will decide in line with their embedded principles.[24] This approach highlights the importance of fuzzy logic implementation as it allows AI systems to go beyond the simple dichotomy of right and wrong.[25] The possibility to make decisions that are sufficiently right or not completely wrong widens the range of possible choices in which humans can identify those acceptable to them. However, the moral strength of humans' decisions is based on a lifelong ethical development that typically begins from childhood, while top-down AI's ethics are passive to external changes. According to Amitai Etzioni and Oren Etzioni, a top-down approach is "highly implausible."[26]

With the bottom-up approach, machines learn from human behavior in multiple situations without a specific base of moral or ethical knowledge.[27] Machines observe how humans behave and react to situations. From these observations, machines create their set of rules to make decisions independently. The main concern with this approach is that humans are not flawless and make mistakes that AI systems may not recognize and consequently absorb as a model of behavior.[28] It is apparent that both top-down and bottom-up approaches present flaws that can hamper the ethical competence of machines and are hard to mitigate.

Assuming that technology can support the development of a full moral agent AI, such a machine would be the most evolved AI system. The application of a top-down approach would mitigate the ability to learn mainly from experiences, because machines rely mostly on human-labeled data or instructions inputted by a limited number of individuals instead of having access to the entire human experience on a specific action/behavior. The bottom-up ap-

proach would expose the machine to the human's natural flaws and misbehavior and will allow for the development of a machine that, like humans, might make mistakes. Arguably, this last scenario might lead humans to over rely on a system that, although similarly inaccurate as humans, could be more effective because of its speed and user friendliness.

An additional issue to consider is about perception of responsibility in relation to decisions made by AI systems. Due to their high level of autonomy to identify and engage military targets while pursuing their mission, the employment of autonomous weapon systems raises concerns about responsibility and accountability in cases of wrong decisions and actions.[29] Mark Ryan raises the issue of responsibility of AI decisions, pointing out that if, on the one hand, it is unfair to assign the responsibility of wrong AI systems' decisions to their designers because these systems can learn; on the other hand, AI cannot be responsible for its decisions because it is not a moral agent.[30]

Ross W. Bellaby takes into consideration those aspects related to the responsibility of military AI systems' failures analyzing different cases involving autonomous weapons or remote-controlled weapons systems. He argues that responsibility goes together with the possibility of making decisions.[31] The rationale is that if an ethical failure happens using a remote-controlled weapon, the human pilot or the human chain of command will be responsible for that failure. However, if an autonomous weapon system makes an ethical error, it would be its AI's responsibility, but AI is not an entity subject to legal action, so the responsibility should go to its developers or to those who decide to employ that system for that mission. While developers might argue that they have written the code a long time before and without the information available at the moment of the failure, the human chain of command might also maintain that they cannot influence decisions and issue orders to avoid the failure.[32] The identification, made on the most objective of bases, of which is the best AI for each specific situation would be a helpful tool to establish, at least if there is some responsibility for having resorted to the wrong AI system.

Eventually, the complete reliance on AI systems can create a gap in the responsibility and accountability chain that ultimately can "create distance from and mitigate the responsibility of the military operators or commanders using the system."[33] The risk is that humans, feeling themselves free from any responsibility, might fail to consider the ethical implication of decisions made by AI systems. However, in all those cases in which there is not a clear, unpredictable technological failure, the responsibility for mistakes made by a full autonomous weapon system while pursing the assigned mission should rest with its chain of command.

Future Battlefield Environments: Capitalizing Advantages of AI in the Decision-making Process

In the near future, combatants will confront enemies capable of conducting multidomain operations (MDOs) that will take place simultaneously in the

air, land, maritime, space, and cyberspace. In MDOs, humans might find it difficult to make fast and, more importantly, timely decisions. It is in such environments that automated systems will be extremely beneficial to support the AI's human-out-of-the-loop decision-making process.[34] This support will be crucial to save time and gain an advantage on the enemy. Anupam Tiwari and Adarsh Tiwari noted that "often the timelines are dominated by the time it takes to move equipment or people or even just the time that munitions are moving to targets. It is important not to overstate the value of accelerating the decision process in these cases."[35] However, this approach does not consider the enduring nature of decision-making processes; once the headquarters issues its order and troops move on the battlefield, the observe-orient-decide-act (OODA loop) cycle keeps on running to maintain the order consistent with changes in the common operational picture. For this reason, AI systems might be far more relevant than it may appear.

Moreover, in relation to the OODA loop, there are similarities in the way machines and humans make decisions. For example, Amitai and Oren Etzioni state that autonomous vehicles "are programmed to collect information, process it, draw conclusions, and change the ways they conduct themselves accordingly, without human intervention or guidance."[36] These vehicles are programmed to approach the decision-making process in the same way militaries do through the OODA loop.

AI systems improve data collection and accelerate the elaboration and update of situational awareness. Operations are information-driven, and success often is on the side of those who possess better situational awareness of the battlefield. More information allows planners to predict enemies' moves and, possibly, preempt them.[37] It is reasonable to think that today's significant AI limitations, for example its highly restricted "ability to recognize images (observe) outside of certain conditions" are transitory.[38] In the future, AI will be able to improve intelligence collection consistently, increasing the sharpness of situational awareness by different applications like, for example, improved image, facial, voice recognition, aggregation of data, and translation.[39] An indication of this future scenario is the U.S. Army development of the capability to deploy swarms of drones to "increase situational awareness with persistent reconnaissance."[40] More refined and vast amounts of information available in less time will allow those who possess this technology to have a decisive advantage over the opponent in terms of situational awareness at the beginning of the OODA loop cycle.

Automated instruments of data processing with the support of AI can provide better intelligence and suggest options for military problems. Genetics, culture, and consolidated expertise heavily influence each decision maker's mental model to process information and produce intelligence.[41] In simple words, the observation phase in a multidomain environment can quickly run out of humans' analysis capabilities, reducing the speed at which military planners can make decisions and act.[42] To mitigate the shortage of analysis capabilities, AI

and ML systems help implement an out-of-the-loop system in which a human's contribution is limited strictly to the necessary, which might be the activation of the system or definition of parameters used to identify an actor on the battlefield as foe.[43] According to Daniel J. Owen, AI will play a prominent role in the transformation of "human decision-makers' abilities to orient by integrating and synthesizing massive, disparate information sources."[44] Nonetheless, no matter whether a top-down or a bottom-up approach is adopted, it is humans who train AI/ML systems.

If the "orientation" phase ends with defining a number of courses of action (COAs), the next phase is when these COAs are compared and weighted to make the decision. Assuming, as a hypothesis, the complete reliability of AI systems, they could decide what option humans should implement. Time has proven to be a critical resource for success, and in the near future it looks as if every fraction of a second could be decisive. Humans are not likely to be self-sufficient in managing situations at the same pace AI/ML can do.[45] Improved autonomous systems trained to implement fuzzy logic can provide accurate and fast decisions.[46]

During the "action" phase, AI can improve force protection. Indeed, AI/ML systems have the capability to run robotics and autonomous systems (RAS). Many different typologies of RAS are being tested to decrease the human involvement in combat and improve the performance of armed forces. Being able to count on advanced autonomy allows for RAS to be able to perform dangerous tasks for longer times and at greater distances while reducing the number of humans at risk.[47]

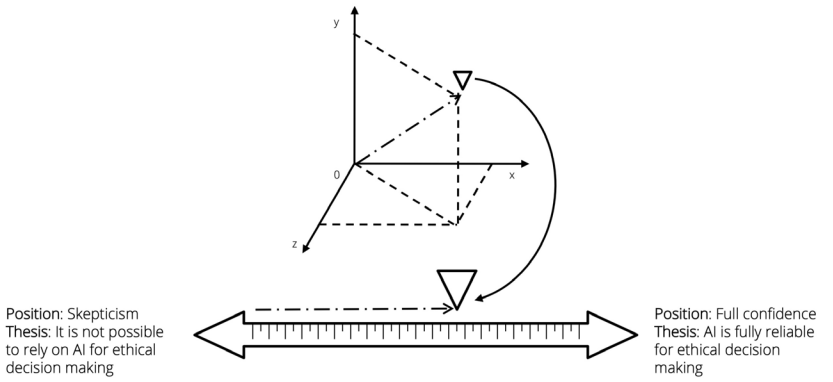
Balancing the Technological, Ethical, and Cultural Factors

In the future application of technological innovations, it is crucial to identify an intermediate point between what appear to be two extreme approaches humans have to AI. On the one side is the absolute skepticism and on the opposite side an unconditional trust to the use of AI. In the authors' view, three factors should be considered to identify such an intermediate point: technological, ethical, and cultural.[48] The position of this intermediate point, which might actually be either located at times closer to one side than the other, depends on the relevance of each of the three factors.

In figure 1, at the point of origin of the axes, the three factors are at their respective minimums. The minimum value represents a condition in which technology does not support data processing; ethics cannot be integrated into decision making; and, from a cultural point of view, it is not possible to accept that a machine can make decisions on behalf of human beings. The triangle represents the point at which the three factors reach the balance, creating the condition for an AI system suitable for effective and ethical decision making. The dotted-line arrow represents the distance from the origin to the balancing point.

Even though currently this balancing point has already moved away from

Figure 1. Balance of factors



x	Technological	From basic algorithm to fully developed AI/ML systems
y	Ethical	From supervised to unsupervised ethical framework
z	Cultural	From culturally unprepared to prepared
0	Origin of axis	
—·—·—→	Resultant	Vector resulting from the combination of factors

Source: courtesy of the authors, adapted by MCUP.

the origin, AI is far from being a fully trustable support for decision making that has an ethical dimension. It is possible to describe the current situation using again the example of the GPS navigator. The technology allows processing of relevant data to identify the position in the world of an individual, relate it to a different geographical point, and evaluate all the variables (time, space, and laws) to provide such individual with the best path to reach the endpoint. The moral implications of this choice are simple enough that it is possible to make this decision ethically acceptable through a basic utilitarian model that maximizes the subject's happiness while decreasing their suffering.

If the priority of the subject is the duration of travel, the AI will develop a path that although it is a longer route perhaps requiring some tolls, it is still the fastest compared to other options. In addition, applications like Google Maps are implementing new features to calculate routes that preserve gasoline consumption to help reduce CO_2 emissions. Finally, human beings are now accustomed to using GPS and have embraced a culture that easily accepts such a tool to support decision making. It is also possible to conclude that humans trust GPS navigators because the three factors of technology, ethics, and culture blend together in a well-balanced, mutually supporting interaction. This example shows that humans trust GPS because they are used to its AI (cultural factor) that does the math right (technological factor) without incurring the risk of being immoral (ethical factor).

However, if one or more of these factors is off-balance, it would not be safe

for humans to rely on AI systems. For example, it is interesting to imagine what might happen when one of the three factors compromise the overall balance. ML technologies might improve decision making, allowing a qualitative leap forward for humankind. Yet, due to the inherent design of its hardware and software, such a technology might be affected by a lack of transparency that could affect how humans control AI/ML systems.[49]

The potential lack of transparency should be offset; first by the possibility of making machine's decisions morally acceptable through an ethical framework suitable for meeting the specific requirements for the task; second by an improved habit of using this technology by humans. The first mitigation avoids morally unwanted second- and third-order effects, while the second reduces humans' natural fear of the unknown. This latter aspect deserves some more explanations.

Human superiority in decision making still exists, but AI might still be extremely helpful in situations in which this superiority is not enough. For example, the ability to always see the big picture, combined with a solid ethical background, makes humans sharper in broad spectrum decision making. Nonetheless, AI's ability to process a more significant amount of data per second could make AI decisive in narrow and particular situations. Humans will have to acknowledge that, under certain conditions, it is possible that the best of their decisions might be worse than the AI systems' worst ones. Indeed, when the enemy launches a missile attack, an accurate but late human decision about a countermissile artillery reaction is more dangerous than a not wholly right yet on time AI decision. This allows at least for mitigating damages due to AI's speed of decision making. In the future, technological improvements will allow the design of increasingly refined AI systems able to make the same types of decisions as humans. However, humans will train these AI systems directly (top-down) or indirectly (bottom-up) according to their knowledge or through their experiences.[50] It is reasonable to believe that, at the end of the training, AI systems will be able to replicate the dynamics of human reasoning very closely; such a reasoning hopefully will include ethical thinking and will have the same fallacies that ethical thinking has in human beings. Nonetheless, humans should make reliance on AI a part of their culture, in particular when situations require processing a disproportionate amount of data in a very limited amount of time. Therefore, it is highly probable that AI systems will still make mistakes, yet given certain conditions (e.g., time available and amount of data), they could be more reliable than humans.

Machines' fallibility might not be a problem, yet humans could hardly accept it. The problem is that, in some situations, especially those that involve people's safety, the same mistake might be more tolerable if made by humans than by a machine. There are two reasons behind this distrust toward AI systems: first, it is accepted that humans can make mistakes while a machine should be flawless; second, there would be nobody to blame when an AI system gets it wrong. Indeed, it is possible to punish a human that has made a mistake, but not a machine.[51]

While these two reasons make it difficult to accept a machine's decision about the safety of human beings, the perceived necessity of a decision-making support tool is at the base of the cultural propensity of humans to accept that an AI system might decide entirely or partially on their behalf.

Humans are committed to research and develop new technologies because they believe that such technologies will improve the well-being of humanity and people's quality of life. This perception affects how much humans are willing to rely on AI. The more difficult it is for human beings to guarantee high standards of speed and effectiveness in a given task, the more they will feel the need for technological support in order to increase their performance and, as a result, they will be more willing to rely on machines. Therefore, humans would safely rely on AI systems as long as they see the machine's worst performance as a better output compared to the human's best performance on the same action.

Conclusion

Having seen the potential that AI has to improve humans' efficiency in ethical decision making, it is crucial for individuals to make every effort to define objective parameters to identify an AI system's balancing point. AI systems should be cataloged and associated with certain situational conditions (e.g., urgency, or the amount of information to be processed) to allow users to identify which ones bring the best benefit to their purpose.

In this way, military commanders could be better positioned to decide which tools to use and under what circumstances. Commanders can drastically reduce the time invested in decision-making processes and be aware of the incomplete suitability of a given system and to implement the necessary arrangements to mitigate the effects of possible errors. The importance of this process lies in the fact that AI is already widespread and accessible to all competitors. Therefore, not being able to optimize the use of AI systems would mean starting with a considerable disadvantage that could compromise the ability to achieve and maintain the initiative on the enemy, thus accepting fighting on the enemy's terms. Very likely the employment of AI in the military decision-making process is unavoidable, and for this reason military leaders and AI developers might study how to build ethics into AI systems. There are different degrees of possible moral machines, from the implementation of basic utilitarian frameworks, up to ethically more complex and sophisticated systems. These different kinds of machines will be able to perform at different complex stages of the decision-making process.

Military leaders should be accountable for the decision they make. This accountability must also endure when AI systems are used to support their decision-making process. Having a catalog that identifies what device is suitable and for what purpose in different situations is a fundamental condition to apportion responsibility on the right individual. If commanders intentionally do not use the appropriate device for a given mission, they are responsible for the decision. Yet, if commanders choose the correct device but the device fails, and

if the follow-on investigation on other actors' responsibility (e.g., AI designers, code developers) determines that none of the actors has a direct responsibility, probably humans should accept that the outcome was unpredictable.

Future studies should investigate how to assign a value to the weight of the three factors at the balance point. As far as technology is concerned, it could be a simple but effective way to rely on the possession or not of specific technical characteristics or certain components. Regarding ethics, it could be helpful to define a scale of values to be associated with a particular ethical model that is purely based on utilitarian logic or can also consider more profound implications or evaluate second- and third-order effects. Finally, the cultural factor could represent the most challenging obstacle to overcome due to its subjective and, in a certain sense, ephemeral nature. However, parameters such as the diffusion among the population or how long a device has been in use can be the starting points to establish values.

Endnotes

1. The use of unsupervised autonomous weapons is growing fast and their use has proved to be an advantage. The case of the 2020 Nagorno-Karabakh war during which Azerbaijan used with excellent results the Israeli-made IAI HAROP Loitering Munitions against Armenia is evidence that such weapons can be extremely powerful. See Brennan Deveraux, "Loitering Munitions in Ukraine and Beyond," *War on the Rocks*, 22 April 2022.
2. Ann-Katrien Oimann, "The Responsibility Gap and LAWS: A Critical Mapping of the Debate," *Philosophy & Technology* 36, no. 3 (2023): 7–16, https://doi.org/10.1007/s13347-022-00602-7.
3. ML does not preclude developing a sense of curiosity. Yet, today AI/ML systems are not able to learn something if humans do not direct them to do so. For example, a thermostat based on an AI/ML system can learn how to improve comfort for humans in their houses, but it is not curious to learn about how to play chess until a human changes its algorithm.
4. Ragnar Fjelland, "Why General Artificial Intelligence Will Not Be Realized," *Humanities & Social Sciences Communications* 7, no. 1 (2020): 2, https://doi.org/10.1057/s41599-020-0494-4.
5. Fjelland, "Why General Artificial Intelligence Will Not be Realized," 7.
6. James E. *Baker, The Centaur's Dilemma: National Security Law for the Coming AI Revolution* (Washington, DC: Brookings Institution Press, 2020), 14.
7. Brian David Johnson, "Autonomous Sentient Technologies and the Future of Ethical Business," in *Ethics @ Work: Dilemmas of the Near Future and How Your Organization Can Solve Them*, ed. Kris Østergaard (Middletown, DE: Rehumanize Publishing, 2022), 69.
8. Stuart J. Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Hoboken, NJ: Pearson, 2009), 2–5.
9. The authors agree with the fact that a strong Kantian view of moral behavior applied to human beings is implausible. Here, we focus on machines whose behavior is determined by the specific input they have received. As Russel and Norvig suggest, machines think rationally and act rationally. To think like humans and behave like humans, they will need to use ML and adopt a nonbinary logic approach.
10. Artificial neural networks (ANNs) are among the most used solution to implement learning skills on machines. The main idea of this technology is to mimic the human brain's neural connections throughout different layers of artificial neural nodes. Incor-

rect outputs generate adaptation in the inner and hidden neural layers until the results are correct. Therefore, machines that are more experienced would be more effective. Yet, this technology still suffers transparency issues because, like a human brain, it is not clear how machines manage these adaptations. This lack of transparency could likely prevent human understanding if and why failures occur in case of false positive output.

11. Bahman Zohuri and Moghaddam Masoud, *Neural Network Driven Artificial Intelligence* (Hauppauge, NY: Nova Science Publishers, 2017), 15–17.
12. James H. Moor, "The Nature, Importance, and Difficulty of Machine Ethics," *IEEE Intelligent Systems* 21, no. 4 (2006): 19–20, https://doi.org/10.1109/MIS.2006.80.
13. Moor, "The Nature, Importance, and Difficulty of Machine Ethics," 20.
14. Moor, "The Nature, Importance, and Difficulty of Machine Ethics," 20.
15. Vincent Conitzer et al., "Moral Decision Making Frameworks for Artificial Intelligence," *Proceedings of the AAAI Conference on Artificial Intelligence* 31, no. 1 (2017): 1–5, https://doi.org/10.1609/aaai.v31i1.11140.
16. Conitzer et al., "Moral Decision Making Frameworks for Artificial Intelligence," 1–5.
17. Noah J. Goodall, "Ethical Decision Making during Automated Vehicle Crashes," *Transportation Research Record* 2,424, no. 1 (2014): 7–12, https://doi.org/10.3141/2424-07.
18. Goodall, "Ethical Decision Making during Automated Vehicle Crashes," 7–12.
19. Goodall, "Ethical Decision Making during Automated Vehicle Crashes," 7–12. What happens in the hidden layers of ANN is obscure to humans. Software developers are not able to understand how ANN-based machines reach their output and how they correct themselves during the process. An interface that translates machine language into a natural language would be beneficial in understanding how machines work and where and why possible mistakes occur.
20. Goodall, "Ethical Decision Making during Automated Vehicle Crashes"; and Alan F. Winfield and Marina Jirotka, "Ethical Governance Is Essential to Building Trust in Robotics and Artificial Intelligence Systems," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, no. 2,133 (2018): https://doi.org/10.1098/rsta.2018.0085
21. Conitzer et al., "Moral Decision Making Frameworks for Artificial Intelligence," 1–5; and Goodall, "Ethical Decision Making during Automated Vehicle Crashes."
22. Amitai Etzioni and Oren Etzioni, "Incorporating Ethics into Artificial Intelligence," *Journal of Ethics* 21, no. 4 (2017): 403–18, https://doi.org/10.1007/s10892-017-9252-2.
23. Etzioni and Etzioni, "Incorporating Ethics into Artificial Intelligence," 406–7.
24. Etzioni and Etzioni, "Incorporating Ethics into Artificial Intelligence," 405–6.
25. Zohuri and Masoud, *Neural Network Driven Artificial Intelligence*, 15–17; and Etzioni and Etzioni, "Incorporating Ethics into Artificial Intelligence," 403–18.
26. Etzioni and Etzioni, "Incorporating Ethics into Artificial Intelligence," 403–18.
27. Etzioni and Etzioni, "Incorporating Ethics into Artificial Intelligence," 406–7.
28. Etzioni and Etzioni, "Incorporating Ethics into Artificial Intelligence," 407.
29. Forrest E. Morgan et al., *Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World* (Santa Monica, CA: Rand, 2020), xiii–xiv, https://doi.org/10.7249/RR3139-1.
30. Mark Ryan, "In AI We Trust: Ethics, Artificial Intelligence, and Reliability," *Science and Engineering Ethics* 26, no. 5 (October 2020): 13–14, 2,749–67, https://doi.org/10.1007/s11948-020-00228-y.
31. Ross W. Bellaby, "Can AI Weapons Make Ethical Decisions?," *Criminal Justice Ethics* 40, no. 2 (2021): 86–107, https://doi.org/10.1080/0731129X.2021.1951459.
32. Bellaby, "Can AI Weapons Make Ethical Decisions?," 95–97.
33. Morgan et al., *Military Applications of Artificial Intelligence.*
34. Anupam Tiwari and Adarsh Tiwari, "Automation in Decision OODA: Loop, Spiral or Fractal," *i-manager's Journal on Communication Engineering and Systems* 9, no. 2 (2020): 1–8, http://dx.doi.org/10.26634/jcs.9.2.18116.

35. Tiwari and Tiwari, "Automation in Decision OODA," 16–17.
36. Etzioni and Etzioni, "Incorporating Ethics into Artificial Intelligence."
37. AI-ML's computing rate is growing constantly, not considering quantum computing. For this reason, the expectation that AI will be able to process large amounts of information is realistic. If the information—any amount of information—fed to the algorithms is accurate and has the correct standard, AI will continue to provide increasingly better outcomes. The problem is when a large amount of information is provided to the AI unfiltered and without a proper application of correct standard.
38. Daniels J. Owen, "Speeding Up the OODA Loop with AI: A Helpful or Limiting Framework" (paper, Joint Air & Space Power Conference, 7–9 September 2021), 159–67.
39. Owen, "Speeding Up the OODA Loop with AI," 159–67; and Baker, *The Centaur's Dilemma*, 30–31.
40. *The U.S. Army Robotic and Autonomous System Strategy* (Fort Eustis, VA: U.S. Army Training and Doctrine Command, 2017), 10.
41. Tiwari and Tiwari, "Automation in Decision OODA," 1–8.
42. Fernando De la Cruz Caravaca, "Dynamic C2 Synchronized Across Domains: Senior Leader Perspective" (paper, Joint Air & Space Power Conference, 7–9 September 2021), 81–89.
43. Tiwari and Tiwari, "Automation in Decision OODA," 1–8.
44. Owen, "Speeding Up the OODA Loop with AI," 159–67.
45. Owen, "Speeding Up the OODA Loop with AI," 159–67.
46. Michael Doumpos and Evangelos Grigoroudis, eds., *Multicriteria Decision Aid and Artificial Intelligence: Links, Theory and Applications* (Chichester, UK: John Wiley & Sons, 2013), 34.
47. *The U.S. Army Robotic and Autonomous System Strategy*, 3.
48. This idea does not necessarily depart from the majority of the literature. This idea strives to provide a better approach on how to assign responsibility when using AI in the decision-making process. Defining parameters that clearly put in relation each AI system with its possible field of employment would be beneficial. For example, defining responsibility for choosing the right AI system for the specific situation. This choice should not be linked only to the question: Is this system able to make this decision? It should comprehend also questions related to the ability to be ethical in the decision-making process and to the humans' acceptance rate for that specific AI system.
49. Alan F. Winfield and Marina Jirotka, "Ethical Governance Is Essential to Building Trust in Robotics and Artificial Intelligence Systems," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, no. 2,133 (2018): 6, https://doi.org/10.1098/rsta.2018.0085.
50. Etzioni and Etzioni, "Incorporating Ethics into Artificial Intelligence," 403–18.
51. Alexander McNamara, "Majority of Public Believe 'AI Should Not Make Any Mistakes'," Science Focus, 6 July 2020.

Colonel John Boyd's Thoughts on Disruption
A Useful Effects Spiral from Uncertainty to Chaos

Brian R. Price, PhD

Abstract: A close examination of John R. Boyd's concept of disruption as recorded in his 1987 presentation, "An Organic Design for Command and Control." This article draws attention to a series of disruptive actions Boyd lists, including uncertainty, doubt, mistrust, confusion, disorder, fear, panic, and chaos, noting that the list begins with the mildest effect but that it progresses regularly toward collapse and chaos. The author argues that Boyd was specific in listing these effects in order and notes that this cycle could be developed into a useful effects spiral, which, once understood, can be catalyzed to enhance enemy disruption in a Joint all-domain operations (JADO) environment. In the postscript, this article argues that officers seeking to operate in a multi- or all-domain environment can benefit from a broad educational base to unlock creativity in approaching wicked problem sets. This creativity, when coupled with concepts like the effects spiral, can enhance traditional maneuver and combat, triggering an opponent's collapse without the need for annihilation.
Keywords: John R. Boyd, OODA, decision cycle, psychological warfare, all-domain, dislocation, strategy, JADO, JADC2, creativity, education, PME

Yours Truly: Operate inside adversary's observation-orientation-decision-action loops to enmesh adversary in a world of *uncertainty, doubt, mistrust, confusion, disorder, fear, panic, chaos* . . .

Dr. Brian R. Price is associate professor of military and strategic studies at the U.S. Air Force's Air Command and Staff College, is course director for the Joint Campaigning (JPME I) capstone course, and recently taught within the Joint All-Domain Strategist (JADS) concentration. He has a BA in political science from the University of California, Los Angeles (UCLA) and a PhD in military history from the University of North Texas. He has published a number of articles, reviews, and several books that focus on the nexus between culture, technology, and war in post-WWII modern and late medieval. The opinions and analyses presented here reflect the views of the author and are not necessarily the views of the Air Command and Staff College, Air University, the Department of Defense, or the U.S. government.

Journal of Advanced Military Studies vol. 14, no. 1
Spring 2023
www.usmcu.edu/mcupress
https://doi.org/10.21140/mcuj.20231401004

> and/or fold adversary back inside himself so that he cannot cope with events/efforts as they unfold.
>
> ~ John R. Boyd[1]

> From John Boyd we learned about competitive decision making on the battlefield—compressing time, using time as an ally.
>
> ~ General Charles C. Krulak[2]

Introduction

This article considers a list of conditions or degrees of collapse noted by U.S. Air Force colonel John R. Boyd in his 1987 briefing, "Organic Design for Command and Control," listed in the above quote. It observes that, when uncertainty and doubt are infused into a complex, adaptive system, that system can follow an entropic, cascading decline as a spiral into confusion, disorder, fear, panic, and chaos. Even if the opponent does not progress far down the spiral, uncertainty and doubt reduce the speed and quality of decision making, slowing the crucial orientation phase of the observe, orient, decide, act (OODA) loop. Further, it notes that Boyd subscribed to a systems view of opponents seen as a holistic organism, featuring multiple centers of gravity (COGs). He proposed that decisive points are often found at connections between COG as critical vulnerabilities, possibly within operational reach, as when well-defended centers of gravity prove difficult to target or as simultaneous catalyzing strikes (*nebenpunkte*). An understanding of the opponent's many nodes and connections is fundamental. Understanding this and how conditions of the effects spiral interact enables planners and commanders to creatively mix physical and informational weapons to degrade decision-making capability or to spark a cascading collapse of the enemy within the human domain.

John Boyd and the OODA Loop

U.S. Air Force colonel John Boyd left a rich legacy of ideas relating to war and competition. Much of it applies to the Department of Defense's quest to operationalize Joint all-domain operations (JADO). Of course, Boyd is most famous for the OODA loop—observe, orient, decide, act—the conflict decision-making heuristic that underpins the Joint all-domain command and control (JADC2) architecture and conceptions of JADO operations.[3] If one is able to "turn inside" the opponent's decision cycle, especially through superior situational awareness, the thinking goes, one can force the opponent to *re*-observe, *re*orient, *re*-decide, and *re*act, capturing the initiative and driving the action. This tempo and initiative-based approach has dominated American and Western thinking about war and competition since at least the post-Vietnam era, displacing conceptions focused more on mass and attrition.[4] Boyd's ideas loom large in this shift, and while he was not the only one advocating this kind of approach, his perceptive synthesis and passionate crusade led the defense estab-

lishment to a new conception of war and conflict. Within the U.S. Marines, his ideas underpin the concept of maneuver warfare.[5]

Boyd's "Organic Design for Command & Control"

Boyd was convinced that survival equated with adaptation, with openness, and that to become isolated or closed off signified defeat, dissolution, and death. In his presentations, he viewed organizations and even nations as collections of organisms forming complex adaptive systems, resilient until they were isolated.[6] The systems view of the environment and the opponent is as old as Carl von Clausewitz, but it was embraced as an outgrowth of the Army Air Corps and Air Force thinking with respect to bombing campaigns, expressed as industrial web theory.[7] Boyd took it much further, synthesizing then-current literature and seeing the enemy system as a collection of entities (what we might today call *nodes*), with crucial connections between.[8]

While another Air Force colonel—John A. Warden III—advocated the targeting of multiple nodes simultaneously, Boyd advocated striking connections between nodes to isolate and disrupt.[9] As he stated directly in his "Patterns of Conflict" briefing at Quantico, on 25 April and 2–3 May 1989, "If you want to subvert or pull apart a guy's center of gravity . . . you want to find out what are those bonds, those connections that permit that organic whole to exist."[10] Though he did not say it, such connections would be more difficult to defend, and thus softer targets than would be the key nodes themselves, often defended as known critical vulnerabilities.[11] In fact, he emphasized that these connections were not necessary physical when he noted, "by striking at those tendons, connections. . . . In other words, you want to generate many non-cooperative centers of gravity."[12] It is a useful principle that targeters, planners, and commanders can profitably use in single-, cross-, multi-, or all-domain operations.

Boyd is also well-known for communicating through marathon briefings, conducted throughout the late 1970s, through the 1980s, and into the 1990s, most famously his monumental "Patterns of Conflict" noted above. In "Patterns," Boyd synthesized his study of conflict, history, and strategy.[13] He concluded that disaggregating the opponent was ultimately the goal in any conflict.[14] His 1987 "Organic Design for Command and Control" applied these ideas directly to command and control (C2).[15] Through 37 slides, he discussed—conceptually—how to create a C2 system that maximizes adaptability, resiliency, and harmony while challenging that of the opponent. In discussing multiple, simultaneous attacks—*nebenpunkte*—he could have been discussing the foundations and aspirations for the Army's multidomain operations or the Department of Defense's Joint all-domain operations (MDO or JADO).[16]

An Effects Spiral

On slide number seven of "Organic Design for Command and Control," Boyd articulated his conditions in what appear to be just a list, but a closer look reveals a subtle but very useful scale of effects:

> Operate inside adversary's observation-orientation-decision-action loops to enmesh adversary in a world of **uncertainty**, **doubt**, **mistrust**, **confusion**, **disorder**, **fear**, **panic**, **chaos**, . . . and/or fold adversary back inside himself so that he cannot cope with events/efforts as they unfold [emphasis added].

Boyd's slides do not characterize the disruptive effects he sought as a spiral, but nonetheless they do have a relationship to one another and were presented by Boyd in what was doubtlessly a carefully crafted order, designed such that "weaknesses thereby generate doubt and uncertainty which magnify into panic and chaos."[17] As effects, they are particularly useful to commanders and planners, because while one might like to immediately cause collapse through shock and awe, it may be much more realistic to seek results that multiply and magnify lesser effects such as uncertainty and doubt. This works in what we might call the human domain, the macrosphere that encompasses belief and calculation of each human in the system, characterized by unique and overlaying segments that clash in any given conflict to create a complex, adaptive system within human society, or soon, human-machine society.

The effects spiral multiplies and compounds small actions that act together to erode trust and move an individual, an organization, or a system further down the spiral toward confusion, disorder, and disaggregation, folding the adversary "back inside himself," as Boyd termed it.[18] The objective is to isolate and divide, creating "many noncooperative conflicting centers of gravity [that] paralyze [the] adversary by denying him the opportunity to operate in a directed fashion."[19] This is very different than conventional planning wisdom, which advocates for attacking the centers of gravity through direct or indirect means.[20] Boyd notes that we might, additionally, attack connections *between* centers of gravity, not just the COGs themselves.

In each of the conditions, recovering is more difficult and time-consuming. Left unchecked or propelled with further momentum, one leads to the next in an effects spiral. Each represents a higher state of disruption, beginning with mere uncertainty and doubt and ending with panic and chaos. An effect could be targeted to cause a particular state, but more likely it will be the combination of multiple effects that begins and accelerates the cycle. The cascade of compounded effects is more than simply multiple effects added together.

Each of these states is based fundamentally on perception, compared to physical reality. To Boyd, perception was comprised of shifting *observations* and feedback, interpreted through *orientation*—itself a variable blend of cultural traditions, previous experience, new information, analysis, synthesis, and genetic heritage.[21] This led to *decision*, which enabled one to *act*. As the author has noted elsewhere, this tempo-based approach to conflict and war is a fundamental shift from an attritional approach, and it depends on viewing the opponent as a system of systems.[22]

As Carl von Clausewitz and Napoléon Bonaparte reminded us, moral fac-

tors are the ultimate determinants in war, and this gets at the essence of an overarching theme within Boyd's presentations—the need to drive the opponent or adversary toward a state of disaggregation, what he characterized ultimately as panic and chaos.[23] Boyd termed this moral conflict, and its essence, expressed in *Patterns*, was to, "[s]urface *fear*, *anxiety* and *alienation* in order to generate many non-cooperative centers of gravity as well as subvert those that adversary depends upon thereby magnifying internal friction" in order to "destroy moral bonds that permit an organic whole to exist."[24]

For Boyd, these factors were seemingly even more important than physical effects, because the idea was to overwhelm the adversary's observation and orientation process that makes sense of (or orients to) environmental actions. The multiple thrust idea is inherent in the concept of what Boyd termed *nebenpunkte*, which Frans Osinga defined as taking "a line that threatens alternative objectives . . . distract[ing] the enemy's mind and forces" from the main effort, *schwerpunkte*.[25]

The temporal approach to war seeks to act before the opponent can properly orient in the OODA loop construct because they are forced to *re*-observe, *re*-orient, *re*-decide, and *re*act. Therefore, rapid decision and compounding effects enable initiative, which can be defined as "the impulsive power resulting from timely decision and action, enabling freedom of maneuver while constraining an opponent's options."[26]

Many observers believe that the temporal aspect in Boyd's conception is essential to achieving these effects.[27] However, time may also be conceived as a maneuver space, and going at a higher tempo may not always be the best approach—nor is a higher tempo necessary for seizing the initiative (consider an insurgent, operating at a slower tempo but achieving initiative in their chosen time scale, surviving beyond the conflict—a strategy of exhaustion).[28] Indeed, the most recent doctrinal publication, *Joint Planning*, Joint Publication (JP) 5-0, notes in the discussion of tempo, "on other occasions, JFCs [Joint Force Commanders] may find advantageous to conduct operations at a reduced pace."[29] Ian T. Brown wrote similarly, concluding, "time and tempo were only two of the many factors used against an opponent to render him incapable of activity; one still sought to isolate and neutralize physical and non-physical strengths and moral bonds simultaneously."[30] While Boyd did emphasize tempo, there is a tendency, as Frans Osinga noted, to equate speed with victory, especially with respect to decision making, but he argued for "dispelling the notion that mere information superiority or superior speed in command and control is the essence of the idea."[31]

In the 1981 versions of "Patterns," Boyd explained:

> Impressions . . . we are trying to . . . get inside adversary system and mask own system against his penetration; create a variety of impressions of what is occurring and what is about to occur; generate mismatches between what seems to be and what is; push adversary beyond his ability to adapt.[32]

While tempo is important, it should not be confused with the goal of merely "out-speeding" the opponent.

As noted elsewhere, decision enables initiative, which conveys degrees of control.[33] While he never explicitly defines initiative, he does offer something that sounds strikingly similar without attaching it to the word "initiative": "improve our capacity for independent action . . . diminish the adversary's capacity for independent action, or deny him the opportunity to survive on his terms, or make it impossible for him to survive at all."[34]

Ultimately, the effects spiral works because it operates against the strongest and weakest links in the system simultaneously—what we might consider the human domain.[35] Dr. Jeffrey Reilly has argued that the human domain is what multi- and all-domain actions seek to influence, comprised of leaders, organizations, and populations.[36] This resonates with Clausewitz's emphasis on the moral and emotional aspects of warfare and with Marine Corps thinking about the nature of war. *Warfighting*, Marine Corps Doctrinal Publication 1, expresses the Marine view of war's characteristics as uncertainty, fluidity, friction, disorder, complexity, violence, and danger; the interaction of physical, moral, and mental forces; its constantly evolving nature; and containing a human dimension, previously mentioned.[37]

The OODA Loop as a Complex, Adaptive System

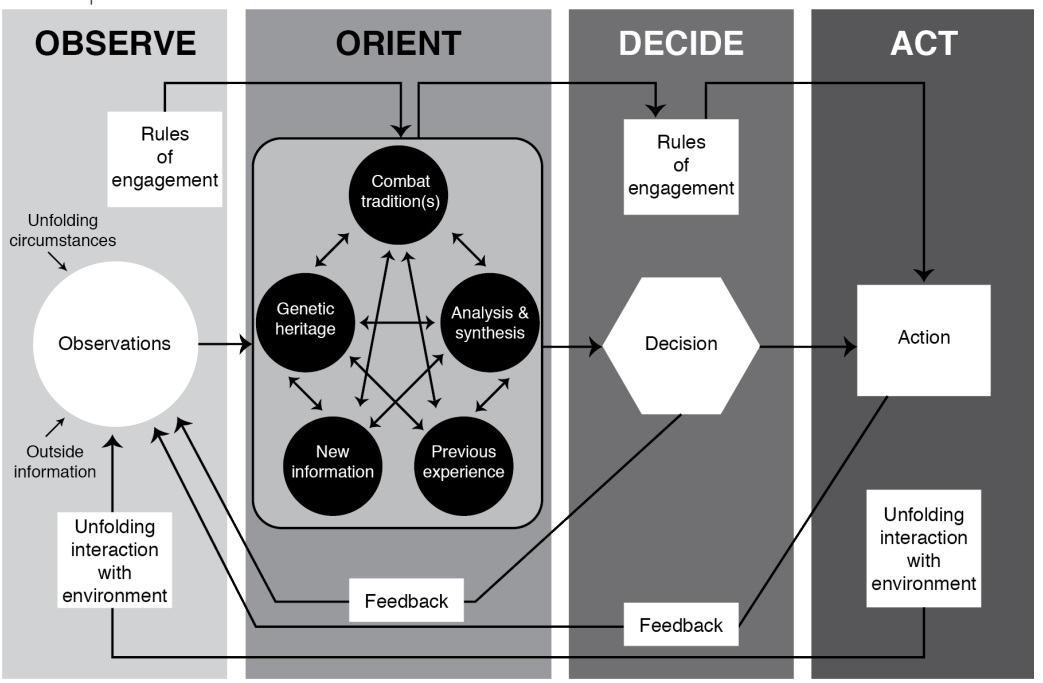
The OODA loop concept underwent significant development during the years of Boyd's writing, progressing from the simple four-step process to a broader view of organizations and organisms.[38] These elements move the OODA loop from a simple heuristic to a more sophisticated model of organic interaction into complex, adaptive systems. As such, they will respond with any kind of interaction, but how they orient and understand will depend on a variety of feedback loops, aspects of orientation, and influences of what and how they observe—in addition to specific efforts at disinformation and deception, which can impact any connection point or points within the system.

Crafting the Environment

In Boyd's 1987 briefing, "Strategic Game of ? and ?," he summarized the crucial importance of his "big theme" for three of his previous projects, which is "one of interaction and isolation."[39]

Boyd saw manipulation of the environment as crucial to victory in states as diverse as full-scale blitzkrieg to guerrilla operations. In conflict, examined in "Patterns," one seeks to create an environment of menace, which he defined as "impressions of danger to one's well-being and survival."[40] *Menace* is the state of being in danger, or the perception of being in danger, which begin the spiral. The Doolittle Raid on Tokyo might be taken as a classic example of an effort to begin the cycle, since it was not the relatively light physical destruction that was the point of the raid, but the psychological effect. It is an atmosphere of menace that energizes the spiral of disruption.

Figure 1. Boyd's expanded OODA loop, based on a 1992 sketch of the loop surviving in the Marine Corps Archives



Note the myriad connections within and between the observation and orientation segments—these are excellent points to inject disruption.
Source: adapted by MCUP.

The idea as expressed in "Patterns" was to create the atmosphere of menace by eroding factors that contribute to cohesion, such as trust and confidence. The safe, comfortable world of the opponent must become amorphous and unpredictable, which creates a fundamental sense of insecurity, anxiety, and menace. In today's multidomain and all-domain approaches, these effects are made, creating tangles across multiple domains simultaneously.[41]

At the same time, he saw that the inverse was true for protecting friendly centers of gravity. In "Organic," he emphasized continual interaction between components, designed to redundantly reassess the changing environment in support of continual adaptation and evolution. It was the connections between elements that was crucial in "Organic," just as it was the connections between components that it was necessary to challenge or sever in "Patterns." Boyd summarized, "the strategic game is one of interaction and isolation."[42]

Character of Modern Warfare and Engagements

Today, the emergence of an effective "reconnaissance-strike complex," blending pervasive surveillance (satellites, ubiquitous sensors, drones) with long-range fires (missiles, stealth, electronic warfare, fifth columns, drones), has created a moment in history when it seems that gathered forces face significant risk—a risk that increases sharply the longer they are gathered.[43]

Clausewitz's conception of a single center of gravity, "the hub of all power and movement," in his famous phrasing, seems to be a characteristic of a foregone era in warfare, at least at the operational level.[44] Indeed, today's COGs (or critical vulnerabilities) are as likely to be informational and economic as they are military and based on force or the threat of force, even at the operational level, because of ubiquitous connectivity and sensors. But this applies even for fielded forces, because of the danger to gathered forces, potently and disastrously experienced by both Taliban fighters and Russian soldiers, and reflected in current Marine Corps efforts to project power into the Pacific area of operations.[45] The characteristics of today's warfare suggest that forces within range of the opponent's long-range fires must be dispersed and mobile, evading detection through the enemy's reconnaissance-strike complex, gathering as briefly as possible to strike with as much speed and secrecy as possible, then dispersing again so long as they remain in range. This suggests that at the operational and tactical levels, multiple centers of gravity exist, a concept Boyd advocated as early as 1989, though he evinced a preference for "vulnerability" over "center of gravity."[46]

Because of their critical nature, centers of gravity are likely to be well defended. As examined above, one can attack the node and/or its connections. Economic and informational nodes may be undefended or defended by means other than the physical. Today's battlefield is connected ubiquitously through military and non-military connections, and without these connections, the disparate elements cannot effectively coordinate. One does not have to destroy; one can use the pallet of defeat mechanisms expressed in *Joint Planning* (destroy, dislocate, disintegrate, isolate, disrupt, degrade, deny, and neutralize) to attack the connections.[47] For example, degrading the quality of a connection by injecting disinformation could begin the spiral and might be well within operational reach, even if destruction of the connection is not—for example, if connection methods are redundant.

Engagements in today's warfare are likely to become what Cyber Command and Special Operations Command term as *persistent*.[48] Unlike the clashing, climactic engagements discussed by Clausewitz and Antoine-Henri de Jomini, much conflict in today's environment takes place across the competition continuum, as the Department of Defense terms it, competition below or hovering just around the threshold of traditional warfighting and dominated by independent economic and informational actors.[49] Because of this persistent engagement by cyber, Special Operations Forces, and informational forces, planting the seeds to be evoked later is common practice, a part of the competition landscape, encouraging commanders and planners to take on a wider vista and a longer-range view of time.

All of this may be seen in the recent Russia-Ukrainian war. Various operational centers of gravity have been identified by both sides that include fielded force concentrations, key leaders, and logistical and C2 hubs, alongside more traditional terrestrial objectives such as cities, bridges, and key terrain features.[50]

Information may be seen as the connective tissue binding centers of gravity within and between Russia and Ukraine, along with their respective allies and supporting partners. Economics might be seen similarly, playing an additional crucial role in terms of stamina and resilience. Persistent presence by both Western SOF forces and advisors such as the California National Guard have forged enduring relationships while preparing Ukraine for conflict.[51] On the Russian side, efforts to drive Ukrainian loyalists out of Crimea and the contested eastern regions, alongside persistent presence of Spetznaz, the Federal Security Service (FSB), Wagner Group, and other entities creates a narrative of its own as well as direct social, cultural, and political effects. [52]

Using the Disruption Spiral

Following Boyd's strategic theory, as expressed across his multiple presentations discussed above, the objective is not to necessarily strike the enemy's centers of gravity directly but to attack the connections between the centers. Indeed, the term *center of gravity* implies a single center, whereas in postmodern warfare any target that may be seen may be struck with lethal and catastrophic force (for example, the Ukrainian HIMARS strike on the Russian barracks in Makiivka, 2022). Postmodern warfare is decentralized, coupled with the ability to rapidly synchronize and gather at the crucial place and time.

At the start of the spectrum, uncertainty can be the minimum effect one might seek in a disruption effort, and it is prevalent in both maneuver and moral style conflicts.[53] On his "Essence of Moral Conflict" slide, Boyd wrote, "*Uncertainty*—impressions or atmosphere generated by events that appear ambiguous, erratic, contradictory, unfamiliar, chaotic, etc."[54] Uncertainty draws from the thinking of Thomas Kuhn on paradigms and Werner Heisenberg with his uncertainty principle; injecting even just a little uncertainty into the opponent's decision cycle may result in a slight delay, which, given the nature of observe/orient, may result in a downward spiral unless counteracted by action taken to retrieve the initiative.[55] In an environment where time compresses toward Dr. Jeffrey Reilly's "OODA point," even a slight delay might be decisive.[56]

In this sense, a suboptimal action may well be better than no action—or the perfect action taken later—since even almost any action might recover the initiative and stop the spiral. Another word for uncertainty might be *ambiguity*, though it is worth noting that entities and individuals have vastly different tolerances for ambiguity, and in a mission-command environment, uncertainty as an effect alone might yield the opposite effect, encouraging creativity, innovation, and freedom of maneuver otherwise constrained by overcentralization.

Doubt is Boyd's next condition on the spiral. Doubt causes an even longer delay in the process of orientation, as information received or previously understood is questioned. In "Patterns," Boyd referred to it as a "moral factor," and he often associated it with fear and anxiety. In the section looking at success factors for blitzkrieg-type operations, he wrote, "broad use of [the] Schwerpunkte concept coupled with fast tempo/fluidity-of-action of armoured teams

and air support permit blitzers to repeatedly reshape strength and rapidly shift it against, or thru [sic], weaknesses thereby generate doubt and uncertainty which magnify into panic and chaos."[57]

Extending further down the spiral, mistrust questions the fidelity of relationships or perceptions of fact. In planning terms, this can be a useful revisit of facts converted from assumptions, but in terms of resistance and cohesion, it acidly chews at the bonds critical for unity of effort. On his moral conflict slide, Boyd wrote, "Atmosphere of doubt and suspicion that loosens human bonds among members of an organic whole or between organic wholes."[58] This does not only apply to fellow humans; it could equally apply to mistrust in key systems or processes. Osinga observed that eroding trust was a crucial element of Boyd's presentation of guerrillas, forcing their opponents to work in "a hostile environment (of menace and uncertainty), which naturally breeds mistrust."[59]

An entity within broken linkages leads to confusion, a state in which information flow is significantly interrupted, requiring a reorientation and realignment of key elements in order to regain cohesiveness. Boyd associated confusion with "contradiction of feeling, indecisiveness, panic," which he arrived at by studying blitzkrieg tactics, Sun Tzu, and guerrilla warfare. Confusion was and is caused not only by fast tempo but *fluidity of action*, challenging further orientation. The object of confusion was to bring about disorder to "shatter cohesion, paralyze effort, and bring about adversary collapse."[60]

Disorder results when confusion multiplies. The overall structure of an organization or organism begins to break down into component parts. Boyd often connected it with confusion in his presentations directly, repeated in the phrase, "confusion and disorder," woven into the spiral with the goal of "present[ing] many (fast-breaking) simultaneous and sequential happenings" that make it hard for the enemy to respond in a "directed fashion." In another instance, he wrote that confusion and disorder, "impedes vigorous or directed activity, hence, by definition, magnifies friction or entropy."[61]

Fear grips individual components when the organization/organism breaks apart, because long-established relationships and stronger bonds yield to growing disruption, and the survival instinct begins to assert itself, eclipsing other concerns. As Osinga notes, Boyd followed J. F. C. Fuller that "a strategist should think in terms of paralyzing, not of killing . . . a man unnerved is a highly infectious carrier of fear, capable of spreading an epidemic of panic."[62] Thus, fear is an accelerant along with the spiral, leading toward panic.

Panic ensues once fear rises to a point where analysis fails to hold disruption at bay and rational thought gives way to raw emotion. A particularly important form of panic is paralysis, the third option in the traditional fight-or-flight conception advanced by Dave Grossman in *On Killing*.[63]

Chaos reigns at the end of the spiral, where an organization (or organism) is fully disaggregated "back inside himself" and there is no cohesive relationship between the parts.[64] Chaos is the opposite of order or law; there should be no corporate will to resist in a state of chaos, though individual components may

still resist as their identity has fully shifted from being a part of the whole to being an individual with fundamental survival instincts.

Using the Spiral of Disruption

One conducts operations in order to fold the adversary "back inside himself" and "maneuver [the] adversary beyond his moral-mental-physical capacity to adapt or endure so that he can neither divine our intentions nor focus his efforts to cope with the unfolding strategic design or related decisive stroke as they penetrate, splinter, isolate or envelop, and overwhelm him."[65] Boyd relates that

> unless such menacing pressure is relieved, [the] adversary will experience various combinations of uncertainty, doubt, confusion, self-deception, indecision, fear, panic, discouragement, despair, etc., which will further:
>
> - Disorient or twist his mental images/impressions of what's happening; thereby
> - Disrupt his mental/physical maneuvers for dealing with such menace; thereby
> - Overload his mental/physical capacity to adapt or endure; thereby
> - Collapse his ability to carry on.[66]

All of this relates fundamentally not so much to observation by the opponent as it does to orientation and overwhelming not just perception but "sense-making," as it is often termed in JADO/JADC2. Orientation, in Boyd's conception, is a product of a variety of influences—cultural traditions, previous experience, and analysis and synthesis, among others. The idea is that the weak points are the connection points in the system, as expressed in Boyd's expanded OODA loop illustration.

Within the JADO construct, the idea is to present the enemy with a "convergence of effects globally, across all domains, to consecutively or simultaneously present an adversary with multiple dilemmas . . . such dilemmas, when presented at an operational tempo that complicates or negates an adversary's response, enable the joint force to operate inside an adversary's decision cycle."[67]

As the director of the Air Force's Joint All-Domain Strategist (JADS) concentration, Dr. Jeffrey Reilly has noted that JADO "recognizes temporarily limited opportunities and deliberately exploits domain interdependencies through access or control of key segments of the domains."[68] This strongly echoes Boyd's intent to attack an adversary's system at the weak points binding elements together, disrupting cohesion and leading to confusion and ultimately disaggregation.

By striking connections, Boyd sought to isolate key elements of the system, as when the Coalition air struck at C2 in the Gulf War, isolating the regime and individual units on the battlefield, which it can be seen clearly launched a

cycle that ended in disorder and chaos. At the very least, initiative was lost to the defending Iraqi forces.[69] Using Boyd's spiral of disruption can be done at all three levels of war: the tactical, operational (what Boyd termed *grand tactical*), and strategic.[70]

Tactically, it is certainly possible to achieve destruction of a whole defender. This tends to be the aim of the direct, battle-centric approach, and it can certainly work. But, following the tenets of Basil Liddell Hart and Sun Tzu, an indirect approach may be less costly (if slower, requiring more patience).[71] Injecting uncertainty and doubt and leveraging Clausewitz's concepts of fog and friction to inject ambiguity may start the spiral of doubt and mistrust that leads to panic/paralysis, disaggregation, and dissolution.[72] The greater the volume of uncertainty, doubt, and mistrust injected simultaneously, the greater the probable rate of slide toward disorder, fear, panic, and chaos.

The larger and more complex an organism, the more likely it will be that resilience will remain with multiple redundant connections connecting key nodes and systems. If the culture enables mission-type command and encourages creative problem solving by educated individuals, resilience will be greatly enhanced. At the operational level, it becomes increasingly difficult and costly to destroy the whole entity, so attacking an entity at its key points of connection may represent a quicker way to leap straight to disorder, injecting more fear, resulting in panic and chaos. From Boyd's perspective, identifying and striking key connections is the best approach while Warden might argue that striking key nodes is a better approach.[73] This strongly implies a multiple COG model would yield the best analysis, rather than a single-center of gravity approach.

In today's world, increasingly, combined forces that multiply joint and all-domain approaches represent a key friendly center of gravity that must be defended. Given that trust is the key bond, the durability of long-term relationships based on shared strategic goals and risk must underpin such relationships and provide "moral strength," in Boyd's conception.[74]

At the strategic level, the scale of an organism suggests that destruction may well be out of reach, short of a nuclear or biological strike—anathema and fortunately likely unavailable. When dealing with an entity on a national scale, multiple redundant pathways and nodes again suggests a multiple center of gravity approach, underscoring the exceedingly difficult task of understanding the opponent at a level sufficient to identify key connection vulnerabilities.

This is especially difficult given the tendency to mirror and project one's own perceptions and conceptions on a thinking opponent. The triangulation of sources of subject matter expertise is one way to mitigate this risk, though it is very hard to do given the pace of most planning teams.

In an alliance, as with combined action, eroding trust is a time-tested way to reduce unity, as with the Iraqi efforts to inject wedges between the Arab states and the West using Scuds in an attempt to draw Israel into the war.[75]

Conclusion

The list of effects proposed by Boyd—uncertainty, doubt, mistrust, confusion, disorder, fear, panic, and chaos—can be used as an effects spiral to create and measure effects in the human domain.[76] For a relatively low cost, one can inject uncertainty or doubt, causing mistrust and confusion and ultimately disorder, fear, panic, and chaos. Through a deliberate combination of physical and informational attacks, each condition leads to the next, serving to progressively isolate centers of gravity so they cannot coordinate and synchronize. In the interlocked and networked modern world, bypassing well-defended centers of strength to strike at an enemy's cohesion through their nodal connections may prove both efficient and effective. And against a peer or near-peer opponent, they may be all that is available in a crisis, or in the case of rapidly eroded fielded forces.

Ultimately, the point on the scale where an individual, an organization, or a leadership team ends up will depend, in part, on the target's resilience, as a function of culture, technology, supply, training, and education when compared to the strength, unexpectedness, and variability of the attack. Reducing the capacity and speed of decision making, or the quality of information available, can be done through both physical and informational means. This approach is valuable at all levels of planning and in all types of conflict, though like all tools, it must be used appropriately; there is no one size or one solution fits all—adaptability and tolerance for ambiguity, however, are crucial. Building an educated force, an energized, flexible set of organizations and individuals that can tolerate ambiguity while maximizing information superiority, is a formula for building a force likely to survive and dominate. Emphasizing human factors alongside technological superiority is a must; neither technology nor human factors can dominate alone in emerging forms of warfare, where physical distances have dramatically increased while the distance between human minds has shrunk.

Postscript: Inoculating against Uncertainty and Doubt

In Boyd's conception, each of the above aspects represents a fundamentally human condition, though each applies also to technological proxies. As human conditions, they are countered by aspects that yield confidence—such as training, a strong esprit de corps, and experience. The West's professionalized armed forces do well with these aspects. But Boyd's spiral also suggests that a clever opponent will try to inject ambiguity and doubt into the cycle—and while coherence expressed through a strong organizational culture (esprit de corps) is helpful in countering it, Western military organizations are not as good at education. Education offers a broader set of adaptable tools and a realization that the world is a much larger place than training normally assumes, yielding tolerance for ambiguity and providing some inoculation of an individual or organization against uncertainty. Education prepares one for the unknown, versus training, which prepares for what is known based on best practices, because education tends to be open-ended and open whereas training is often a closed

system. In Boyd's terms, one cannot survive and adapt without openness and change, because the world constantly changes. Education offers cross-domain knowledge that can be exceedingly useful in developing approaches to wicked problems.[77] To realize the benefits of planning, response, and strategy, Boyd argued:

> By an instinctive see-saw of analysis and synthesis across a variety of domains, or across competing/independent channels of information[, one must] . . . spontaneously generate new mental images or impressions that match up with an unfolding world of uncertainty and change.[78]

Endnotes

1. John R. Boyd, briefing slides, "Organic Conception for Command and Control," 1987, copy in the author's collection, slide 7; available also in John R. Boyd, *A Discourse on Winning and Losing*, ed. and comp. Grant T. Hammond (Maxwell AFB: Air University Press, 2018), 217–44.
2. Charles C. Krulak, "Letter to the Editor," *Inside the Pentagon*, 11 March 1997, author's collection.
3. This is not to imply that agreement with the OODA loop is universal; though widely accepted, challengers include Stephen Robinson, *The Blind Strategist: John Boyd and the American Way of War* (Dunedin, New Zealand: Exile Publishing, 2021). Nonetheless, the OODA model is widely accepted and is a foundational element of the Joint all-domain command and control (JADC2) architecture, as noted in the unclassified architectural graphic and in the U.S. Air Force's "USAF's Operating Concept for Information Warfare," 30 March 2022, author's collection.
4. For the best treatment of Boyd's ideas in their scientific and military context, see Frans P. B. Osinga, *Science, Strategy and War: The Strategic Thinking of John Boyd* (New York: Routledge, 2006); and John Andreas Olsen, ed., *Airpower Reborn: The Strategic Concepts of John Warden and John Boyd* (Annapolis, MD: Naval Institute Press, 2015). For a hagiographic but insightful biography from the perspective of the Defense Reform Movement, see Robert Coram, *Boyd: The Fighter Pilot Who Changed the Art of War* (New York: Little, Brown, 2005). For a less hagiographic but also less detailed treatment, focused more on Boyd's ideas, see Grant T. Hammond, *The Mind of War: John Boyd and American Security* (Washington, DC: Smithsonian, 2001). For a strongly argued critique of Boyd's work, particularly his use of history, see Robinson, *The Blind Strategist*. For the author's own treatment, Brian R. Price, *Eagles, Falcons & Warthogs: Gen. "Bill" Creech, Col. John Boyd and the Struggle to Remake the Tactical Air Forces in the Wake of Vietnam* (Annapolis, MD: Naval Institute Press, forthcoming).
5. For an excellent summary of Boyd's contribution to maneuver warfare, see Ian T. Brown, *A New Conception of War: John Boyd, the U.S. Marines, and Maneuver Warfare* (Quantico, VA: Marine Corps University Press, 2018), https://doi.org/10.56686/9780997317497.
6. For an expansion and an excellent summary, see Frans P. B. Osinga, "The Enemy as a Complex, Adaptive System: John Boyd and Airpower in the Postmodern Era," in *Airpower Reborn: The Strategic Concepts of John Warden and John Boyd, ed. John Andreas Olsen* (Annapolis, MD: Naval Institute Press, 2015), 48–92.
7. Carl von Clausewitz, *On War*, trans. Peter Paret and Michael Howard (1832; repr., Princeton, NJ: Princeton University Press, 1984), 184. "Another reason for not placing moral factors beyond the scope of theory is their relation to all other so-called rules. The effects of physical and psychological factors form an organic whole which, unlike a metal alloy, is inseparable by chemical process." First articulated in the interwar period

by the Air War Plans Division, in the strategic bombing plan for Germany (AWPD-1). See Howard David Laine, "AWPD-1: America's Pre-World War II Plan for Bombing Germany" (master's thesis, Virginia Polytechnic Institute and State University, 1991). For industrial web theory as part of the U.S. Army Air Corps Tactical School's approach, see for example Brian D. Laslie, *Architect of Air Power: General Laurence S. Kuter and the Birth of the U.S. Air Force* (Lexington: University Press of Kentucky, 2017), 32; Scott F. Murray, *The Moral and Ethical Implications of Precision-Guided Munitions* (Maxwell AFB: Air University Press, 2007), 11–13; and Robert A. Pape, *Bombing to Win: Air Power and Coercion in War* (Ithaca, NY: Cornell University Press), 62–63.

8. For Boyd's extensive use of scientific, psychological, and philosophical literature, see especially Osinga, *Science, Strategy, and War*, chaps. 3 and 4, 52–127.
9. John Warden's thinking on targeting developed over time, beginning with discussions on key targets in his 1988 work, *The Air Campaign* (Washington, DC: Potomac Books, 1988), 6, which flows into a discussion of centers of gravity. By the time he was articulating his "five rings" theory, articulated later that year in "Global Strategy Outline" (unpublished, see Olsen below, bibliography), he posited that "simultaneous attacks . . . launched against multiple target-sets within each of the five rings [leadership, C2, infrastructure and industry, population and agriculture, fielded forces], the effect would be exponential," as summarized by John Andreas Olsen in *John Warden and the Renaissance of American Air Power* (Washington, DC: Potomac Books, 2007), 110. As Osinga notes, Boyd sought to "isolate an opponent [or node] and in due course it will lose internal cohesion and external support, its delayed and misinformed reactions will be ineffective, and it will fail to adjust correctly to the changed environment." Osinga, *Science, Strategy, and War*, 72. Recently, Ian T. Brown summarized it this way: "Boyd characterized moral conflict as a style of warfare that sought to deliberately fray or sever those bonds in a way that reduced an opponent to a chaotic assortment of frightened, mistrustful, and isolated individuals [again, one could also read 'nodes']"; and Brown, *A New Conception of War*, 110.
10. John Boyd, "Patterns of Influence," 1989, transcribed in Brown, *A New Conception of War*, 228. One of Brown's superb contributions to those seeking to understand Boyd's ideas is the "Patterns of Conflict" briefing transcription, since only a very few recordings of his briefings are known.
11. The idea of "critical vulnerability" has a specific doctrinal meaning, used in critical factor analysis as a central method to analyze enemy centers of gravity (COGs) in Joint planning doctrine. See *Joint Planning*, Joint Publication 5.0 (Washington, DC: Department of Defense), GL-7, where a critical vulnerability is defined as "an aspect of a critical requirement which is deficient or vulnerable to direct or indirect attack that will create decisive or significant effects." For critical factor analysis, see Dale Eikmeier, "Logical Method of CoG Analysis," *Military Review*, September–October 2007, 62–66, where Eikmeier adds to the doctrinal definition, "The smaller the resources and effort applied and the risk and cost, the better." At least by 1995, as expressed in the Quantico presentation cited in note 10, Boyd preferred the term vulnerability to center of gravity; see the exchange between Boyd and an unidentified colonel in Brown, *A New Conception of War*, 217–18. In fact, doctrinal acceptance of the ideas of critical factor analysis was suggested in 1989 by Capt John Schmitt as author of *Marine Corps Operations*, Marine Corps Doctrinal Publication 1, just published at the time of Boyd's "Patterns of Conflict" seminar at Quantico. Schmitt argued, "we should focus our efforts against a critical enemy vulnerability," a fundamental tenet of maneuver warfare theory as expressed in MCDP-1 in both 1989 and 1997.
12. John Boyd, "Patterns of Conflict," 1989, 213.
13. Osinga, *Science, Strategy, and War*, 1–2.
14. Boyd, "Patterns," slide 175, "The Game is to. . . . Pull adversary apart, produce paralysis, and collapse his will to resist," reprinted in Brown, *A New Conception of War*, 249, cited also multiple times in Osinga, *Science, Strategy, and War.*
15. Boyd, "Organic Conception of Command and Control," in *A Discourse on Winning and Losing*, 217–54.

16. *Nebenpunkte* as a concept appears 15 times in Boyd's September 1981 version of "Patterns of Conflict." In the 1989 "Patterns" briefing transcribed by Ian T. Brown, it appears only once, in the summary on slide 175 (249), though this is only a published excerpt of a full transcription. Osinga explains the concept: "If you take a line of distraction that threatens alternative objectives . . . an idea Boyd was to come to refer to as Nebenpunkte." *Science, Strategy, and War,* 35. This idea forms a core part of conceptions of multi- and all-domain operational concepts. For the U.S. Army, this was expressed clearly in "The U.S. Military in Multi-Domain Operations," TRADOC Pamphlet 525-3-1, 2018. "Long range ground fires complicate enemy defenses by forcing the enemy to react to multiple forms of attack simultaneously against a number of different systems for which it does not have an effective counter," 33. The Air Force's Joint All-Domain Strategist (JADS) concentration approach—roughly equivalent of the Army's SAMS—is built around the synchronization of all-domain operations to achieve simultaneous effects.
17. Osinga, *Science, Strategy, and War*, 159.
18. Boyd, "Patterns of Conflict," 1989, slide 141, 244.
19. Boyd, "Patterns of Conflict," 1989, 215.
20. *Planning*, iv–22. In *Planning*, the importance of COGs is described: "Success requires protecting the friendly COGS while defeating the enemy COG."
21. Boyd's OODA loop sketch, 1996 version from "The Essence of Winning and Losing," 28 June 1995, rev. 1996.
22. Brian R. Price, "In Pursuit of Decision Advantage in JADO/JADC2: Conceptual Background & Definitions," *Aether* (forthcoming).
23. In *On War*, Clausewitz states, "Material activity is never directed against material force alone; it is always aimed simultaneously at the moral forces which give it life, and the two cannot be separated," 137. Additionally, he states, "We must return once more to this subject, already touched upon in Chapter Three of Book Two, since the moral elements are among the most important in war," 184. Napoléon's dictum of "the moral is to the physical three to one" appears in Boyd's presentations, but without specific attribution.
24. Boyd, "Patterns of Conflict," 1981, slide 113. Note: the author has chosen to smooth a few of Boyd's phrasings in terms of adding articles and occasional punctuation.
25. Osinga, *Science, Strategy, and War*, 35 and 154. Boyd defined this clearly in "Patterns of Conflict," 1981: "Employ *Cheng/Nebenpunkte* as basis to repeatedly tie-up, divert, stretch-out, or drain-away adversary attention and strength in order to expose vulnerabilities and weaknesses for decisive stroke(s) by *Ch'i/Schwerpunkt*," 132.
26. Price, *Decision Advantage.*
27. Steven Metz, "The U.S. military will be the first post-modern state combatant, attaining greatly amplified speed and precision by the integration of information technology and development of a system of systems which link together methods of target acquisition, strikes, maneuver, planning, communication and supply . . . time will be the key element. Postmodern militaries will use speed and knowledge to bring the conflict to a quick resolution." Cited in Osinga, *Science, Strategy, and War*, 251.
28. Price, "Time as a Domain," *Over the Horizon* (blog), forthcoming.
29. *Joint Planning*, Joint Publication 5-0 (Washington, DC: Department of Defense, 2020), iv–36, where the full paragraph attempts to address the need for nuance in discussing time: "The tempo of warfare has increased over time as technological advancements and innovative doctrines have been applied to military operations." One may think of the Taliban or Ukrainian defending forces seeking to draw out the tempo of a campaign to seek exhaustion rather than attrition. *Joint Planning* continues: "In many situations, JFCs may find it advantageous to maintain an operational tempo that stretches the capabilities of both friendly and enemy or adversary forces. On other occasions, JFCs may find it advantageous to conduct operations at a reduced pace. During selected phases of a campaign, JFCs could reduce the pace of operations, frustrating enemy or adversary commanders while buying time to build a decisive force or tend to other priorities in the OA such as relief to displaced persons. During other

phases, JFCs could conduct high-tempo operations designed specifically to overwhelm enemy defensive capabilities."

30. Brown, *A New Conception of War*, 144.
31. Osinga, *Science, Strategy, and War*, 177.
32. Boyd, "Patterns of Conflict," 1981, slide 120.
33. Price, *Decision Advantage.*
34. Boyd, "Strategic Game of ? and ?," slide 14.
35. *Warfighting*, Marine Corps Doctrinal Publication 1 (Washington, DC: Headquarters Marine Corps, 1997), presents a less comprehensive view of what the authors term the "human dimension." The concept of a human domain better characterizes the impact of humans, even on a technologically advanced battlefield, as central.
36. Jeffrey M. Reilly, "Multi-Domain Operations" slide, multiple briefings, author's collection. Reilly's conception features the overriding importance of control over the electromagnetic spectrum (EMS), with cyber as a subset of EMS that enables space, air, maritime, and land operations, all working to influence the human domain. The Joint-All Domain Strategist approach taught within the SAMS-like concentration at Air Command & Staff teaches students to recognize and attack "temporally limited opportunities . . . [to] exploit domain interdependencies through access or control to key segments of the domains." *Joint All-Domain Strategist (JADS) Planning Guide* (Maxwell AFB: JADS, 2020), 1. Note how this is different from Boyd's idea of attacking connections between organic elements of the system as a whole—the JADS method focuses on disrupting fielded forces and command, control, communications, computers, intelligence, surveillance, and reconnaissance (C4ISR) to influence leaders, organizations, and populations, a much more operationally focused expression of this closely related idea.
37. *Warfighting*, 5–18. The human dimension appears on 13–14, but it focuses more on the effect of will, noting however that it is central in war. In today's operations, because of ubiquitous connectivity, the author would argue that the human dimension is even more encompassing and crucial, with emotive factors such as narrative and identity overshadowing traditional intellectual factors such as interest calculation and cost-benefit analysis. Therefore, establishing an action's narrative as expressed to and as seen by key audience segments requires deep knowledge of key segments, their beliefs, and how clashes are interacting with other perspectives in real time. This in turn requires knowledge of and access to those same segments, a major challenge for policy makers, planning staffs, commanders, and those charged with what we today term, imprecisely, "information operations." All of this forms part of the human domain, in addition to leaders, organizations, and populations. As any marketer knows, segmenting the market and developing a deep understanding of each segment is the key to success.
38. One needs only compare the very simple four-step articulation of the OODA loop replicated countless times in military and business literature (observe, orient, decide, act) and the far more complex model shown far less frequently, which takes into account the elements of systems theory and complexity in the form of key feedback loops and conceptual lenses through which all actions are viewed, key in the orientation phase. For the fundamental articulation of Boyd's OODA loop, see his 1981 discussions of the loop in "Pattens of Conflict"; for the full model, see Boyd, "A Discourse on Winning and Losing," June 1995, slide 4. Frans Osinga discusses the distinctions between the simple tempo-based OODA loop and the nuanced, uncertainty-seeking approach as his central theme in chapter 7 of *Strategy, Science, and War*, 234–57. See also Daniel H. Abbot, "A History of the OODA Loop," in *The John Boyd Roundtable: Debating Science, Strategy, and War*, ed. Mark Safranski et al. (Ann Arbor, MI: Nimble Books, 2009), 1–5.
39. Boyd, "Strategic Game of ? and ?,", slide 30.
40. Boyd, "Patterns of Conflict," 1981, slides 19, 35, 104, 111, 113–14, 117, 121, esp. 113–14.
41. There are myriad examples, but a good expression may be found in the Air Force's *Department of the Air Force's Role in Joint All-Domain Operations*, Air Force Doctrinal

Publication 3-99 (Maxwell AFB: Department of the Air Force, 2021), "success requires the convergence of effects globally, across all domains, to consecutively or simultaneously present an adversary with multiple dilemmas," 1.

42. Boyd, "Strategic Game of ? and ?," slide 33.
43. Justin Bronk, "The Impact of Reconnaissance Strike Complex on Ground Maneuver," RUSI, 15 June 2021, YouTube video, 35:57.
44. Clausewitz, *On War*, 485. At the strategic level, the will of the population and/or leadership and key organizations is often thought to be the COG, depending on a number of factors, such as government and communications infrastructure. Even Clausewitz noted the conception of an opponent as a kind of organic entity, where cohesion was crucial, "in war as in the world of inanimate matter the effect produced on a center of gravity is determined and limited by the cohesion of the parts," 486.
45. "Airstrikes Target Taliban Location Disclosed in Social Media, Leaving Several Dead," Khaama Press, 18 March 2019; Derek Eaton et al., *Supporting Persistent and Networked Special Operations Forces (SOF) Operations: Insights from Forward-Deployed SOF Personnel* (Santa Monica, CA: Rand, 2017), https://doi.org/10.7249/RR1333; "Persistent Engagement" (CyberCom capabilities and operational approach briefing to the Joint All-Domain Strategist class of 2022, UNCLASSIFIED slide 4, Maxwell AFB); Jeff Schogol, "Russian Soldier Gave Away His Position with Geotagged Social Media Posts," *Task and Purpose*, 3 January 2023; and *Tentative Manual for Expeditionary Advanced Base Operations* (Washington, DC: Headquarters Marine Corps, 2021).
46. "Patterns of Conflict," 1989, 217–18.
47. *Joint Planning*, iv-41–iv-42. Note that the coherency of this model has been challenged by colleagues Ann Mezzell and Wes Hutto, "Bridging the Gap: Military Strategy in Theory and Doctrine," draft, December 2022.
48. Derek Eaton et al., *Supporting Persistent and Networked Special Operations Forces (SOF) Operations*; and "Persistent Engagement."
49. *Joint Operations*, v-1–v-8. This edition of *Joint Operations* builds on the earlier "Joint Concept of Integrated Campaigning" (JCIC), dating from 2018. The 2022 *Joint Operations* integrates much of the JCIC's content and shifts the emphasis on operations for the joint force from contingency planning to campaign planning and globally integrated operations.
50. Olafimihan Oshin, "US Has Helped Ukraine Target Russian Generals: Report," *Hill,* 4 May 2022; "Ukrainian Forces Target Four Russian Concentrations of Troops, Equipment in the South," *Kyiv Independent, 5* January 2023; David Axe, "Ukraine's Missiles Are Blowing Up Russian Supply Hubs—Some Inside Russia," *Forbes*, 27 June 2022; and Yaroslav Trofimov, "Ukraine Targets Key Bridge as It Prepares Counteroffensive in South," *Wall Street Journal*, 20 July 2022.
51. William Leasure, "U.S. Special Operations Forces Train with Ukrainian Counterparts," Department of Defense, press release, 25 July 2017; and Jim Garamone, "Ukraine-California Ties Show Worth of National Guard Program," press release, Department of Defense, 18 March 2022.
52. Paul Waldie, "How Ukrainian Loyalists Are Bearing the Brunt of Russia's Intervention," *Global and Mail*, 6 March 2014; Yevgeny Prigozhin, "Ukrainian Army Reportedly Attacks Wagner Mercenary Base in Occupied Donbas," Yahoo! News, 2 January 2023; Tor Bukkvoll, "Russian Special Operations Forces in Crimea and Donbas," *Parameters* 46, no. 2 (2016): https://doi.org/10.55540/0031-1723.2917; and Serhey Hayday, "Ukrainian Military Carried Out a Precision Strike on Wagner Base at Stadium in Kadiivka, Luhansk, Governor Says," *New Voice of Ukraine*, 10 June 2022.
53. Boyd, "Patterns of Conflict,"1981, slide 103.
54. Boyd, "Patterns of Conflict,"1981, slide 113; and 1987 version, slide 122.
55. For Kuhn's influence on Boyd, see Michael Loh, oral history interview with Brian R. Price, 15 and 31 January 2018, 4–5, author's collection. Gen Loh was a longtime associate of Boyd's. For Kuhn and Heisenberg, see Frans. P. B. Osinga, "John Boyd and Airpower in the Postmodern Era," 56, 64, 74.
56. In Dr. Reilly's conception, the "OODA Point" is the point at which available decision

time becomes so short as to be impossible without the aid of artificial intelligence. Jeffrey M. Reilly, "OODA Point: The Need for an Airman's Approach to Operational Design" (working paper, Air Command & Staff College, Maxwell AFB, November 2020).

57. Boyd, "Patterns of Conflict," 1981, slide 80. As noted in note 5 above, Boyd's use of history was profoundly influenced by the writings of Basil Liddell Hart and the memoirs of World War II German generals, which have been characterized by recent historians as inaccurate. Critics such as Stephen Robinson, writing in *The Blind Strategist: John Boyd and the American Way of War,* challenge Boyd's interpretation of how the Germans succeeded at the operational level, and based primarily on this critique, challenges the OODA loop and the whole of maneuver warfare.
58. Boyd, "Patterns of Conflict," 1981, slide 113.
59. Osinga, *Science, Strategy, and War*, 170.
60. Boyd, "Patterns of Conflict," 1981, slides 90, 113, 126–27.
61. Boyd, "Organic Conception of Command and Control," slide 20.
62. Osinga, *Science, Strategy, and War*, 34.
63. Dave Grossman, *On Killing: The Psychological Cost of Learning to Kill in War and Society* (1995; repr., New York: Open Road Media, 2009), see especially chapter 2.
64. Boyd, "Organic Conception of Command and Control," slide 7.
65. Boyd, "Patterns of Influence," 1981, slide 144.
66. Boyd, "The Strategic Game of ? and ?," slide 44.
67. *The Department of the Air Force's Role in Joint All-Domain Operations*, Air Force Doctrinal Publication/Space Doctrinal Publication 3-99 (Washington, DC: Department of the Air Force, 2021), 1.
68. Jeffrey M. Reilly, "JADO Briefing," slide 4, "Continuum of Domains," October 2020, unclassified, author's collection.
69. Note that the Iraqi example is interesting also because the Iraqis believed they had seized the initiative even using a defensive posture, because they hoped to lure the allied forces into attacking into the teeth of the defenses, weakening the force and allied resolve. In Joint doctrine, the offense is often equated with initiative, but this may represent a flawed understanding of initiative; it should be seen not equivalent to offense, but the power to reduce the opponent's options while maximizing one's own—it is the motive power to make something happen. This can be done on offense or defense. See Price, "Decision Advantage," forthcoming.
70. Boyd, "Patterns of Conflict," 1981, slides 39, 46, 60, 78, 82.
71. Basil Liddell Hart's classic work, *Strategy* (New York: Praeger, 1954). For Sun Tzu, although there are many translations, the author is grateful to colleagues and friends Drs. Robert Kerr and John Minney for alerting me to a phenomenal and nuanced translation, *Sun Zi Art of War: An Illustrated Translation with Asian Perspectives and Insights*, trans. Chow-Hou Wee (Singapore: Pearson, 2003). As is evidenced by the inclusion of both within Boyd's "Patterns of Conflict" briefing slides, it is also widely recognized that he drew a great deal from both.
72. Clausewitz, *On War,* "War is the realm of uncertainty; three quarters of the factors on which action in war is based are wrapped in a fog of greater or lesser uncertainty," 101.
73. See notes 9 and 10.
74. Boyd, "Patterns of Conflict," 1981, slide 111. Boyd "suggests that *moral strength* represents mental capacity to overcome menace and uncertainty."
75. Norman Scharzkopf and Peter Petre, *It Doesn't Take a Hero: The Autobiography of General H. Schwartzkopf* (New York: Bantam, 1992), 416–17; and Richard P. Hallion, *Storm over Iraq: Air Power and the Gulf War* (Washington, DC: Smithsonian Books, 1982), 181–82.
76. Interestingly, this may be even more important in the cyber realm. Current artificial intelligences and machine learning systems learn to deal with the expected; their tolerance for ambiguity varies, but when the circumstances exceed known patterns, responses can range from the unexpected to shut down.
77. See especially David Epstein, *RANGE: Why Generalists Triumph in a Specialized World*

(New York: Riverhead Books, 2019). For a summary and application, see Brian R. Price, "David Epstein's RANGE—Specialization and Cross-Domain Problem Solving: What It May Mean for Education, Training, and Problem-Solving in the Armed Forces," *Over the Horizon* (blog), 31 January 2022.

78. Boyd, "Strategic Game of ? and ?," slide 58.

Future Bioterror and Biowarfare Threats for NATO's Armed Forces until 2030

Dominik Juling

Abstract: The article argues that advances in biotechnology and other transformations of the threat environment will increase the risk for North Atlantic Treaty Organization (NATO) forces of being confronted with a biological, particularly a genetically modified, weapon by 2030.
Keywords: bioweapon; biowarfare; bioterrorism; chemical, biological, radiological, and nuclear; CBRN, future warfare

Introduction

At the beginning of the COVID-19 (coronavirus disease) pandemic, caused by the virus SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2), the dangers posed by biological attacks or the strategic effects of pandemics were discussed in national security debates. Now, one catastrophe follows the next, and the Russian war of aggression dominates the security agenda. In the foreseeable future, however, we will not be able to erase new, natural biological threats from the agenda. For example, the 2022 monkeypox outbreak, with a first outbreak cluster in the United Kingdom, reminds us that smaller outbreaks of transmissible diseases are a constant companion of humanity. Nevertheless, the security dimension of pathogens has fundamentally changed in the twenty-first century. It will change even more in the future.

This article explores the next generation of warfare in terms of biological threats by the year 2030. Because of few precedents in the area of biological warfare or biological terror and the partial look into the future, the article, and especially its target audience and substantive focus, is broad. Because biological

Dominik Juling studies conflict studies at the London School of Economics and Political Science and environmental science at Yale University. Previously, he graduated from the Technical University of Munich in political science with a focus on technology. He has work experience with the German Armed Forces, NATO, and the George C. Marshall European Center for Security Studies. His academic interests are diverse but mainly focus on the interaction of climate change and conflict.

Journal of Advanced Military Studies vol. 14, no. 1
Spring 2023
www.usmcu.edu/mcupress
https://doi.org/10.21140/mcuj.20231401005

threats often involve difficult-to-control spread of germs, North Atlantic Treaty Organization (NATO) forces were chosen as the major threatened group for this article, rather than focusing on the U.S. Marine Corps alone. Consistent with the U.S. Marine Corps' *Force Design 2030* and the *NATO 2030* initiative, the time horizon of 2030 was chosen. The former is a comprehensive modernization and restructuring program for the U.S. Marine Corps within the 2030 time horizon. Key points of the program include modernizing equipment, improving cooperation with the U.S. Navy, adapting tactics and strategy to modern weapons, threats and surveillance technology, and better internal talent management. While the *Force Design 2030* report talks a lot about emerging military technologies and hostile area denial, it does not talk about the possibility of biological methods of area denial and their countermeasures. This article is intended to draw attention to potential threats that must also be considered in the restructuring of the U.S. Marine Corps.[1]

Within the framework of the *NATO 2030* initiative, an innovation and reorientation plan comprising nine proposals, it states that NATO also wants to defend its technological lead in the field of biotechnology. In addition, NATO's new Chemical, Biological, Radiological and Nuclear (CBRN) Defence Policy, which has been in place since 2022, provides a comprehensive overview of NATO's policy on biological threats, but it often remains comparatively vague. This article will help to provide examples and further information on threats.[2]

The threats studied may stem from state actors, nonstate actors, unknown origins, or accidents. Consequently, the research question is: "What are possible future bioterror and biowarfare threats for NATO's Armed Forces by 2030?" While past and current events and examples are used throughout the article, the goal is to identify and broadly assess potential future threats. The hypothesis for the article thus assumes that advances in biotechnology and other transformations of the threat environment will increase the risk for NATO forces of being confronted with a biological, particularly a genetically modified, weapon by 2030. The article will show how and why the author comes to this conclusion. In doing so, the article will attempt to demonstrate that future biological threats by 2030 pose a serious but underestimated threat to NATO.

To provide an entry point and broad overview of the topic, the article provides a short history of biowarfare and bioterrorism and discusses the future biological threat environment, influential megatrends, emerging and disruptive technologies, possible biological threats by 2030, current and future means of delivery, and possible actors. It is argued that the threat from deliberately deployed biological agents will increase and change in nature by 2030. Unlike, for example, chemical weapons, biological weapons have not been tactically or strategically usable against humans because of their potentially uncontrolled spread, even to unprotected friendly forces, coupled with their highly complex production and stabilization outside of laboratory conditions. However, advances in biotechnology in modifying existing pathogens and creating entirely new ones now make it possible to circumvent these previous barriers and

produce limited biological weapons for the first time. At the same time, it has already become cheaper, easier, and safer to produce dangerous agents, even more so by 2030. New technologies are also helping to deliver biological agents more effectively. In dual-use terms, by 2030 numerous civilian biotechnological successes will create a vast array of possibly ill-intentioned weapons that will provide NATO's hostile actors with a wide range of methods.

History of Events and Developments Involving Potential Biological Weapons until 2022

As early as 1932, Japan engaged in a massive biological weapons program that resulted in the deaths of at least 10,000 prisoners of war by 1945. It is estimated that more than 200,000 additional civilians and soldiers were killed by Japanese biological weapons during military field operations. Various pathogens and means of delivery were systematically studied. After the end of the Second World War, further nonlethal experiments with biological weapons were conducted by the United States.[3] Particularly noteworthy are the results of a series of ethically highly controversial experiments on unknowing civilians in America. At that time, about 800,000 people in San Francisco were infected with a harmless bacterium. A ship was used to disperse the organisms in the air, but a dispatch with airplanes is also known. In secret tests in the New York City subway, there were even more estimated infections with the harmless bacteria noted. Light bulbs filled with microbes were thrown onto the tracks to distribute the bacteria. At least 239 known tests were conducted between 1949 and 1969, demonstrating the potentially massive spread of deliberately released bacteria.[4] The Soviet Union had a similarly comprehensive biological weapons program. In 1979, four years after the Biological Weapons Convention came into force, there was a very serious accident involving anthrax spores in a laboratory in what is known today as Yekaterinburg, Russia. Due to a missing filter, the area around the laboratory was contaminated and at least 66 people died.[5] Based on testimony from high-level former employees of the Soviet Biopreparat Research Agency, it can be inferred that the Soviet Union worked intensively to develop, mass produce, and test delivery methods of highly lethal biological weapons. Strains were repeatedly modified and improved. The goal was to create weapons that avoided precautionary measures or aftertreatment and were effective quickly and lethally.[6]

The first significant attack in modern history using bioweapons and defined as terroristic occurred in 1984, when followers of cult leader Bhagwan Shree Rajneesh infected 751 citizens of The Dalles, Oregon, with salmonella. Forty-five people were hospitalized. The precipitator was the sect's intention to gain seats in the local county circuit court.[7] In 1990, another cult began more comprehensive attempts to use biological weapons. Professor Barry Kellman reports on Aum Shinrikyo:

> In April 1990, Aum attempted to attack the Japanese parliament with botulinum toxin aerosol. In 1992, Aum sent a mission to Zaire to assist

> in the treatment of the Ebola virus disease victims in order to find a sample of the Ebola strain to take back to Japan for culturing purposes. In June 1993, the cult tried to release poison at the wedding of the Japanese crown prince. Later that month, Aum attempted to spray anthrax spores from the roof of a building in Tokyo. All these attacks were unsuccessful and resulted in no casualties.[8]

Even though the cult's chemical weapons program proved to be deadlier, a well-equipped laboratory was found with various biological substances that were used to successfully cultivate bacteria and viruses.[9]

Since the World Health Organization (WHO) announced the eradication of smallpox in 1980, there has been a debate about whether the last remaining virus strains in laboratories should be destroyed. There has also been much discussion of the possibility of terrorist use, as humanity has become very vulnerable following the suspension of vaccination.[10] At present, the United States and Russia still have small stocks of smallpox strains, which are kept in highly secure laboratories. According to the WHO, no other laboratory has official access to the virus.[11] However, since the attacks of 11 September 2001 (9/11), the general debate on chemical, biological, radiological, and nuclear (CBRN) weapons has been broadened again to include other pathogens. This was also strongly reinforced by the anthrax letters sent only a week after the devastating al-Qaeda attacks. Of the 22 infected, 5 died. The perpetrator was, according to an FBI investigation, a professional Army biological researcher with access to all the essential materials.[12] Also in 2001, the book *Germs: Biological Weapons and America's Secret War* was published only a few weeks after 9/11 and remained at number one on the *New York Times* bestseller list for more than two weeks. It contained a number of investigative novelties about the United States' biodefense projects.

After 2001, it became known that al-Qaeda had already been pursuing a practical bioweapons program since the beginning of 1998. In 1999, the terrorist group recruited a Pakistani biologist to develop biological weapons in a laboratory in Kandahar. In 2001, a biochemist from the al-Qaeda network may have been able to isolate a lethal anthrax strain.[13] The actual progress of al-Qaeda's anthrax research was more advanced than global leaders suspected, but the group was never able to produce a viable bioweapon.[14]

In 2003, there was the first case of letters filled with ricin toxin in the United States. The perpetrator is unknown still today. Ricin toxin is a plant material, so there is no infection and reproduction as with microbes. Al-Qaeda terror cells in Great Britain, Spain, Italy, Turkey, Sweden, and Germany were also planning attacks with ricin toxin in 2003. Suspects were arrested in Great Britain, Spain, Italy, and France.[15] In 2004, ricin toxin contamination was detected in a building in Washington, DC. Until 2009, this was the last major incident involving material that could be used as a biological weapon, with a potential terrorist background.

Since then, there have been a number of incidents up to 2021 due to the relatively easy production of ricin toxin. Most of the recorded cases have occurred in the United States. The lethality of ricin toxin is illustrated by the example of Bulgarian dissident Georgi Markov, who was killed in an assassination in London in 1989 by only 0.2 milligram of the agent.[16] Significant incidents since 2009 include ricin letters sent to American politicians in 2013, ricin toxin in the hands of a right-wing militia in the United States, attempted orders via the darknet, and possession of ricin toxin in 2018 and ricin-powder-filled letters again in 2020.[17] The darknet is a variety of networks that are shielded or hidden from public access. The attempt by a jihadist living in Germany in 2018 to carry out an attack with ricin toxin stands out, as he was believed to have had contact with members of Islamic State and managed to produce potentially lethal ricin toxin on his own. He followed internet tutorials on how to make explosives and extract ricin toxin with rudimentary resources.[18] But also, in Iraq and Syria, the Islamic State tried to obtain functioning biological weapons. A laptop discovered in Syria in 2014 contained many different instructions for the construction, storage, and delivery of weapons of mass destruction.[19] However, the Islamic State's focus seemed to be on chemical weapons, especially after 2014.

A study by the U.S. National Consortium for the Study of Terrorism and Responses to Terrorism that was examining 74 nonstate actor incidents involving biological agents from 1990–2011 concludes that use of an agent, possession of a nonweaponized agent, and attempted acquisition are the most common events. Other categories not recorded as often include plot, interest, possession of a weapon, threat with possession, and attempted use of an agent. The most common types of perpetrators involved in attacks during the period studied are cults and lone actors.[20]

As in many other areas, the ongoing COVID-19 pandemic is also a turning point in the field of bioweapons. Since 2020, there have been a number of different scientific papers examining the link between COVID-19 and terrorism. Experts at University College London's Jill Dando Institute of Security and Crime Science found evidence as early as May 2020 that extremist groups were calling for the virus to be deliberately spread and to infect religious or ethnic groups particularly deemed adverse. Likewise, conspiracy theory narratives that SARS-CoV-2 was designed as a biological weapon became established.[21] The deliberate spread of SARS-CoV-2 was particularly discussed by parts of the American neo-Nazi scene, who set their sights on a violent collapse of the current system to establish a White ethno-state afterward. In right-wing Telegram channels, for example, the door handles of non-Whites, Jews, or FBI facilities were indicated as targets for the application of infectious saliva. Initially, the approach was also discussed in jihadist circles, as the Western states were most affected toward the beginning of the pandemic. In April 2020, an alleged Islamist was arrested in Tunisia for planning to deliberately spread SARS-CoV-2 among local security forces. In addition, many experts agree that COVID-19

has served as a great inspiration for various groups of different orientations that have already considered researching or acquiring biological weapons.[22] Various religious groups of different faiths see COVID-19 as a kind of revenge of God, without actively wanting to contribute to its spread.[23]

In summary, it can be said that, as with chemical weapons, the procurement or attempted procurement of dual-use equipment, which could potentially be used for biological weapons production, has increasingly shifted to the internet since 2009. Here too, in addition to the regular online shops, the so-called darknet is once again playing a prominent role. As a relatively easy-to-obtain toxin, ricin toxin has played an increasingly important role since 2009, and the motivations of nonstate actors have generally been diversified. However, ricin is more suitable for attacking individuals or small groups, since a large-scale attack in the open is logistically difficult and would not be very effective. A major attack with biological weapons predicted by some analysts before 2010 was not realized until the end of 2022. Effective weaponization of SARS-CoV-2 has been partially attempted, but it has not been measurably successful, as all attempts were under primitive conditions.

Warnings about antibiotic-resistant bacteria, vaccine resistant viruses, and the creation of completely new pathogens (chimeras) are also not new and were already voiced, for example, by the authors Tom Mangold and Jeff Goldberg in 1999. In their 1999 prediction, it will take about 20 years before genetic engineering can completely circumvent current biological countermeasures.[24]

The World in 2030

Clearly, the environment for an analysis of biological threats will be different by the year 2030. The author does not attempt to draw a coherent picture of the security world of the future, but rather to identify some factors that are important for the future biological threat environment. One is the overall geopolitical evolution of NATO's relationships with other state and nonstate actors. In a more cooperative world, the role of new treaties and their compliance in dual-use research and biological agents is an important variable of the future. In this context, the future monitoring and prevention of proliferation of pathogens for production and distribution is also an important factor. Another relevant factor is the political stability of countries with significant biotechnology research laboratories and stockpiles of potent pathogens. In the event of insufficient protection of the facilities or political unrest and upheaval, the hazardous materials could fall into the wrong hands.

Other factors are additional natural pandemics through 2030 and the long-term effects of COVID-19 on future strategic considerations within NATO, its member states, and among potentially hostile actors. The consequences of Russia's war of aggression, the following build-up of capabilities, shifts in foreign policy paradigms in some NATO countries, and a potentially more uncooperative international order will also play into the future of a biological threat environment. Add to this a huge number of potential black swan events, ranging

from doomsday cults to false flag attacks to extortionist criminal groups. Equally unpredictable, of course, are future conflicts and their associated events. The next section discusses a number of megatrends that, unlike the variables identified in this article, have already begun in the past and will continue to have a relatively reliable impact through 2030 and beyond.

Megatrends through 2030

Climate change as a megatrend through 2030 is having a significant impact on future biological threats. It has long been known that climate change will lead to a further geographic spread, as well as a net increase in transmissions of infectious diseases.[25] The Euro-Atlantic area in particular will be affected by new species emigrating from the south. The deliberate introduction of already found pathogens or vectors to new habitats farther north might be a terrorist method, made possible in part by climate change. At the same time, permafrost is thawing in many places, revealing frozen pathogens that might not be present today. For example, a child died in Siberia in 2016 from anthrax that was frozen in the permafrost, but smallpox and dangerous influenza strains can also potentially thaw in the Arctic region and be transmitted to humans. Similarly dangerous are much older and completely unknown pathogens that are buried several meters deep in the soil and could come to the surface by 2030.[26] Terrorist use is unlikely but not impossible. An additional factor, accelerated by climate change, is that in many cases natural disasters are followed by infectious disease outbreaks and epidemics. This is mainly due to displacement, which is mostly negatively connected to the availability of safe water and sanitation facilities, the degree of crowding, and the availability of health care services.[27] Another impact is that due to the decrease of global animal and plant biodiversity, large populations from one species potentially have advantages in dispersal in an imbalanced manner. Thus, insects and vectors used as bioweapons can more effectively attack plants, humans, and animals while transmitting and reproducing diseases.

Another set of megatrends such as population growth, migration, urbanization, and demographic change also interact with biological threats to NATO forces through 2030. Poor sanitary conditions in densely populated and rapidly growing megacities make the spread of pathogens more likely. NATO nations are experiencing steady demographic change that includes a rapidly growing older segment of society that is more vulnerable to many transmittable diseases.

Due to ongoing globalization and worldwide trade, especially online, it can be assumed that it will continue to be possible to order and deliver laboratory and medical equipment online through 2030. Similarly, pathogens can spread rapidly and potentially undetected in a short time due to the long-distance transport of people and animals.

The next megatrends identified by the author are inequality and poverty. However, meat consumption has often risen as a result of the greatly increased standard of living in China, for example. While total meat production in other

parts of the world has increased only slightly since 1990, the amount in Asia has doubled. But individual consumption has also risen sharply in China and Brazil since 1990, while individual consumption in many NATO member states has declined slightly since around 2010.[28] It should be noted that there is a clear link between infectious diseases and meat production.[29] In particular, inadequate hygiene and safety measures, as well as factory farming, contribute to new zoonotic viruses and epidemics.[30] Due to various reasons, including high meat consumption, experts suspect that several and more severe pandemics will follow in the future.[31] However, a significant decrease in global meat consumption is not expected. In addition, more meat consumption significantly increases greenhouse gas emissions, which in turn increases biological hazards associated with climate change. Local poverty and inadequate government resources will continue to contribute to the inability to contain and prevent local outbreaks of infectious diseases in a timely manner through 2030, potentially posing a threat to nations far away.

The next megatrend through 2030 is briefly discussed in terms of digitalization and technological advances. As described in more detail in the next section, advances in biotechnology and medicine, as well as in the field of bioinformatics, are already contributing to major breakthroughs in the manipulation of bacteria, viruses, and animals. Bioinformatics is an interdisciplinary science that uses computer-assisted methods to try to generate new findings in the fields of biotechnology and medicine. This trend is very likely to continue by 2030 and further breakthroughs may be recorded. In addition, the advanced methods already known today for manipulating and producing pathogens are expected to become cheaper, easier to use, and possibly more widespread by 2030. This depends on whether there will be stronger regulations in this area in the future. However, it is very likely that civilian research and genome databases with potent pathogens that are freely available on the internet will be expanded by 2030 and could still be misused. The internet also facilitates recruitment and communication between nonstate actors hostile to NATO. Just as today, by 2030 the internet will likely make it possible to communicate encouragement and support for the development or terrorist deployment of bioweapons regardless of location.

The final megatrend cluster identified by the author is hybridization and asymmetric warfare. Both trends pose a certain threat in a world in 2030 in which limited-use biological weapons can wreak havoc on the enemy, but not on the enemy's own forces. In addition, there is the possibility of concealing the origin of, for example, a local epidemic or the possibility of biological weapons that are not lethal to humans. In a hybrid conflict, an adversary actor could, for example, also want to cause economic damage or supply shortages and target livestock populations or agriculture. In a hybrid conflict, it would also be possible to use pathogens against NATO forces to incapacitate soldiers for a longer period of time without causing them permanent harm. In a possible future asymmetric conflict between now and 2030, it must be expected that

the facilitated production and delivery of limited biological warfare agents will allow a heavily outnumbered actor to pretend that it has the ability to establish a certain balance against a perceived superior adversary.

Overall, for the complex 2030 threat environment, a broad set of important variables and longer-lasting megatrends suggest that there are several indications that by 2030 the threat of deployment may be higher and the impact more severe. In the next section, special attention is given to emerging and disruptive technologies through 2030 that are important for the design, production, and delivery of potential biological weapons.

Emerging and Disruptive Technologies until 2030

This section of the article will outline how new technologies are having a major impact on biological weapons by 2030. Before analyzing specific technologies in more detail, however, the author first wants to point out that biological weapons not only have a purely military use, but also, like other weapons of mass destruction, have a particular impact on politics and society. With a large number of digital devices connected to the internet, online media, and the peculiarities of social networks, actors could use the threat or deployment of biological weapons to spread panic and fear. Allison E. Betus, Michael K. Jablonski, and Anthony F. Lemieux examine the important role of media in our increasingly digitalized world as follows:

> Violent acts initiate media coverage, as well as word-of-mouth transmission, functioning as a gateway that draws attention to the terror group and its messages in a manner that increases the salience of the communication; then media provides additional information contextualizing the original act. Media coverage may make the group initiating the communication look more dangerous or powerful than is warranted.[32]

It is thus becoming increasingly clear that CBRN threats are not only reflected in new hardware, but also increasingly affect the virtual information and communication space, as well as the public perception of a real or perceived threat.

A research paper by the NATO Centre of Excellence Defence Against Terrorism identifies a countervailing mechanism for the interaction of terrorism and technological progress. In general, military and civilian innovations influence each other with a reciprocal push and pull mechanism. This also benefits nonstate actors, who usually focus on adapting and refining existing and proven dual-use technology for their own purposes.[33] In addition to easy obtainable dual-use goods, high-tech equipment and material is mostly stolen from professional armed forces, bought on the black market, or supplied by state actors. In *NATO Strategic Foresight Analysis: 2017 Report*, one of six chapters is devoted exclusively to future technologies. The report describes, among other things, the rate of technological advances, the number of individuals with access to the internet, the potential of adversary non-state actors' access to new technologies,

the international interconnectedness, the amount of data collected, and an increase in the number of sensors in the world. At the same time, it is becoming more difficult for states, international organizations, or other frameworks to effectively regulate potentially dangerous technologies. This is due, among other things, to the rise of dual-use devices, effects of globalization, an increase in the power of the commercial sector, and the rapid pace of market maturity of new technologies, where democratic mechanisms can often be slow to react.[34]

The first tangible technologies under consideration are user friendly AI applications and web scrapers, which can already easily search large amounts of information about a certain online topic on the internet or in a database, for example about pathogens. AI can then theoretically analyze or even interpret the results. If no powerful computer hardware is available, capacity can be rented via cloud services. This intersection could well be classified as digital dual-use. The consequence is that gene combinations can be tested on the computer before they are cultivated. This saves time and resources and can be used to develop pathogens with specific properties. The process of producing a large number of molecules by combining and varying different chemical components using modern methods also exists in chemistry.

One of the most important future technologies described in this article are modern biological applications. These include genetic engineering, synthetic biology, and biochemistry. Again, this is an area of dual-use research. Genetic engineering is the direct genome manipulation of organisms, including clustered regularly interspaced short palindromic repeats (CRISPR) gene editing that is probably one of the most important scientific breakthroughs of recent times. Especially in the field of biological weapons and nonstate actors, this is a method that can be misused with serious consequences. The special advantage is that, compared to prior methods, it provides easier, cheaper, and more precise additions or removal of parts of the genome while the organism is alive. Thus, in the future, it will be reasonably easy to turn bacteria, viruses, fungi, plants, and humans into genetically modified organisms.[35] In general, this field is well researched and there are many publications available, as vaccines, for example, are also being developed using similar methods. For instance, a research paper on the synthesis of horsepox was published in 2017. Dr. Tom Inglesby, director of the Center for Health Security at the Johns Hopkins Bloomberg School of Public Health, sees this as increasing the risk of smallpox synthesis.[36] In the future, it is believed that despite often grave ethical concerns and attempted political regulation, research will continue to advance. It is often difficult to regulate and identify dual-use applications early enough. However, strategic considerations and scientific great-power competition also play into this technology, as China, in particular, has recently become known for advances in genetic engineering, which are often seen as ethically critical.[37]

One of the many different aims of synthetic biology is to produce synthetic cells (i.e., synthetic life). In 2019, a synthetic bacterium was created for the first time from an artificial sequence of genomes.[38] In this way, even very dangerous

bacteria could theoretically be created as if from a construction kit. Research is currently being done on this with the aim of producing a synthetic drug delivery platform.[39] However, viruses can also be transported and distributed by synthetic bacteria. Advances in synthetic virology are particularly relevant to this study. In the future, it is expected that any virus whose DNA/RNA (deoxyribonucleic/ribonucleic acid) is available can potentially be reverse engineered, bringing viruses that have been eradicated back into circulation. Currently, the National Library of Medicine has a large database called the National Center for Biotechnology Information Virus (NCBI Virus), which contains the genetic data of nearly all known viruses, as well as other microorganisms and mammals.[40] There is an important report by the U.S. National Academy of Sciences, commissioned by the U.S. Department of Defense in 2018, which describes three particularly dangerous scenarios of synthetic biology. In addition to the already described technique of reproducing viruses with genetic code from the internet, it also mentions the possibility of making bacteria resistant to antibiotics and the possibility of programming microbes in such a way that they slowly poison people through their metabolism. The last method could lead to death after a long time and thus disguise the crime. Much more difficult to implement, but theoretically possible, is a so-called gene drive that automatically spreads through the population, altering people's DNA.[41]

The field of biochemistry is also important, as research into, for example, metabolism processes in cells, signal molecules, or enzymes must also be considered in the effect of biological weapons. The exact impact of this area of research up to 2030 cannot be forecasted precisely, but it is certain that the impact will be significant.

A new development that could potentially have an impact on chemical and biological weapons is microreactors in the form of a continuous flow reactor. Fundamentally, the idea is to allow chemical reactions to take place in a very small device. Advantages compared to large reactors include scalability, on-site and on-demand production, as well as a high reaction yield.[42] The small reactors can be scaled up to almost any size, and expensive, large, and complicated synthesis facilities in batch reactor design are no longer necessary, as the cult Aum Shinrikyo once built them. A 2013 study, however, stresses that the use of microreactors for the production of chemical weapons is limited. Nevertheless, future technological advances may well enable a broader range of warfare agents.[43] Advances in micro-enzymatic reactors are also expected in the field of biology.[44] This could help future terrorists or state actors to produce small quantities of toxic agents in almost any place in the world without significantly putting themselves at risk during production. Although the implications are not yet well understood, the cultivation of pathogens could also benefit from the technology.

Current and future often dual-use developments in nanoscience also offer many overlaps with biological weapons and means of delivery. But not only potentially lethal applications are being developed; nanoscience also supports

modern material sciences, engineering, and production. For some years now, several armed forces have been researching machines frequently called nanobots. However, this often refers to insect-size unmanned aerial vehicles (UAVs), which does not correspond to the "nano" definition. Nevertheless, these bionic insects, which are often only 2–3 mm in size and are capable of flying, can, for example, deliver a highly potent poison unnoticed to many locations.[45] In a swarm, technical systems could be manipulated, disrupted, or destroyed. However, real nanobots (i.e., nano-size synthetic drug carriers) are also not unlikely in the future. For example, a group of Chinese researchers undertook the first successful tests for targeted tumor treatment in 2018.[46] On the other hand, such carriers could also be used for the targeted transport of viruses and toxins. Bacteria have been used as drug carriers for similar applications for some time now. Theoretically, however, it is also possible to manipulate unmodified or transgenic insects with the help of nanotechnology, for example to increase the effect of distributed biological warfare agents.[47]

Other applications of nanotechnologies are very small computers, which will be important for small means of delivery and monitoring of production of biological and chemical warfare agents.[48] In general, by 2030, nano-size technologies are expected to make the dual-use laboratory equipment needed for biological weapon production, among other things, cheaper, more effective, smaller, and more flexible.[49] In addition, future attacks with nanotubes may offer entirely new possibilities for disguising origin and lethality. A researcher at American University explains: "For example, nanotubes could be used to deliver only the lethal parts of the anthrax virus—without the signature protein that is recognizable to the immune system." The researcher identifies three main dangers in linking nanoscience and potential biological weapons. First, rudimentary nanotechnology labs are already available on the internet for under $500 USD. Second, the technology makes it easier and cheaper to produce, disguise, and transport biological warfare agents. And third, the technology is not sufficiently regulated, which could lead to an asymmetric arms race that threatens the overall strategic security of major countries.[50]

The dual-use problem in the CBRN sector, which has already been mentioned several times in this article, has been recognized for some time. For this reason, the informal multilateral export control regime known as the "Australia Group" has been in existence since 1985. It deals with dual-use technologies, which can be misused for the production of chemical and biological weapons, among other applications. The NATO countries and the European Commission are members, but Russia and China, for example, are not, which makes international control much more difficult. Nevertheless, the group offers expertise in identifying potential dual-use applications. Additionally, after 11 September 2001, there were great efforts to provide weaponizable research with guidelines and, in some cases, regulations. For example, after a research report on the synthetic production of a polio virus was published in 2002, the U.S. government set up a high-level advisory body to draw up guidelines against the terrorist use

of biological research.[51] *Biotechnology Research in an Age of Terrorism*, a comprehensive work on the state of the art at that time, was published in 2004.[52] In 2012, the book *Innovation, Dual Use, and Security* was published, in which, in addition to the biological risks, attention was also drawn to the potential chemical risks. It contains a 300-page in-depth overview of many intersecting issues.[53] In 2016, a case came to light in which a Chinese company exported a synthetic opioid called carfentanil unregulated to countries in the Euro-Atlantic area. However, this chemical is so potent that it has already killed several unknowing drug users. Terrorist use could not be ruled out.[54] This incident is exemplary for several substances and devices. Furthermore, on the Chinese state level, there have been concerns from some NATO member states in recent years. The country is pursuing civil-military integration in many scientific fields, often resulting in dual-use goods.[55] In 2021, the United States accused China of not clearly distancing itself from weaponizable research in the biological field:

> China continues to develop its biotechnology infrastructure and pursue scientific cooperation with countries of concern. Available information on studies from researchers at Chinese military medical institutions often identifies biological activities of a possibly anomalous nature since presentations discuss identifying, characterizing and testing numerous toxins with potential Dual Use applications.[56]

Other countries that the United States accuses of a possible dual-use biological weapons program are North Korea and Iran. Russia is accused of not having properly destroyed "BW items specified under Article 1 of its past BW program."[57] An increase in civil-military dual-use research in the CBRN field poses the risk of openly available knowledge being misused for malicious purposes. The next section will take a closer look at the actual level of research in 2023 and what developments are possible by 2030.

Possible Biological Threats by 2030

Without question, the biological threats of the future are increasingly severe. The individual threats are often incomprehensible for nonexperts, as biological warfare can be carried out by using viruses, bacteria, fungi, insects, or plants. Almost all animals are possible vectors, and in the far future, even mechanical products or highly manipulated organisms could also be possible vectors. In addition, synthetic biology, nanotechnology, and DNA manipulation open up a whole new range of possibilities for modifying or even completely rebuilding or recreating viruses and bacteria. The latter are called designer pathogens. These technological advances were foreseeable for some time, and yet they only came to public attention because of the global pandemic. But as complex and diverse as the possible types of biological weapons are, so are the techniques to enhance the efficacy of biological weapons through biological engineering. A 2013 report in the *Dartmouth Undergraduate Journal of Science* lists the possible techniques for weaponizing biological materials. These include the manip-

ulation of bacteria; the aforementioned designer pathogens; the destruction or replacement of individual genes in the context of misused gene therapy; stealth viruses that only unfold their effect in the body after external or internal activation; host swapping diseases that, for example, specifically jump from domestic cats to humans; designer diseases that, for example, cause artificial cancer; and personalized biological weapons. The latter spread approximately asymptomatically in the population and only have an effect on certain genetic characteristics of a person or group of people.[58]

In his 2002 contribution to *The Counterproliferation Papers* of the U.S. Air Force Counterproliferation Center at Air University, Michael J. Ainscough describes the threats that could become reality by 2030. Based on findings of the JASON Defense Advisory Panel in 1997, Ainscough describes six future threats. First, he talks about binary biological weapons that can be used for extortion or safe handling. For this, a harmless host bacterium and a virulent plasmid would be isolated separately and threatened with the release of the associated second component, which would then interact to produce its effect. As far as designer genes are concerned, the researcher concludes that these have long been state of the art with simple modifications at the time of the study. Future designer pathogens will have far more complex capabilities and will be able to exhibit a whole range of modified characteristics. Regarding gene therapy, he writes:

> There are two general classes of gene therapy: germ-cell line (reproductive) and somatic cell line (therapeutic). Changes in DNA in germ cells would be inherited by future generations. Changes in DNA of somatic cells would affect only the individual and could not be passed on to descendants. Manipulation of somatic cells is subject to less ethical scrutiny than manipulation of germ cells.[59]

Already 25 years ago, viruses were used as vectors to insert genes into mammalian cells. This genetically engineered virus was successfully used to prevent rabies in wildlife. Likewise, viruses were successfully used as vectors for mousepox viruses 25 years ago. This allowed vaccination of mice to be circumvented, which died shortly afterwards. The concept of stealth viruses is not new in nature. In this case, an initially unnoticed virus could enter human cells and wait for an external or internal signal. One related example are oncogenes, which are mutated genes that cause cancer as soon as they are activated. Some viruses have segments of DNA that mimic oncogenes. Other substances, bioregulators, physical processes, or external influences such as ultraviolet light could thus activate the virus. Ainscough also writes about host-swapping diseases and designer diseases. In the future of 2030, it could be possible to create the suitable pathogens for a certain disease pattern. This would make it possible, for example, to temporarily shut down the immune system or induce cell death in certain cells.[60] Twenty years later, Ainscough's prognoses are all proving to be increasingly technically feasible. Except for complex designer pathogens and diseases, all predictions are applicable in the year 2023.

Although some of the possible applications mentioned have not yet been achieved in practice, thanks to the aforementioned CRISPR-Cas9 gene-editing technology and the general progress in the field, it is only a matter of time before the biological weapons mentioned are successfully tested within military or civil dual-use research. Another extremely problematic aspect is that CRISPR is not a high-tech technology that is only available in secure laboratories. At the current rate, it is foreseeable that in the world of 2030, manipulated and synthetic biological substances could take on an almost everyday character. But how difficult is it really for future actors to actually develop and deploy one of these methods themselves?

Based on the state of the art in 2015, researcher Zian Liu of the University of California, Berkeley, concludes that there are five potential barriers that could prevent nonstate actors without access to professional laboratories from creating novel biological weapons. First, it is not easy to create a properly protective research environment that will secure the actor adequately. Secondly, although it is possible to order all the necessary materials on the internet, very specialized equipment for very dangerous substances and many test runs cost up to $30,000 USD. If an already dangerous bacteria or virus strain are used as an initial substance, a screening of the person placing the order is usually requested. However, there are sometimes great differences in this respect worldwide. Nevertheless, there are already mechanisms that automatically subject the online ordering of several suspicious materials to a closer examination. An example is the code of conduct for gene synthesis published by the International Association of Synthetic Biology in 2009. Fourth, it is often standard practice to modify existing research for one's own purposes. However, specific research on modern biological weapons is of course top secret. But it is still possible to gather information from civilian dual-use literature, but this requires a higher degree of specialist expertise. Fifth, the actor would have to undertake potentially extensive testing and adjustments prior to deployment. Such tests can easily arouse suspicion in various ways. The author also describes that there is already an established community of so-called biohackers in many countries around the world. Determined nonstate actors might join such an often anonymous internet hobby community to act more effectively.[61]

At the same time, of course, it is also possible that such a biohacker could lose control of a potentially dangerous agent as a result of an accident, since generally weaker standards of safety are observed in amateur labs. Liu's six-year-old remarks must also be seen in the light of the fact that more advanced technologies are already available on the internet now. In the future, it will probably be even easier to circumvent the barriers as, for example, the aforementioned small flow reactors and CRISPR-Cas9 applications become widely marketable.

All in all, synthetic or DNA-engineered biological weapons can potentially cause enormous damage, but a closer look reveals that, at least for nonstate actors, production is currently not as easy as it might seem. By 2030, however, some of the current barriers are expected to be significantly lower. Although it

is possible to learn the fundamentals via internet courses, in most cases a solid academic education is needed to gain practical experience with the laboratory equipment. Compared to genetically modified agents, existing natural pathogens may pose an even greater danger, as slightly less experience is required to weaponize them. There is also more publicly available research and potential natural source sites for such pathogens. In 2014, for example, a Tunisian jihadist did not even attempt to produce complicated pathogens, but instead records were found on their laptop of how the causative pathogen of plague (*Yersinia pestis*) can be isolated from infected animals and subsequently weaponized. The chemist and physicist would presumably have had the theoretical prerequisites for creating his own strain, but it seems the costs were too high compared to the benefits.[62] He was caught without carrying out an attack.

It would also be relatively easy for nonstate actors to take advantage of a natural outbreak to infect themselves and then infect as many other people as possible. Breaking an imposed quarantine during a disease outbreak for political reasons could also be classified as terrorism, as people could be killed indirectly. Such intentions, as well as acting as a so-called superspreader, are entirely possible, as already described in the section on SARS-CoV-2. However, it is relatively difficult to deliberately infect oneself with a naturally occurring virus as the first carrier. Another comparatively simple biological weapon that could be used for attacks in the future is the mass breeding of insects. This can lead to effective attacks on crops, but as soon as the insects are to be used as vectors for diseases against humans, a greater effort might be required, although it might still be much less than that of producing a synthetic pathogen. The use of insectoid vectors proved to be very effective in the operations carried out by the Japanese during the Second World War. Other biological agents already used in the past, such as anthrax and ricin toxin, might also potentially be used in the future again. Currently, the Centers for Disease Control and Prevention lists more than 20 dangerous bioterrorism agents, which they subdivided into three categories.[63]

In addition, there is the danger of developments by state actors that could be misused for terrorist purposes by employees, fall into the hands of nonstate actors, be released as a result of an accident, and could be used intentionally or as part of a covert operation. The unconfirmed efforts of the People's Republic of China operating a disguised dual-use bioweapons program are a cause for concern.[64] It is also very problematic that various states have not ratified international agreements and, in some cases, do not adhere to international standards, which could facilitate proliferation to potentially adversarial nonstate actors. The internet, and its global expansion, will continue to play a fundamental role in the future through legal and illegal orders, educational courses, and specialized biohacking communities, as well as the latest research and publicly accessible DNA/RNA databases.

With a prospective application in mind, a distinction must be made between how demanding it is to produce or obtain a specific biological weapon.

As with chemical weapons, greater effectiveness goes mostly hand in hand with more difficult acquisition and are thus less likely to be used. This rough prediction may be obsolete by 2030, as technological advances lower the threshold for acquisition while increasing lethality. As emphasized in the introduction, it is important to note that various current and future biotechnological developments have the potential to limit and thus to a certain degree control transmissible biological weapons.

Current and Future Means of Delivery for Biological Material

Due to the often-unstable nature of biological pathogens outside the laboratory, methods of dissemination are also important. In the following, current and conceivable methods by 2030 are examined in more detail. A whole range of bombs, including cluster bombs and balloon bombs, were developed for use with biological weapons at the beginning of the Cold War. Many of these developments were aimed at destroying enemy crops with plant pathogens. In the Second World War, Japan used, among other things, ceramic bombs filled with pathogens. While most chemical weapons can be stored for longer periods of time in their means of delivery and can be used relatively effectively by many methods, biological weapons usually require a much more cumbersome procedure. Due to the high impact energy of nonbraked bombs and missiles, successful dissemination of a biological agent is not likely. Parachuted bombs with a large-scale dispersal mechanism are more likely to succeed. However, anthrax spores are nevertheless known to survive dispersal by low-yield explosion, as found for example in the American E61, E120 or M143 cluster bomb submunitions developed in the 1960s.[65] However, a careful explosive delivery system for sophisticated bioweapons is very difficult for nonstate actors to achieve on their own. A civilian aircraft could be bought or rented for the drop of a bomb or cannister, but the overall cost of such a venture is very high compared to the possible outcome.

Easy to control, maneuverable, low-cost UAVs with a comparatively high payload designed for the civilian market have become quite popular in the last decade and see regular combat operation, for example in the Ukraine war of 2022. In addition, camera technology is becoming smaller and smaller, batteries come with improved storage capacity, and small and lightweight flight controllers, accelerometers, and GPS (Global Positioning Systems) are becoming increasingly widespread. Thanks to mass production, mostly in the People's Republic of China, models are now available in many price ranges and payload sizes. In the meantime, a large market has also established itself with do-it-yourself components with which mission-oriented UAVs can be built relatively easily. This can be done both as a fixed-wing aircraft and as a multicopter or helicopter. In recent years, a growing market has also emerged that specializes in professional applications and offers more expensive, but still affordable, products. In the United States alone, almost 750,000 commercial and recreational drones

are currently registered.[66] At the same time, effective defense against these commercial UAVs remains a major challenge. In practice, it is also difficult to distinguish between registered and legal drone flights and potential attacks.

At an event organized by the Center for Arms Control, Energy, and Environmental Studies in 2011, some interesting points were made in relation to UAVs. For example, a simulation was mentioned in which 900g of weapons-grade anthrax would be released 100 meters above a large city. With appropriate winds, about 1.5 million people would be infected and tens of thousands would die despite strong containment measures. At the same event, the TAM-5 model aircraft was mentioned, which flew automatically for 39 hours in 2003 and traveled more than 3,000 km over the Atlantic.[67] Since 2009, more and more UAVs have been configured as multicopters. These models usually cannot fly as far or as long as fixed-wing aircraft, but they are more maneuverable and usually easier to operate. Modern remote-controlled aircraft can fly far faster than 500 km/h; modern quadrocopters far faster than 200 km/h. For professional applications, there are now drones with a payload of more than 100 kg.[68] In 2016, British prime minister David Cameron warned that UAVs could disperse radioactive material in massive quantities over cities. He is probably alluding to the wide availability of automated crop duster UAVs, which are in fact a low-effort, high-impact means of delivery for terrorists, especially when many people are crowded together in the open. Instead of radioactive material, however, chemical or biological material could be effectively disseminated.[69] State actors with access to professional technologies have resources to develop further technical solutions tailored to the agent. Manned aircraft for the deployment of CBRN material have been little considered by nonstate actors. In the past, Aum Shinrikyo tried to modify a Mil Mi-17 helicopter to spray toxic gas over Tokyo.[70] In 2001, an al-Qaeda terrorist traveled to the United States to possibly prepare an attack with a crop duster plane.[71]

In addition to aerial deployment, CBRN material can also be deployed from the ground. The direct application of pathogens, as in the 1984 Rajneeshee bioterror attack, can be considered a ground-based attack. The same applies to attempts to deliberately transmit SARS-CoV-2 or other viruses to, e.g., door handles or from person to person. This category also includes assassinations with biological warfare agents.

A subcategory of biological warfare is entomological warfare. There are two fields of application, because insects can be used to act directly as weapons or to spread pathogens. But noninsectoid animals can also be used to deliberately spread pathogens. This type of warfare was first systematically studied and applied during the Second World War. Japan was particularly involved; the empire infected Chinese populations with plague-infected fleas and cholera-spreading flies. This mode of transmission proved catastrophically effective. Yellow rats were also bred in large numbers for use as vectors.[72] After the war, the Soviet Union, among others, researched ticks as vectors. According to their own statement, an automatic insect breeding facility was developed.[73] Such a

facility was also planned in the United States, where mosquitoes and fleas were successfully tested as vectors and were dropped from airplanes.[74] But nonstate actors have also recognized the advantages of insects as biological weapons. For example, in 1989, after a letter from a group called "the Breeders" was found, "peculiar patterns of Mediterranean fruit fly infestation in southern California that year" were detected.[75] More recent cases have not been detected. In principle, it is easier to use insects as weapons than to successfully infect vectors with deadly diseases without endangering oneself. Major financial damage or famine due to crop shortfalls can be a consequence that is not directly fatal to humans.

As already indicated, the biological field is probably the most significant for the future. The possibilities of releasing and spreading a fully developed pathogen are very diverse and almost impossible to prevent. In jihadist circles, for example, one of the terrorists could be the first carrier, while other types of terrorists might want to harm a specific person or group of people. From poisoned water to public salad buffets, there are many methods. In the future, however, genetically manipulated or even synthetic bacteria, insects, or other animals will be particularly useful as vectors. Such animals can be bred or designed according to the requirements at hand (e.g., to reproduce and spread particularly quickly or to deliver the pathogen particularly effectively). Similarly, in the future it will often be difficult to distinguish manipulated animals from nonmanipulated animals. Thus, the origin of the outbreak can be concealed, which presents potential for a state attack disguised as a terrorist attack, or vice versa.

Biological means of delivery of pathogens can already be prepared with the help of artificial hatcheries or programmed to reproduce themselves as quickly as possible. The latter might be a logistically more effective solution, although manual incubation requires less expertise in the field of molecular biology. In the future, modified organisms may be able to identify and attack certain people or groups of people on the basis of certain characteristics or infect them specifically with the transported pathogen. Similarly, carrier animals could be manipulated to feel comfortable in other climates or environments and attack the local population or displace native species. Climate change would accelerate such intentions. It is also possible that by 2030, technologies will exist that can artificially control insects or small animals, turning them into covert weapons. Currently, this already works with beetles. In this way, CBRN materials could be delivered unnoticed to a specific target without attracting attention. A pathogen that has a deliberately long delay to disease onset or death built in can be used to spread unobtrusively in humans or animals before it is detected.

In addition to the ways of delivering biological material already discussed, there are other ways that can be used to contaminate soil, water, or plants. The perpetrator can either use one of the previously explained systems, such as an agricultural UAV. A simpler way is to distribute the agent personally in unguarded places. Biological agents such as anthrax are likely to contaminate soil permanently. The two best-known examples are Gruinard Island in Scotland and Vozrozhdeniya Island in what is now Uzbekistan and Kazakhstan.

Both were partially contaminated by tests with *Bacillus anthracis*, the cause of anthrax; studies proved the extreme persistence of the biological weapon in soil during initial decontamination attempts.[76] To alert the public to the dangerous situation on the island, unknown perpetrators sent two packages of soil samples from Gruinard Island almost 40 years after the initial release of anthrax agent. One of the packages actually contained anthrax spores.[77] The island was then thoroughly decontaminated. The former Soviet biological weapons test site in the Aral Sea was also decontaminated in 2002 with funds from the United States, because many anthrax cultures were not sufficiently destroyed by the Soviets. Nevertheless, it is likely that live spores could still be found in unknown locations on the island. *Yersinia pestis*, known as plague, and smallpox virus have also been experimented with on the Soviet testing area but are not likely to have survived until today.[78]

The deliberate poisoning of water, mostly of human drinking water, has been discussed many times in the past. In such a case, it is known as a point source. In fact, in 1972, two teenagers tried to poison Chicago's drinking water with biological agents, but they did not come close to achieving their goal.[79]

The deliberate poisoning of plants or livestock with biological agents is a very broad field of application that has been studied and partially applied since before the Second World War. In the past, Germany, France, Japan, Iraq, the United Kingdom, the United States, and the Soviet Union pursued such programs, sometimes on a large scale.[80] The means of delivery are either vectors or insects themselves, but the use of anticrop fungi and other transmissible plant diseases has also been successfully tested. Once applied to a plant, it then serves as both the means of delivery and the target of the weapon. As with soil contamination, there are theoretically multiple motivations for terrorists to engage in agro-terrorism. Agro-terrorism can often be closely linked to entomological warfare methods. For more information, see the section on animals as a means of delivery. Jonathan Ban of the Chemical and Biological Arms Control Institute lists some motivations:

> Some actors may be motivated for the same reasons as other terrorist actions—to attract attention to a cause, incite fear, disrupt society, or demonstrate a capability with the intent of exacting political concessions. Other actors may be prompted by different motives—economic interest, sabotage, or revenge.[81]

He lists several cases in which crop poisoning was threatened or carried out. In the described cases, chemicals like mercury or cyanide were used for poisoning, but not self-transmitting biological weapons. Also, the alleged medfly attacks in California in 1989 had food production, in this case mass-produced fruits, as a target.[82] The Federation of American Scientists provides information on further incidents of biowarfare against agriculture: "In 1985 and 1988, Iraq conducted field tests of wheat cover smut to demonstrate its effectiveness as an anti-crop agent. Iraq also produced canisters designed to disperse the fungal

agent over Iranian wheat fields. In Sri Lanka in the early 1980s, a group of Tamil separatists threatened to spread non-endemic plant diseases among rubber and tea plantations in a scheme to undermine the government."[83]

In the section on emerging technologies, the potential and current areas of application of nanotechnology in the CBRN sector have already been outlined. There is also a future field of application in the area of means of delivery. Future systems can use the bionic advantages of real living beings and combine them with the advantages of technical applications. Since only a few grams of various toxins or pathogens are often needed to have a lethal effect or to start an epidemic compared to current nuclear weapons, for example, nanorobots are also suitable for delivering the material. Also, camouflage as, for example, a mechanical rat or bird is possible to outsmart security measures of military premises or essential personnel. It is unlikely that nonstate actors will be able to build and operate such complex military high-tech means of delivery, but a dual-use application of such technologies is not impossible by 2030.

Fully autonomous vehicles are certainly part of the future of 2030. With autonomous UAVs, the damage of even low-quality CBRN weapons can be increased by automatically matching and selecting between multiple detected targets. Reprogramming requires IT skills, but these can also be obtained by terrorist groups. Deployed en masse, autonomous vehicles can carry out many different conceivable types of attacks and cause increased panic among the population, which is further exacerbated by the use of CBRN material. Autonomous drones can also target, for example, crowds of people with CBRN material, move on, and attack new identified targets. This saves CBRN material and makes the attack more effective, as even agricultural drones have a rather limited capacity when it comes to creating a deadly concentration of an agent in the air.

Possible Actors

The last and final section provides an overview of possible actors up to the year 2030. Earlier in the article, China and its dual-use biotechnology activities were discussed in more detail. Of the potentially hostile state actors, however, North Korea must also be mentioned, whose possible bioweapons program is explained in two reports as well as the Russian Federation, about whose current bioweapons allegations there is also a detailed article.[84] In the case of both countries, however, there is no definitive evidence. On Iran and a possible bioweapons program, sources are comparatively sparse.

Starting with state actors that may have sophisticated and resource-intensive capabilities to research, produce, and deploy biological weapons, it must never be forgotten that former state actors, like defectors or disloyal soldiers, may also get their hands on these biological weapons or sophisticated weapons get stolen or lost. In today's world and the world of 2030, there are also pseudo-nonstate actors who ostensibly operate autonomously but are significantly supported by a state actor. In addition to economically, religiously, and politically motivated

actors, there are also cults that stand out from other groups in the field of nonstate actors, since their goal may well be the extermination of all human life without limitation. Other nonstate actor groups that could theoretically plan to use biological weapons by 2030 are ecoterrorists, extreme conspiracy theorists, cyberterrorists, internal staff, renegade scientists, or laboratory security personnel. The third major category is unintentional accidents in laboratories or accidents involving members of the biohacker community at home. For example, at least two accidents occurred in coronavirus laboratories in China in 2004, and the local outbreak of foot and mouth disease in the UK in 2007 was traced to a laboratory in Surrey.[85] The fourth category is incidents, outbreaks, and attacks of unknown origin, which is not unlikely in the context of possible hybrid warfare by 2030.

Conclusion and Overall Threat Potential

In conclusion, NATO forces will find themselves in an increasingly dangerous biological threat environment by 2030. Despite the diverse threat environment, the alliance must credibly ensure that it can continue to operate actively in the aftermath of biological weapons attacks. Despite the high potency of biological agents, the issue is often treated only half-heartedly in armed forces and often remains a secondary consideration in national security strategies, despite the COVID-19 pandemic as an illustrative example. This article shows that there are virtually no limits to future biological weapons. This type of weapon of mass destruction has the potential to fundamentally change the future of warfare. As Ainscough's prognosis shows, this is not necessarily a new conclusion. The hypothesis is thus confirmed, although it is clear that forecasts for the future are always merely educated assumptions and that a large number of unknown factors play a decisive role in the real outcome.

It is very difficult to quantify the threat of future bioweapon attacks on a scientific basis. At the end of 2022, there is no concrete evidence that any actor is planning or threatening to use biological weapons in the near future. Nevertheless, the threat environment is evolving in a direction that fundamentally increases biological threats. Likewise, the progress of biotechnology will sooner or later lead to the development of limited transmissible bioweapons. So far, uncontrolled spread has deterred actors from using transmissible bioweapons. If, by 2030, it is possible to effectively limit biological weapons or make them nonlethal and endow pathogens with individual capabilities and attributes as designer pathogens, biowarfare could indeed establish itself as an alternative to traditional types of kinetic warfare in the future.

NATO forces must work closely together to develop effective counterstrategies and stay at the forefront of research to identify threats and develop effective countermeasures, as stated in the *NATO 2030* agenda. Additionally, the U.S. Marine Corps should address biological threats more thoroughly. At the same time, the defensive nature and safe conduct of their own biological research must always be made clear at the international stage and a treaty structure

adapted to the changed conditions of our time, in particular with the People's Republic of China, must be sought diplomatically. It must be reliably ensured that, despite a lower barrier, the use of biological weapons will continue to elude the interest of any actors in the future.

For further research, the author recommends the development of effective counterstrategies to future biological weapons attacks and an outlook on what biotechnological advances potential adversaries could use to make their soldiers more capable and resilient in the future.

Endnotes

1. *Force Design 2030* (Washington, DC: Headquarters Marine Corps, 2020).
2. Jason Blessing, Katherine Kjellström Elgin, and Nele Marianne Ewers-Peters, eds., *NATO 2030: Towards a New Strategic Concept and Beyond* (Washington, DC: Henry A. Kissinger Center for Global Affairs, Johns Hopkins University, 2021).
3. Kate Charlet, "The New Killer Pathogens: Countering the Coming Bioweapons Threat," *Foreign Affairs* 97, no. 3 (May/June 2018): 178–85.
4. George W. Christopher et al., "Biological Warfare: A Historical Perspective," *Journal of the American Medical Association* 278, no. 5 (1997): 412–17.
5. Matthew Meselson, "The Sverdlovsk Anthrax Outbreak of 1979," *Science* 266, no. 5,188 (1994): 1,202–8, https://doi.org/10.1126/science.797370.
6. Michael J. Ainscough, *Next Generation Bioweapons: The Technology of Genetic Engineering Applied to Biowarfare and Bioterrorism*, Future Warfare Series 14 (Maxwell Air Force Base, AL: Air University, 2002).
7. James Weaver, letter to the editor, "Slow Medical Sleuthing," *New York Times,* 24 April 2001.
8. Barry Kellman, "Biological Terrorism: Legal Measures for Preventing Catastrophe," *Harvard Journal of Law and Public Policy* 24, no. 2 (2001).
9. Kellman, "Biological Terrorism."
10. "Smallpox as a Biological Weapon," *Journal of the American Medical Association* 281, no. 22 (1999): 2,127–37, https://doi.org/ 10.1001/jama.281.22.2127.
11. "Smallpox," World Health Organization, accessed 11 May 2023.
12. "U.S. Officials Declare Researcher Is Anthrax Killer," CNN, 6 August 2008.
13. Rolf Mowatt-Larssen, *Al Qaeda Weapons of Mass Destruction Threat: Hype or Reality?* (Cambridge, MA: Harvard Kennedy School, Belfer Center for Science and International Affairs, 2010).
14. "Al Qaeda's Bio Weapons," CBS News, 31 March 2005.
15. Mowatt-Larssen, *Al Qaeda Weapons of Mass Destruction Threat*.
16. "Ricin and the Umbrella Murder," CNN, 23 October 2003.
17. "Two Georgia Men Convicted in Ricin Plot Against U.S. Government," Reuters, 17 January 2014; "*Breaking Bad* Fan Jailed over Dark Web Ricin Plot," BBC, 18 September 2015; "Henderson Man Sentenced for Unlawfully Possessing Ricin," press release, U.S. Attorney's Office, Eastern District of Texas, 2 August 2018; and Clare Hymes, "Woman Accused of Sending Ricin Letters to White House Charged with Making Threats Against the President," CBS News, 22 September 2020.
18. Florian Flade, "The June 2018 Cologne Ricin Plot: A New Threshold in Jihadi Bio Terror," *CTC Sentinel* 11, no. 7 (August 2018).
19. Harald Doornbos and Jenan Moussa, "Found: The Islamic State's Terror Laptop of Doom," *Foreign Policy*, 28 August 2014.
20. "Ricin Letters Mailed to President and Senator," START, April 2013.
21. Nadine Salman and Paul Gill, "Terrorism during the COVID-19 Pandemic," UCL JDI Special Series on Covid-19: No. 13, May 2020.

22. Gary Ackermann and Hayley Peterson, "Terrorism and COVID-19: Actual and Potential Impacts," *Perspectives on Terrorism* 14, no. 3 (2020): 59–73.
23. Arie W. Kruglanski et al., "Terrorism in Time of the Pandemic: Exploiting Mayhem," *Global Security: Health, Science and Policy* 5, no. 1 (2020): 121–32, https://doi.org/10.1080/23779497.2020.1832903.
24. Tom Mangold and Jeff Goldberg, *Plague Wars*: *The Terrifying Reality of Biological Warfare* (New York: St. Martin's Press, 2001), 92.
25. C. Drew Harvell et al., "Climate Warming and Disease Risks for Terrestrial and Marine Biota," *Science* 296, no. 5,576 (June 2002): 2,158–62, https://doi.org/10.1126/science.1063699.
26. Amelie Bottollier-Depois, "How Climate Change Could Expose New Epidemics," Phys.org, 16 August 2020.
27. Isidore K. Kouadio et al., "Infectious Diseases Following Natural Disasters: Prevention and Control Measures," *Expert Review of Anti-Infective Therapy* 10, no. 1 (2012): 95–104, https://doi.org/10.1586/eri.11.155.
28. Hannah Ritchie, "Which Countries Eat the Most Meat?," BBC, 4 February 2019.
29. Romain Espinosa, Damian Tago, and Nicolas Treich, "Infectious Diseases and Meat Production," *Environmental and Resource Economics*, no. 76 (2020): 1–26, https://doi.org/10.1007/s10640-020-00484-3.
30. Laura Spinney, "Is Factory Farming to Blame for Coronavirus?," *Guardian*, 28 March 2020.
31. Bhaskara Reddy and Milton Saier Jr., "The Causal Relationship between Eating Animals and Viral Epidemics," *Microbial Physiology* 30, no. 1 (2020): 2–8, https://doi.org/10.1159/000511192.
32. Allison E. Betus, Michael K. Jablonski, and Anthony F. Lemieux, "Terrorism and Intergroup Communication," Oxford Encyclopedia of Communication, 26 October 2017, https://doi.org/10.1093/acrefore/9780190228613.013.409.
33. Afzal Ashraf and Anastasia Filippidou, *Terrorism and Technology* (Ankara, Turkey: Centre for Excellence Defence Against Terrorism, n.d.), 8.
34. "Allied Command Transformation Strategic Foresight Work," NATO, accessed 3 May 2023, 45–55.
35. Rasmus O. Bak, Natalia Gomez-Ospina, and Matthew H. Porteus, "Gene Editing on Center Stage," *Trends in Genetics* 34, no. 8 (2018): 600–11.
36. Kim Riley, "Bioterrorism Threats Require Common Global Experimentation Oversight, Expert Says," Homeland Preparedness News, 10 August 2017.
37. David Lawrence, "Genetic Engineering and Human-Animal Hybrids: How China Is Leading a Global Split in Controversial Research," Conversation, 3 September 2019.
38. Julius Fredens et al., "Total Synthesis of *Escherichia coli* with a Recorded Genome," *Nature* 569, no. 7,757 (2019), 514–18.
39. Pinero-Lambe et al., "Programming Controlled Adhesion of E. coli to Target Surfaces, Cells, and Tumor with Synthetic Adhesins," *ACS Synthetic Biology* 4, no. 4 (2014): 463–73, https://doi.org/10.1021/sb500252a.
40. "NCBI Virus," National Library of Medicine, accessed 11 May 2023.
41. Ian Sample, "Synthetic Biology Raises Risk of New Bioweapons, US Report Warns," *Guardian*, 19 June 2018.
42. Claire Delacour, "Why Use a Microreactor for Chemical Processes?," European Training Network for Continuous Sonication and Microwave Reactors, accessed 3 May 2023.
43. Andreas Zaugg, Julien Ducry, and Christophe Curty, "Microreactor Technology in Warfare Chemistry," *Military Medicine Science* 82, no. 2 (2013): 63–68, https://doi.org/10.31482/mmsl.2013.009.
44. Anita Salic and Bruno Zelic, "Synergy of Microtechnology and Biotechnology: Microreactors as an Effective Tool for Biotransformation Processes," *Food Technology & Biotechnology* 56, no. 4 (2018): 464–79.
45. Gary Sheftick, "Army Developing Robotic Insect," U.S. Army, 17 December 2014.

46. Suping Li et al., "A DNA Nanorobot Functions as a Cancer Therapeutic in Response to a Molecular Trigger in Vivo," *Nature Biotechnology* 36, no. 258–64 (2018): https://doi.org/10.1038/nbt.4071.
47. Jeff Daniels, "Mini-nukes and Mosquito-like Robot Weapons Being Primed for Future Warfare," CNBC, 17 March 2017.
48. "DARPA Microsystems Exploration Seeks Revolutionary Advances in Military Embedded Computing Technologies," Military and Aerospace Electronics, 16 July 2019.
49. Margaret E. Kosal, "The Threats from Nanotechnology," *Bulletin of the Atomic Scientists* 75, no. 6 (2019): 290–94, https://doi.org/10.1080/00963402.2019.1680054.
50. Nicholas Winstead, "The Applications and Implications of Nanotechnology," American University, 15 April 2020.
51. Erika Check, "Terror Watchdog Set up for 'Dual Use' Biology," *Nature* 428, no. 109 (2004): https://doi.org/10.1038/428109a.
52. *Biotechnology Research in an Age of Terrorism* (Washington, DC: National Academies Press, 2004).
53. Jonathan B. Tucker, ed., *Innovation, Dual Use, and Security: Managing the Risks of Emerging Biological and Chemical Technologies* (Cambridge, MA: MIT Press, 2012).
54. Erika Kinetz and Desmond Butler, "Chemical Weapon for Sale: China's Unregulated Narcotic," *Berkshire (MA) Eagle*, 7 October 2017.
55. Meia Nouwens and Helena Legarda, "China's Pursuit of Advanced Dual-use Technologies," International Institute for Strategic Studies, 18 December 2018.
56. *Adherence and Compliance with Arms Control, Nonproliferation, and Disarmament Agreements and Commitments* (Washington, DC: Department of State, 2021).
57. *Adherence and Compliance with Arms Control, Nonproliferation, and Disarmament Agreements and Commitments*.
58. Mackenzie Foley, "Genetically Engineered Bioweapons: A New Breed of Weapons for Modern Warfare," *Dartmouth Undergraduate Journal of Science* (Winter 2013).
59. Ainscough, *Next Generation Bioweapons*.
60. Ainscough, *Next Generation Bioweapons*.
61. Zian Liu, "Bioweapons . . . for Dummies?," *Bulletin of the Atomic Scientists*, 28 September 2015.
62. Bruce Goldman, "How-to Manual for Making Bioweapons Found on Captured Islamic State Computer," *Scope*, 3 September 2014.
63. "Bioterrorism Agents/Diseases," Centers for Disease Control and Prevention, accessed 4 May 2023.
64. Elsa B. Kania and Wilson Vorndick, "Weaponizing Biotech: How China's Military Is Preparing for a 'New Domain of Warfare'," Defense One, 14 August 2019.
65. Jeanne Guillemin, *Biological Weapons: From the Invention of State-Sponsored Programs to Contemporary Bioterrorism* (New York: Columbia University Press), 101.
66. Federica Laricchia, "Consumer and Commercial Drones Worldwide—Statistics and Facts," Statista, 20 April 2022.
67. Eugene Miasnikov, *The Threat of the Use of Small UAVs by Terrorists: Technical Aspects* (Moscow, Russia: Center for Arms Control, Energy and Environmental Studies, Moscow Institute of Physics and Technology, 2004).
68. "Griff 300 Review: Drone that Can Lift 500 Pounds," Drone Tech Planet, accessed 4 May 2023.
69. David Hambling, "Could ISIS Really Attack the West with a Dirty Drone?," *Popular Mechanics*, 8 April 2016.
70. Suzuki Nathie, "Prophet of Terror: The Story of Shoko Asahara, Aum Shinrikyo, and the Danger of Religious Terrorism," *Suzuki's Thoughts* (blog), 14 January 2019.
71. Mowatt-Larssen, *Al Qaeda Weapons of Mass Destruction Threat*, 16.
72. Jeffrey A. Lockwood, "Insects: Tougher than Anthrax," *Boston Globe*, January 2010.
73. Jonathan Ban, *Agricultural Biological Warfare: An Overview* (Alexandria, VA: Chemical and Biological Arms Control Institute, 2000), 3.
74. William H. Rose, *An Evaluation of Entomological Warfare as a Potential Danger to the*

United States and European NATO Nations (Dugway, UT: U.S. Army Dugway Proving Ground, 1981).

75. Ban, *Agricultural Biological Warfare*, 4.
76. Zaria Gorvett, "The Deadly Germ Warfare Island Abandoned by the Soviets," BBC, 28 September 2017.
77. "Biological Warfare: Dark Harvest," *Time*, 9 November 1981.
78. Zaria Gorvett, "The Deadly Germ Warfare Island Abandoned by the Soviets."
79. Michael Miner, "The Terrorist Mind—A Look Back at a 1972 Plot to Poison," *Chicago Reader*, 25 September 2012.
80. Ban, *Agricultural Biological Warfare*.
81. Ban, *Agricultural Biological Warfare*.
82. Ban, *Agricultural Biological Warfare*.
83. "Biowarfare Against Agriculture," Case Studies in Agricultural Biosecurity, accessed 4 May 2023.
84. Hyun-Kyung Kim, Elizabeth Philipp, and Hattie Chung, "North Korea's Biological Weapons Program: The Known and the Unknown," Belfer Center for Science and International Affairs, Harvard Kennedy School, October 2017; Elisa D. Harris, *North Korea and Biological Weapons: Assessing the Evidence* (Washington, DC: Stimson Center, 2020); and Robert Petersen, "Fear and Loathing in Moscow: The Russian Biological Weapons Program in 2022," *Bulletin of the Atomic Scientists*, 5 October 2022.
85. Robert Walgate, "SARS Escaped Beijing Lab Twice," *Genome Biology*, no. 4 (2004): https://doi.org/10.1186/gb-spotlight-20040427-03; and Andrew Alderson, Richard Gray, and Patrick Hennessy, "Foot and Mouth Lab Failure Causes Outbreak," *Telegraph*, 5 August 2007.

Sovereignty, Cyberspace, and the Emergence of Internet Bubbles

Eldar Haber, PhD; and Lev Topor, PhD

Abstract*:* The cyber domain emerged as a perfect platform for international struggle over power and influence. International powers are actively engaged in cyber proxy warfare due to the relatively low risk of escalation, various enforcement challenges, and the vagueness of international law within this realm. These indirect conflicts might lead some global powers to close or restrict their virtual borders to avoid or reduce the plausibility of cyber proxy warfare or unwanted foreign influence in general. The formation of such restricted networks, articulated in this article as "internet bubbles," is already shaping within the realm of actors like Russia, China, North Korea, and Iran. The authors argue that liberal democracies like the United States might be at a severe disadvantage to fight against cyber proxy warfare due to legal and constitutional barriers. But at the same time, the emergence of platform governance and self-regulation might be proven as a new force within these proxy wars and reshape its boundaries.
Keywords: international security, cybersecurity, internet, proxy warfare, sovereignty

Introduction

> *From Stettin in the Baltic to Trieste in the Adriatic an iron curtain has descended across the Continent*.
>
> ~ Winston Churchill

Winston Churchill's quote refers to the Soviet Iron Curtain—the nonphysical boundaries dividing Europe at the end of World War II.[1] Today, countries worldwide are forming digital iron curtains

Dr. Eldar Haber is an associate professor at the Faculty of Law, University of Haifa, Israel. Dr. Lev Topor is a visiting ISGAP scholar at the Woolf Institute, Cambridge, UK, and a senior research fellow at the Center for Cyber Law and Policy (CCLP), University of Haifa, Israel.

Journal of Advanced Military Studies vol. 14, no. 1
Spring 2023
www.usmcu.edu/mcupress
https://doi.org/10.21140/mcuj.20231401006

within their efforts to preserve sovereignty and control public opinion. Russia, China, North Korea, and Iran, to name a few key examples, control and restrict their cyber domains to prevent foreign intervention. In contrast, liberal democracies like the United States currently lack substantial legislative freedom to similarly control and restrict their cyber domains and are therefore becoming more susceptible to foreign interference of various types. This was demonstrated very recently with the conflict between Ukraine and Russia (2022–23)—the latter restricted domestic media to try and restrain opposition to this conflict while it also disseminated anti-Ukrainian and anti-Western propaganda to try and undermine Western support for Ukraine.

The main argument of this article is that some countries worldwide will attempt, and many times succeed, to form their own restricted internet networks—"internet bubbles"—for the purpose of avoiding undesired foreign influence and to better govern and control their domestic affairs.[2] The authors further argue that these internet bubbles position nondemocracies better than democracies to gain and preserve cyber sovereignty, considering the difficulty to attribute cyberattacks and propaganda.[3] However, these internet bubbles are not hermetically sealed, and the rise of platform and corporate governance might aid democracies to govern their virtual borders from foreign influence.

Examining the hypothesis begins with discussing the rise in cyber warfare and foreign interventions through cyber means. Notably, the internet was always subjected to hacking and manipulations by foreign agents, often conducted through its backbone.[4] But cyber warfare and other forms of foreign interventions became more common and prominent for many countries worldwide recently, directed mostly against the West and the United States, and might in turn threaten sovereignty. Cyberattacks were directed at the state not only directly but also through private parties, serving as a state's beneficiary proxy, as exemplified within the cyberattack by North Korea against Sony Pictures Entertainment in November 2014.[5] Further, it is only natural for a sovereign state to protect itself from malicious foreign interventions. Yet, authoritarian states also seek to limit foreign civil and cultural influences.

Methodologically, the authors examine the arguments, suggestions, and predictions with a traditional international relations approach and treat each international actor as a unitary actor seeking to gain complete sovereignty and independence. This argument is based on traditional theories of international relations, sovereignty, and proxy warfare, as well as a legal analysis of cyber proxy wars from both international and domestic law perspectives. Since it is an extreme and obvious case of undesired foreign influence, the focus of this article is on cyber proxy warfare.

This article examines and compares the three cyber domains of three global powers—the Russian, Chinese, and American domains—to predict how international actors will use cyber warfare against their adversaries, while keeping their own cyber domains safe. Finally, other modalities are suggested that can

replace the necessity of creating internet bubbles—a suggestion that is derived from the comparison of the American, Russian, and Chinese cases.

Interestingly, the February 2022 invasion of Russian forces into Ukraine and now the conflict between the parties has demonstrated that cyber warfare is limited. It is an important tool for times of peace and times of tensions and mainly for disseminating propaganda. However, in times of kinetic conflicts, the utility of cyber warfare is limited simply due to the fact that it takes kinetic means like infantry, tanks, jets, and other weapons to conquer land. The conflict between Ukraine and Russia has also demonstrated the argument about the strategic need of an internet bubble. That is, putting values like democracy, liberalism, and human rights aside, Russia has restricted its internet and media to deny any anti-governmental and pro-Western influences.

Sovereignty, Conflict, and Cyber Proxy Wars: Setting the General Framework

Nations generally desire to control their internal affairs. That is, they seek the ability to control their domestic affairs, control their population, as well as to control their ability to make foreign policy decisions like engagement in trade, war, or diplomatic relations in general.[6] In the context of this article, cyberspace is a platform upon which states can fulfill this desire of control, especially regarding their domestic affairs. Politicians and their constituents in the United States, the European Union (EU), Russia, and China have grown increasingly nervous about letting capital, goods, and people move freely across their borders and the threat of terrorism or even the COVID-19 pandemic only made this more prominent.[7] In the age of information and cyberspace, politicians and their constituents are also concerned about the type of information crossing into their digital borders.[8]

States are also willing to engage in conflicts over their sovereignty. In the twentieth and twenty-first centuries, major wars and conflicts have all been characterized by the involvement of foreign powers in the affairs of other actors or those of their allies and beneficiaries.[9] For the conceptual purpose of this article, conflict between states can emerge, among other ways, mainly in two ways: first, when actors disagree about their mutual international affairs. Second, when actors try to influence and intervene in the domestic political arena of other states.[10] In this regard, Fredric S. Pearson suggested that there are six key reasons for states' interventions in others' affairs: (1) they wish to acquire territory or domains; (2) to protect social groups; (3) to protect economic interests; (4) to protect military or diplomatic interests; (5) they intervene due to ideology; and, lastly, (6) to keep or adjust the regional balance of power.[11]

States may acquire control over matters through peaceful negotiations, military pressure, or any other use of power—soft, hard, smart, or sharp power.[12] In the context of this article, and the question of power and influence through cyberspace, the question of how one can measure sharp power such as disinformation or cyberattacks arises. This is rather complicated and has no definitive

answers yet, partially since many executions of power are made through proxies carrying out cyberattacks, blurring or hiding the involvement of an international actor in another's affairs, thus making the attribution of the hostilities even more difficult. Furthermore, the actual victim can be considered a state proxy itself, as one might treat Sony Pictures Entertainment as such within the abovementioned cyberattack against it in 2014. Moreover, even when a victim state can point at the perpetrator state or actor, traditional military or economical retaliation is often more difficult to justify than when dealing with kinetic actions, and the attribution problem often renders deterrence slow, blunt, and ineffective.[13] Furthermore, following Karl W. Deutsch and Andrew Mumford's theme, when states consider ideology, interests, and risks, they tend to opt for the use of proxies.[14]

The conceptual soil on which the conflict is now fought, in this respect, is cyberspace itself. Cyberspace allows states a rather high degree of anonymity and detachment from their actions. The difficult forensic process of attributing an attack to a specific perpetrator makes the internet an ideal tool for waging a proxy war.[15] While states are legally responsible for activities undertaken through their proxies, holding them responsible will depend on proof (i.e., the attribution of the proxy's actions to its patron). However, some actions—such as spreading disinformation and online propaganda—are currently not even considered illegal on the international level and states use this to influence other international actors and even to resist traditional hard power such as the case of Russian disinformation against the North Atlantic Treaty Organization (NATO) in the Baltic region.[16]

States strive to control their affairs, including the type and nature of information their citizens consume. Even liberal democracies seek to restrict influence on public opinion if this influence is malicious. The extreme case of such influence is cyber proxy warfare like foreign mis/disinformation campaigns and for this reason this article demonstrates this argument with examples of cyber proxy wars. Following the discussion of sovereignty, conflict, and traditional proxy wars, the authors define the term *cyber proxy wars* to further elaborate the argument. Combining and extending the definitions of proxy wars by Deutsch and Mumford, cyber proxy wars could be defined as international conflicts between two foreign powers fought on or using the cyber domain, disguised as actions taken by unrelated international actors or entities, made in an attempt to influence an actor's strategic outcomes; for instance, where both the attacker and the victim can be a proxy.[17] In the age of information and technology, cyberspace—through cyber warfare—serves as the perfect arena to avoid direct conflict while trying to obtain Pearson's six goals for intervention. In general, cyber warfare can be defined as using cyber weapons as well as the domain itself in order to execute strategies and policies that undermine and influence other international actors. These acts can be executed by all forms of international actors.[18] The characteristics and associated benefits of cyber tactics make them very attractive for use by states and even terror groups alike.[19]

The Formation of the Russian and Chinese Internet Bubbles

In this section, the authors focus on Russia and China since their internet bubbles are still relatively premature, although they are constantly growing, and in contrast to the North Korean *Kwangmyong* internet bubble, are still not entirely hermetic. The North Korean internet bubble is almost hermetically sealed off, and all information and communications, in and out of the country, are controlled by Kim Jong-un and his government.[20] The article also addresses the actual structure of cyberspace and argues that all layers of cyberspace can be restricted—physical, logical (data routing), information, and users. These layers affect many things such as regulation and the relations and interactions between all users, states and people alike.[21] In practice, as exemplified by the Russian and Chinese examples, restricting physical and logical layers can lead to a more restricted internet bubble while controlling information and users in cyberspace can lead to a more subtle internet bubble. That is, for instance, the fact that North Korea is physically and virtually disconnected from the global grid makes the North Korean internet very restricted, more so than Russian legislation against foreign information.

Unfortunately, while writing this article, one of the largest international conflicts since World War II between sovereign nations erupted between Ukraine and Russia. Although Moscow sought to have its grip on what it perceived as its "backyard" already in 2013–14 and even before, the current 2022–23 conflict—or war—between Ukraine and Russia demonstrates the argument about the strategic need for internet bubbles but also demonstrates the limited magnitude of cyber warfare.[22] In fact, Russia acts in two spheres of information and one of warfare. First, Russia seeks to restrict and control its domestic affairs, through control of media and information, to oppose any domestic criticism regarding the invasion of Ukraine.[23] Second, it disseminates anti-Ukrainian and anti-Western propaganda globally to undermine international support for Ukraine.[24] Third, Russia puts more effort in kinetic warfare than in cyber warfare simply due to the fact that its aims are kinetic—Moscow wants to conquer land and one does not conquer land just with cyber means but coupled with kinetic means like weapons, tanks, and infantry (that is, cyberattacks are secondary to the main effort).[25]

The Russian Internet Bubble

Russia has the potential to pose the largest threat to the United States, the European Union, and other democracies in general.[26] Its influence over global affairs is probably not lesser than its predecessor, the Soviet Union, as Moscow influences almost every major actor, in every region of the globe, and, as was uncovered in 2016, on its main adversary, the United States.[27] The Russian Federal Council has in fact emphasized the increasing importance of cyber warfare and use of cyber-related actions to accommodate and complement other types of acts in the international relations arena.[28]

Russia does not only utilize cyberspace to exert its influence on a global scale but also to protect itself from foreign cyber influence. Russia is working to create its own protective shield, a Russian internet bubble. Moscow's December 2019 successful attempt to unplug itself from the internet was just another step toward total domestic control of its domestic cyber domain—*RuNet*.[29] The Russian exclave Kaliningrad was connected to Russia via the Kaliningrad Cable, owned by Rostelecom, in 2021, further expanding its capabilities of internal communication.[30] Moscow had begun the process in early 2000s when it established control over television and the press—an act that allowed it to gain more control over information consumed by its citizens.[31] Moscow then turned to address cyberspace, and the developing Russian internet bubble is meant to deal with the technological and psychological aspects of the internet and its use by Russian citizens. For example, the Yarovaya Law, Russia's "sovereign internet" law, the Russian mass communications surveillance system (SORM), and the law making Russian applications mandatory on smartphones are all examples of the legal regulation of the technological aspects of the internet—namely and mostly the logic layer, which Russia seeks to gain control of.[32] The psychological aspect of the Russian cyber domain is controlled through the "fake news" law, the law concerning disrespect, and a recent law regarding foreign agents' activities.[33] The efforts on the part of the technological aspect are aimed to regulate outside sources, while those on the part of the psychological aspect are aimed to discourage Russian citizens from criticizing the authorities and cooperating with outside forces.[34]

In the context of the Russian "special military operation" in Ukraine, the abovementioned restriction of domestic information and media and the influence campaigns on foreign audiences allow Russia to implement sharp power. While leaks, anonymous communications, and rogue media allow Russian citizens a glance at the outside world, mass media is generally protected against unwanted information about the conflict in Ukraine and thus antigovernmental sentiments are limited.[35]

The Chinese Internet Bubble

China is another key global adversary of the United States and is much closer in diplomatic and military relations to Russia than to the United States, a fact that downgrades the United States from the global premier to some extent.[36] At home, China has successfully gained almost complete control of its internet since the early twenty-first century, restricting social media such as Facebook, Twitter, and even Tinder; blog platforms such as WordPress; some email providers; and even search engines such as Google. As an alternative, China allows for domestic social media platforms and other service providers to operate like WeChat, Weibo, Tencent, Baidu, and many others.[37] China also restricts access to messaging applications such as Telegram, Signal, and WhatsApp. Furthermore, platforms such as YouTube, Netflix, the *New York Times*, the BBC, and even Wikipedia are all restricted in China.[38] As previously mentioned, China's

internet is also currently locked behind the "Great Firewall"—a national project aimed at monitoring and censoring available online content through various means and methods—which can be conceptually compared to the Soviet Iron Curtain.[39]

China's foreign policy and cyber activities are aimed to protect the Chinese Communist Party (CCP) and to ensure domestic stability, territorial integrity, modernization, and economic growth; or, in other words, to ensure Chinese sovereignty and national security.[40] In December 2016, China has released its first *National Security Strategy*, which states that there can be no national security without cybersecurity and further reaffirms that "cyberspace sovereignty is an important part of state sovereignty." China's cybersecurity law, which acts as the baseline of its cyber regulation, came into effect in June 2017, alongside many other additional laws and policies that were enacted to ensure complete regulation of the internet—to ensure the CCP's "cyber sovereignty."[41]

As a national strategy, China addresses mainly the economic, political, and military spheres of cyberspace. As Amy Chang noted in 2014, there are six main issues promoted by the CCP: (1) economic development through cyber industrial espionage on other countries, including the United States; (2) pro-Communist propaganda and control over domestic information, as discussed; (3) utilization of offensive cyber operations to express discontent with acts of foreign powers; (4) development of military cyber capabilities both of infrastructure and of personnel; (5) maintaining intelligence and continuous reconnaissance of the cyber capabilities of China's adversaries; and (6) promotion and justification of domestic surveillance.[42] These six issues are executed by China's global footprint in the technological domain, especially as the Sino-American trade competition intensifies. Chinese companies like Huawei are perceived by the West as a challenge because China has found a way to penetrate the West not just with propaganda but with hardware and software as well. Yet this, of course, is a topic for another full article.[43]

China's strict control of its domestic internet and its general cyber sovereignty means that by now China effectively has an internet bubble. In comparison, Chinese internet regulation is stricter than its Russian counterpart, and in fact, it applies to all four layers of cyberspace: the Chinese government controls the physical layer through the regulation of routers, switches, servers, and other hardware in general. It commands the logic layer through its control of Domain Name Systems (DNSs), Internet Protocols (IPs), software, and websites. Power over the information layer is achieved through state censorship, and as a result, China also controls the user layer as the state manipulates and shapes users' experiences.[44] However, it should be noted that the restrictions imposed on the user layer and in part the information layer as well are not bulletproof as Chinese citizens and foreigners often employ workarounds, such as virtual private networks (VPNs) to bypass web restrictions.[45]

American Internet Regulation and Deproliferation: Responses to Foreign Insurgency

Cyber proxy wars are more challenging than kinetic ones from the legal and sovereignty aspects. They also negatively affect liberal democracies such as the United States more so than non or less-democratic states and might even threaten democracy. This is due, partially at least, to legal constraints and barriers for forming internet bubbles that serve to mitigate the dangers and harms of cyber proxy wars. The authors argue that the power to control parts of the internet, and the lack thereof, might eventually challenge the proper functioning of some governance forms, perhaps especially those whose legal regimes highly value and protect free speech. To further articulate the differences between cyber and kinetic proxy wars, one must first understand how some legal regimes might contest cyber proxy wars differently than kinetic ones. To do so, this article examines the two potential legal methods whereby cyber proxy wars are likely to be handled: international law and domestic law.[46]

The first realm that might affect cyber proxy wars is the international sphere, and more specifically international public law.[47] If international law prohibits proxy wars—then, prima facie, they should not be conducted. In reality, however, international law likely fails to regulate the conduct of states regarding proxy wars in general and cyber proxy wars in particular. Aside from political or otherwise economic barriers for such conduct regulation, the international sphere might prove trickier than one might presume, especially regarding cyber operations, which lack effective regulation within the realm of international law, as the article will discuss further.[48]

In the kinetic world, it is evident that states have almost full sovereignty over what occurs within their physical boundaries and can thus exercise various rights in response to certain hostile foreign acts, such as the right to self-defense.[49] The question of whether and how a state could respond to a nonphysical exercise of foreign powers within its own domain, be it the physical or cyber one, does not enjoy great legal certainty at this time. The answer would greatly depend on the characterization of the act and perhaps the harm that it caused, but also on a formal acknowledgment of state sovereignty in its cyber domain and its legal boundaries.[50]

Theoretically, the general legal status of cyber proxy wars could be inferred from that of regular proxy wars—those that existed prior to the emergence of the cybernetic ones. Proxy wars were formally acknowledged within the realm of international law since the 1980s.[51] While the legal framework around proxy wars consists of a patchwork of international treaties and customary law, it does establish legal obligations binding states to act responsibly in their use of proxies.[52] These rules and obligations establish, for example, a constraint on the use of force and the responsibility of a sponsor state for "internationally wrongful acts" committed by its sponsored proxy.[53] Conversely, the enforcement of such legal obligations is scant at best and thus lacks substantial teeth.[54]

To understand the extent to which cyber proxy wars could be regulated through international law, one might suggest setting a framework for interpreting cyber proxy wars under the existing legal framework, in equivalence to physical proxy wars. As previously suggested, building on Deutsch and Mumford's definitions of kinetic proxy wars, one might define cyber proxy wars as international conflicts between foreign powers, disguised as acts carried out by unrelated international actors in an attempt to influence an actor's strategic outcomes, using or fought on the cyber domain.[55]

Actions carried out as part of cyber proxy war campaigns might implicate and breach, inter alia, the existing international norms of nonintervention, as well as those prohibiting the "threat or use of force" and "armed attack" against a foreign state, similar to how kinetic actions by states and proxies could breach the same norms. If one further considers viewing cyber acts as acts of war, then, at least theoretically, they must first meet the requirements of *jus ad bellum* and then the law of armed conflict and international humanitarian laws.[56]

Nevertheless, cyber proxy wars are even more challenging to regulate than their kinetic predecessors. First, as previously mentioned, the problem of attribution is greatly enhanced in the cyber realm, adding difficulties to prove or even know which state was behind an attack. But aside from these challenges, applying the traditional legal framework of the international laws of war to cyber operations raises many difficulties regarding fulfilling traditional definitions and requirements originally meant to be applied to, and fulfilled by, the kinetic, physical world of war and its elements. Put simply, since international law only determines which physical actions would justify physical responses, the major challenge would be determining when a cyber action would amount to and equal such an action as to justify and make legitimate a response.[57]

This challenge served as one of the main reasons for the writing of the *Tallinn Manual*—a nonbinding expert's opinion on how international law should interpret and apply to cyber activities with respect to the law of war.[58] The manual addresses the issue of applying these norms to cyber operations and offers an interpretation of when a cyber conduct would breach each of them.[59] In international law, it is only when an action amounts to an "armed attack" that the right of self-defense may be invoked, allowing the injured state to respond to the hostilities.[60] Since states generally seek to retain sovereignty within their own cyberspace, they thus generally also enjoy the inherent right to act in self-defense in the face of an armed attack.[61]

According to the manual, one must examine whether the act in question constitutes either an intervention, a threat or use of force, or an armed attack; all depending on the purpose of the act, the target, and its impact. A state may therefore exercise its right to self-defense only when it is the target of a cyber operation that rises to the level of a kinetic armed attack.[62] Not every act conducted as part of a cyber proxy war will fall under the manual's or international law's requirements, as they will not constitute an "armed attack" and will therefore not qualify as an actionable act.[63] Even if the international law of war

were to unequivocally apply to cyber proxy wars, it would only allow for a very narrow and limited opportunity to respond, even if difficulties like attribution were overcome. With time, states might sign bilateral agreements in cyberwarfare, modeled on the Cold War-era arms control treaties.[64] But for now, such agreements do not exist, and international law currently lacks legally binding or sufficiently enforceable norms.

As the article previously established that the formation of internet bubbles will greatly affect cyber proxy wars, the authors further argue that domestic laws will greatly shape these bubbles and their creation. As laid down, cyberspace consists of several different layers, each of them facilitating the next.[65] The state could generally govern any layer of the internet in its effort to create an internet bubble: it could control the physical infrastructure; control the logic or information layer; or directly regulate end users, much like in the example of Russia's *RuNet*.

Regarding cyber proxy wars, however, the regulation of end-users might not advance the state's objective to a great extent since these users might not be under a state's jurisdiction. They might, for instance, be foreign agents residing outside of it, and pursuing them would prove highly difficult, expensive, or generally ineffective. The state might therefore be left focusing its efforts mainly on the first three layers. Control of the first three layers—the infrastructure, logic, and information layers—is gained mostly through control of those who operate and maintain them: the online intermediaries (i.e., internet service providers [ISPs] who maintain the different infrastructure and online platforms).

Thus, the creation of internet bubbles relies heavily on how online intermediaries are regulated by the state—or how much the government can control and command them. This is where the Chinese-Russian and American approaches greatly differ, walking down diverging paths. The American legal system generally abstains from imposing any form of direct or indirect liability on online intermediaries. Under the American liability regime—it would be highly difficult, if not almost impossible, to mandate or control an American internet bubble for almost any reason, let alone to combat cyber proxy wars. One of the main reasons that the United States might be in a severe disadvantage in defending against cyber proxy wars, or properly responding to them, is that its legal regime makes it highly difficult to form internet bubbles. The U.S. approach largely relies on the market to self-regulate, based partially on Adam Smith's monumental notion of the "invisible hand."[66]

This approach is currently articulated under the Communication Decency Act (CDA). While this 1996 act was originally aimed at curbing online pornography, large parts of it were deemed an unconstitutional infringement on free speech by the U.S. Supreme Court and thus struck down, all but keeping one highly influential section—known as section 230.[67] Under the current prevalent interpretation of section 230, the CDA grants broad immunity for online intermediaries, and they generally are not liable for third-party content they host.[68] This exemption from liability generally grants broad immunity to any

ISP, regardless of the legality or legitimacy of the content hosted, while ISPs are also entitled to remove offensive or otherwise objectionable content from their platforms when acting in "good faith." [69]

It is thus challenging from an American perspective to regulate ISPs, and thereby the content that is present in online platforms, as long as such content is considered protected speech under the current Supreme Court's libertarian stance on the First Amendment, and as long as section 230 remains intact. Such regulation might impede free speech, guaranteed by the First Amendment, which is considered a highly important human right for various reasons, but perhaps mostly plays an important role in protecting democracy.[70] In other words, the U.S. approach would generally abstain from obliging ISPs to act as censors, not because all content must remain online at any cost, but because the government should not, and legally speaking cannot compel intermediaries to make such judgments, at least for the time being.[71]

However, many argue that it is both possible and desirable to amend section 230 and that, at the very least, some internet intermediaries should bear legal liability in some instances.[72] The United States had, in fact, experienced some recent changes in its view regarding intermediary liability when former president Donald J. Trump claimed he intends to create a so-called "internet kill-switch" for national security purposes.[73] In the context of the COVID-19 pandemic, Twitter had begun tagging some of the tweets made by Trump as factually false, while adding informational links to news articles.[74] In response, Trump signed an executive order that allowed the Federal Communications Commission to craft rules that will govern internet intermediaries under section 230.[75] Even with this new order, and after the Capitol riot on 6 January 2021, Twitter had permanently banned Trump's account over the "risk of further incitement of violence," and Trump's attempt to file a lawsuit against them for doing so eventually failed.[76] But more importantly, such regulation did not last long, as U.S. president Joseph R. Biden decided to revoke Trump's executive order that targeted section 230.[77]

Still, section 230 is not a constant. Reshaping or even revoking section 230's safeguards to intermediaries might enjoy bipartisan support, at least at this time, as reflected in the view of President Biden, among other U.S. senators, and there are few proposed bills that aim to do so.[78] Other bills might also directly tackle foreign disinformation on social media, adding some exceptions to section 230.[79] Currently, however, more than 25 years after its enactment, section 230 remains intact, and other legislative attempts to limit its scope failed for now.

There are many facets to choosing a liability model, and it greatly depends on the legal jurisdiction in question. It is not our purpose here to discuss which liability model is more optimal in general (if such normative evaluation could even be objective), or to show how choosing one model would impact human rights and liberties differently than another.[80] Rather, this article aims to exemplify how the American liability regime comes into play within the context of

cyber proxy wars, and to further shed light on the potential future path that those susceptible to such wars might take (or are already taking)—if they keep their current legal approach to intermediaries.[81]

Even if section 230 changes over time, it is difficult to see how the United States would directly create an internet bubble, as such a move stands in stark contrast to the American notion of free speech. Any form of regulation that will attempt to create a U.S. internet bubble by infringing upon free speech, through any means of controlling one or more of the different layers, will very likely be constitutionally challenged, and thus subjected to strict scrutiny—the highest and almost impossible threshold that the state must pass to prove the lawfulness of such regulation.[82] Of course, if the president declares a "war or threat of war" or even "a state of public peril," then they might be able to exercise various authorities such as taking control over "wire communications" under a 1934 act—including the internet.[83] Therefore, at least in its territory, the U.S. president might be able to control the internet without even adhering to Congress.[84] With that being said, it is highly unlikely that this authority will be easily exercised, especially not in the context of proxy wars, cybernetic or not.

There could be some other forms of regulating intermediaries that could potentially also affect cyber proxy wars. One example could be using advertisement rules or other forms of mandatory disclosures regarding those who purchase online ads.[85] Following the 2016 U.S. election interference, some states had in fact passed election laws that obligate ISPs to disclose information about the identity of those who purchased political ads.[86] But, aside from potential practical difficulties, like that of acts of concealment by an actor within a proxy war, laws of this nature (if challenged in court) will not likely be deemed constitutional as they are considered compelled speech, which could also infringe upon the First Amendment.[87] Moreover, in the context of this article, such disclosure laws will only tackle potential cyber proxy wars from a very limited aspect—serving a narrow solution to a much wider challenge.

Means to Preserve Sovereignty in Cyberspace

The United States might make use of other potential means, which do not include the creation of an internet bubble per se, in its effort to preserve sovereignty and resist cyber proxy warfare. One means is to actively restrict transactions between U.S. entities and parent Russian or Chinese companies, essentially banning their use in the United States. Former president Trump had attempted to do so with ByteDance and Tencent (the parent companies of TikTok and WeChat, respectively).[88] The problem here is that such means are highly limited as it only targets a fraction of intermediaries and is less relevant for U.S. companies as long as section 230 remains intact.

A more plausible means is that of the market self-regulating. Under this argument, it is upon private actors—like ISPs—to regulate the kind of harmful conduct involved in cyber proxy war campaigns. In other words, the American approach, which created the governmental barrier of noninterference within

an invisible hand perspective, might also drive the market to respond to cyber proxy wars, and thus, even without forming an internet bubble, mitigate at least some of the risks to the United States from them.

The question of whether this approach advances the rationales of free speech or not could be debatable, but it is beyond the scope of this article.[89] Here, in this context, the authors merely strive to show how these constitutional barriers and the legal regime in the United States could be used as a tool by other jurisdictions within these cyber proxy wars. The problem with market self-regulation are its numerous potential failures. It is perceived as unlikely that for-profit companies will self-regulate their platforms, even despite ongoing cyber proxy wars, unless such self-regulation proves economically beneficial for them.

The market, however, could be nudged to combat these wars, at least to some extent. The government or other policy makers could, for instance, warn companies that they *might* be regulated if they do not act in a self-regulatory manner, which will, at the very least, reduce the scope of these proxy wars and their perceived damages and negative effects. Consider the congressional response to the Facebook-Cambridge Analytica data breach—an example of how one might use ISPs to influence politics (and advance their own agenda)—that could demonstrate how Congress might pressure or nudge online intermediaries to act without the need for direct legislation or regulation.[90]

Furthermore, these online platforms often aid the government under what is termed as public-private partnerships (PPPs)—collaborations between governments and online intermediaries in managing online behavior.[91] The authors have witnessed such PPPs in American history and more closely within some of the secret surveillance programs that Edward Snowden revealed in 2013.[92] If properly incentivized to "voluntarily" assist the government, online intermediaries might assume a role as a cyber proxy for governance responses to cyber proxy wars.[93] In the post-Snowden era, Congress further granted authorization for ISPs to monitor their information systems, operate defensive measures, and share "cyber threat indicators" or "defensive measures" for a cybersecurity purpose.[94]

But all in all, the United States might just attempt to respond to cyber proxy wars by utilizing other means at hand, which might prove simply more feasible. It might deploy its political, economic, or otherwise kinetic strength to directly or indirectly combat those who operate against it within the cyber domain.[95] The limits of such means, however, lie within those sovereign powers who are less reliant or dependent on American political or economic support. While the United States, on the other hand, might find itself heavily reliant on foreign powers who may already have an internet bubble in place and therefore places it in a severe disadvantage in fighting the cyber proxy wars.

Still, even without resorting to kinetic wars, the United States might simply act aggressively within the cyber realm directly against its adversaries.[96] It might also begin to heavily regulate what enters its kinetic and digital borders to some extent. In the physical realm, the United States could respond to proxy wars by

banning specific imports.[97] In the digital realm, it might regulate end-users by banning specific apps or regulate the market by banning or otherwise restricting transactions between U.S. companies and foreign ones—relying on national security arguments. Such tactics had been taken regarding the Chinese-owned apps TikTok and WeChat.[98]

Eventually, without a significant shift in the American perception of online intermediaries' regulation, the solution to cyber proxy wars will probably lie elsewhere than with the formation of a U.S. internet bubble. The United States will likely continue to have an open internet, as opposed to an isolated bubble, albeit with independent market forces, as part of a notion of self-regulation, likely to intervene more to address harmful effects. Under corporate social responsibility or other incentives, we are likely to see platform governance on the rise, which could eventually include a direct response to cyber proxy wars. We have already begun to witness how some social media companies, like Facebook or Twitter, are forming their own oversight boards, intended to make principal decisions regarding content moderation.[99]

And truly, corporate governance is on the rise and might prove useful as a shield against foreign influence. Content moderation and the removal of accounts that are linked to domestic political influence is constantly occurring around the world, such as in Ukraine, Iran, Russia, to name but a few examples.[100] These platforms are already shaping the scope of national security in many countries.[101] On the other hand, there are still limits for such influence, especially when for-profit companies wish to stay in the market. To exemplify, when the Russian government was dissatisfied with Twitter's content removal procedures, it almost immediately slowed it down for users.[102] Thus, one of the problems with platform governance is that eventually these for-profit companies act to increase their revenues outside of the United States as well.

Only time will tell whether such an approach could work for the United States. Perhaps internet bubbles make a more direct and efficient way of handling cyber proxy wars. But they do come with costs in terms of human rights and liberties, and if ISPs do a rather decent job in combating these proxy wars—even if not as good as with a strict liability regime in place—then this trade-off might prove worthwhile. It would be unfortunate if nondemocratic states will eventually misuse democratic and liberal values against those same states who attempt to safeguard them.[103]

Finally, it is important to note that while the United States and Russia-China serve as opposing examples for domestic law regimes, other legal regimes could be placed along the spectrum between a liberal democracy and a nondemocracy.[104] Indeed, the greater fear and challenge might lie within those other legal jurisdictions that desire to implement a regulatory regime rather similar to that of the United States, but lack any meaningful other powers—be it political, economic, military, or otherwise—to challenge and engage with cyber proxy wars without adhering to direct legislation or intervention. Lacking strong constitutional safeguards such as those of the United States, countries may resort to

legislation and deeper intervention in cyber space, and internet bubbles might form in many countries as a defensive measure. This could, in turn, eventually negatively affect online free speech rather dramatically, and subsequently affect and perhaps even threaten democracy and liberalism itself. This concern grows more severe as we move further away from the United States' end of the political spectrum toward that of Russia, China, North Korea, and their likes.

Conclusion

The authors began their article with Churchill's note on the Soviet Iron Curtain, which existed to serve the Soviet regime and enable it to both control its domestic affairs and avoid extensive international influence. With the proliferation of the internet, along with a toothless international law system regarding cyberattacks, influence, and espionage, countries now seek to gain more control with a contemporary iron curtain of their own, thus gaining cyber sovereignty meant to avoid or resist foreign influence and intervention. Furthermore, cyber deterrence or retaliation is becoming almost impossible due to the practice of cyber proxy warfare—cyberspace is an evasive and anonymous proxy. Countries like Russia, China, Iran, and North Korea could resist foreign influence by creating their sovereign internet bubbles and gain power by influencing countries that lack such bubbles. These internet bubbles could cover all of the layers that make cyberspace what it is: Russia's *RuNet* experiment, China's sovereign internet, and North Korea's *Kwangmyong* intranet project all exemplify the bubbles created by the physical layer (aimed to protect hacking and eavesdropping), the logic layer (aimed to protect from computational manipulations), the information layer (aimed to protect from disinformation or malicious software), and the user layer (aimed to protect from manipulative users).

To some extent, liberal democracies such as the United States are in a severe disadvantage in this regard. That is, while Russia and China, lacking meaningful legal and constitutional restraints, are dealing with the deficiencies of the cyber domain and the lack of binding international law, the United States is left behind due to its democratic values and governance. Cases like the 2016 presidential elections intervention by Russia or COVID-19 disinformation might all serve as examples in which the United States failed to properly protect itself from foreign threats. In contrast, the conflict between Ukraine and Russia has demonstrated, at least as of this writing (May 2023) that restricting and regulating domestic internet and media are important strategic tools to undermine foreign propaganda and antigovernmental sentiments. Still, Moscow is not safe from its own domestic arena as internal rifts and power struggles intensify in Russia but many are not directly connected to Western propaganda. The practice of cyber proxy warfare might further allow foreign powers to attack their adversaries through targeting nonstate entities and institutions associated with them, as exemplified by the Sony Pictures case.

How can the United States and other similar democracies protect themselves and remain sovereign in the age of (dis)information and cyber warfare?

As the authors argue and predict throughout the article, if the United States will not eventually "catch up" with internet-related restrictions to stand strong against its global adversaries, it will be up to private intermediaries to self-regulate such threats. The United States is not likely to form a hermetic internet bubble, but if platform governance fails, it might strive to find other ways to influence ISPs or use other means to aid them in the fight over sovereignty and control. As the authors suggest, the practice of cyber proxy warfare has in fact influenced international orders and norms. Now, the only questions are how and whether they will succeed. Otherwise, perhaps even true liberal democracies will begin to form their own internet bubbles and the internet will transform into something different altogether.

Endnotes

1. Winston Churchill, address at Westminster College, Fulton, Missouri, 5 March 1946.
2. The term *internet bubbles* was coined by the authors in the context of cyber sovereignty. The term mostly exists in the context of the internet bubble of the 1990s, a.k.a. the dot-com bubble. It can also refer to a generation that grew up in a digital bubble and social media.
3. Richard J. Harknett and Joseph S. Nye Jr., "Deterrence and Dissuasion in Cyberspace," *International Security* 41, no. 3 (2017): 44–71, https://doi.org/10.1162/ISEC_c_00290; and Laura Rosenberger, "Making Cyberspace Safe for Democracy—The New Landscape of Information Competition," *Foreign Affairs*, no. 99 (May/June 2020): 146–59.
4. The Internet Backbone is an infrastructure of mainly fiber optic cables that connects multiple nodes and effectively connect various networks worldwide. See Nazli Choucri and David D. Clark, *International Relations in the Cyber Age: The Co-Evolution Dilemma* (Cambridge, MA: MIT Press, 2018), 33–65, 101–22.
5. Kristina Daugirdas and Julian Davis Mortenson, "United States Responds to Alleged North Korean Cyber Attack on Sony Pictures Entertainment," *American Journal of International Law* 109, no. 2 (2015): 419.
6. Andreas Osiander, "Sovereignty, International Relations, and the Westphalian Myth," *International Organization* 55, no. 2 (2001), 251–87; Gene M. Lyons and Michael Mastanduno, eds., *Beyond Westphalia?: State Sovereignty and International Intervention* (Baltimore, MD: Johns Hopkins University Press, 1995); and Robert O. Keohane and Joseph S. Nye Jr., "Power and Interdependence in the Information Age," *Foreign Affairs* 77, no. 5 (1998): 81–94.
7. Rawi Abdelal and Adam Segal, "Has Globalization Passed Its Peak?," *Foreign Affairs* (2007): 103–14; and Josh Salisbury, "Omicron: More Countries Restrict Foreign Travellers over New Variant Fears," *Evening Standard*, 30 November 2021.
8. Alexander Lanoszka, "Disinformation in International Politics," *European Journal of International Security* 4, no. 2 (2019): 227–48, https://doi.org/10.1017/eis.2019.6.
9. Aysegul Aydin, *Foreign Powers and Intervention in Armed Conflicts* (Stanford, CA: Stanford University Press, 2012), 1–5.
10. Osiander, "Sovereignty, International Relations, and the Westphalian Myth."
11. Frederic S. Pearson, "Foreign Military Interventions and Domestic Disputes," *International Studies Quarterly* 18, no. 2 (1974): 259–90.
12. Joseph S. Nye Jr., "Soft Power," *Foreign Policy*, no. 80 (1990): 153–71, https://doi.org/10.2307/1148580; Joseph S. Nye Jr., "Get Smart: Combining Hard and Soft Power," *Foreign Affairs* 88, no. 4 (July/August 2009): 160–63; Ernest J. Wilson III, "Hard Power, Soft Power, Smart Power," *Annals of the American Academy of Political and Social Science* 616, no. 1 (2008): 110–24, https://doi.org/10.1177/0002716207312618; Ra-

chelle Faust, " 'Sharp Power': Rising Authoritarian Influence," National Endowment for Democracy, 5 December 2017; and Christopher Walker and Jessica Ludwig, "The Meaning of Sharp Power: How Authoritarian States Project Influence," *Foreign Affairs*, 16 November 2017.

13. Harknett and Nye, "Deterrence and Dissuasion in Cyberspace," 44–71.
14. Karl W. Deutsch, "External Involvement in Internal War," in *Internal War: Problems and Approaches*, ed. Harry Eckstein (New York: Free Press of Glencoe, 1964); Andrew Mumford, *Proxy Warfare* (Cambridge, UK: Polity Press, 2013), 1–29; and Andrew Mumford, "Proxy Warfare and the Future of Conflict," *RUSI Journal* 158, no. 2 (2013): 40–46, https://doi.org/10.1080/03071847.2013.787733.
15. Harknett and Nye, "Deterrence and Dissuasion in Cyberspace"; Michael N. Schmitt and Liis Vihul, "Proxy Wars in Cyberspace: The Evolving International Law of Attribution," *Fletcher Security Review*, no. 1 (2014): 53–72; and Michael N. Schmitt, ed., *Tallinn Manual 2.0 on the International Law Applicable to Cyber Warfare* (Cambridge, UK: Cambridge University Press, 2017).
16. Tim Maurer, " 'Proxies' and Cyberspace," *Journal of Conflict and Security Law* 21, no. 3 (2016): 383–403, https://doi.org/10.1093/jcsl/krw015; Pnina Shuker and Lev Topor, "Russian Influence Campaigns Against NATO in the Baltic Region: Spread of Chaos and Divide et Impera," in *The Russian Federation in Global Knowledge Warfare*, ed. Holger Mölder et al. (Cham, Switzerland: Springer, 2021), 295–314, https://doi.org/10.1007/978-3-030-73955-3_15.
17. Deutsch, "External Involvement in Internal War"; Mumford, *Proxy Warfare*, 1–29; Tim Maurer, *Cyber Mercenaries: The State, Hackers, and Power* (Cambridge, UK: Cambridge University Press, 2018), 3–28; and Andreas Krieg and Jean-Marc Rickli, "Surrogate Warfare: The Art of War in the 21st Century?," *Defence Studies* 18, no. 2 (2018), 113–30, https://doi.org/10.1080/14702436.2018.1429218.
18. Jason Andress and Steven Winterfeld, *Cyber Warfare: Techniques, Tactics and Tools for Security Practitioners* (Amsterdam, Netherlands: Syngress, 2013), 1–14.
19. Herbert Lin and Amy Zegart, eds., *Bytes, Bombs, and Spies: The Strategic Dimensions of Offensive Cyber Operations* (Washington, DC: Brookings Institute Press, 2019), 1–8.
20. There is, however, an available domestic intranet, known as *Kwangmyong*, which allows access to domestic websites and emails, but these are controlled by the government. "North Korea: Connection Denied: Restrictions on Mobile Phones and Outside Information in North Korea," Amnesty International, 9 March 2016; "North Korea: Tightened Controls on Communications with the Outside World Leave Families Devastated," Amnesty International, 7 March 2016; and Cheng Chen, Kyungmin Ko, and Ji-Yong Lee, "North Korea's Internet Strategy and Its Political Implications," *Pacific Review* 23, no. 5 (2010): 649–70, https://doi.org/10.1080/09512748.2010.522249.
21. Yochai Benkler, *The Wealth of Networks: How Social Production Transforms Markets and Freedom* (New Haven, CT: Yale University Press, 2006), 23; Choucri Nazli and David D. Clark, "Integrating Cyberspace and International Relations: The Co-Evolution Dilemma" (ECIR Working Paper No. 2012-3, MIT Political Science Department, 6 November 2012); and Nazli and Clark, *International Relations in the Cyber Age*, 33–65, 101–22.
22. Jonathan Masters, "Ukraine: Conflict at the Crossroads of Europe and Russia," Council on Foreign Relations, 14 February 2023; and Jon Bateman, Nick Beecroft, and Gavin Wilde, "What the Russian Invasion Reveals About the Future of Cyber Warfare," Carnegie Endowment, 19 December 2022.
23. Anton Troianovski and Valeriya Safronova, "Russia Takes Censorship to New Extremes, Stifling War Coverage," *New York Times*, 4 March 2022.
24. Tetyana Klug and Rachel Baig, "Fact Check: Russia's Disinformation Campaign Targets NATO," DW, 13 February 2023.
25. Jon Bateman, "Russia's Wartime Cyber Operations in Ukraine: Military Impacts, Influences, and Implications," Carnegie Endowment, 16 December 2022.
26. Stephen Blank, "Threats to and from Russia: An Assessment," *Journal of Slavic Military Studies* 21, no. 3 (2008): 491–526, https://doi.org/10.1080/13518040802313746;

and Andrea Shalal, "U.S. Air Force Leader Sees Russia as 'Biggest Threat'," Reuters, 9 July 2015.

27. *Russian Strategic Intentions: A Strategic Multilayer Assessment (SMA) White Paper* (Washington, DC: Department of Defense, May 2019).
28. Совет Федерации, "Концепция Стратегии Кибербезопасности Российской Федерации" [Russian] (The Concept of the Cyber Security Strategy of the Russian Federation); and Michael Connell and Sara Vogler, "Russia's Approach to Cyber Warfare," CNA, March 2017.
29. Russia's successful attempt to unplug from the global internet was an effort to test an internal (or domestic) internet that would allow internal communication without being connected to external information flow and without relying on external services. This was done by unplugging Russia from submarine fiber optic cables that connect Russia to the global internet. See Jane Wakefield, "Russia 'Successfully Tests' Its Unplugged Internet," BBC, 24 December 2019.
30. "Cable Compendium: A Guide to the Week's Submarine and Terrestrial Developments," CommsUpdate, 19 February 2021.
31. Daria Litvinova, "Human Wrongs: How State-Backed Media Helped the Kremlin Weaponise Social Conservatism" (Reuters Institute Fellowship Paper, Oxford, UK, University of Oxford, July 2018); and Jill Dougherty, "How the Media Became One of Putin's Most Powerful Weapons," *Atlantic*, 21 April 2015.
32. The Yarovaya Laws (Russian: Закон Яровой) are a set of bills (374-FZ, 375-FZ) passed in 2016 that amend preexisting counterterrorism laws and deal with the regulation and monitoring of cyberspace. The bills compel the telecommunication industry to allow Russian authorities access to their data by providing them the encryption and decryption keys necessary to decode and monitor online data. For the bill amending counterterrorism measurements, see: Законопроект № 1039149-6 [Russian]. The Russian Sovereign Internet Law provides Moscow the capacity to turn off connections within Russia or with the worldwide web entirely in case of an emergency. See Isabelle Khurshudyan, "The Kremlin Is Notorious for Global Meddling Online. But Controlling Cyberspace at Home Has Been Trickier," *Washington Post*, 26 January 2020. SORM—System for Operative Investigative Activities is Russia's mass communications surveillance. Russian law gives Russia's security service, the FSB, the authority to collect, analyze, and store all data that was transmitted or received on Russian networks. See James A. Lewis, "Reference Note on Russian Communications Surveillance," Center for Strategic and International Studies, 18 April 2014; and Anton Zverev and Gabrielle Tétrault-Farber, "Putin Signs Law Making Russian Apps Mandatory on Smartphones, Computers," Reuters, 2 December 2019.
33. On 18 March 2019, Vladimir Putin signed two laws passed by the Russian Federal Assembly aimed at countering the creation and dissemination of fake news. The laws establish fines for knowingly spreading fake news and procedures for internet service providers to block access to websites disseminating fake news. The laws are the Federal Law on Amending Article 15-3 of the Federal Law on Information, Information Technologies and Protection of Information (№ 31-ФЗ) and the Federal Law on Amending the Code of Administrative Violations (№ 27-ФЗ). Mark Bennetts, "Russia Passes Law to Jail People for 15 Days for 'Disrespecting' Government," *Guardian*, 6 March 2019. The Russian Foreign Agent Law applies to individuals who distribute online information and receive funds from foreign sources. Individuals who distribute foreign media can also be labeled as foreign agents. Russian citizens and foreign visitors can be labeled as foreign agents. Интерфакс, "Путин подписал закон о физлицах-иноагентах," 2 декабря 2019.
34. Lev Topor and Alexander Tabachnik, "Russian Cyber Information Warfare," *Journal of Advanced Military Studies* 12, no. 1 (Spring 2021): 112–27, https://doi.org/10.21140/mcuj.20211201005.
35. Rob Picheta, " 'It's All a Lie': Russians Are Trapped in Putin's Parallel Universe. But Some Want Out," CNN, 27 February 2023.
36. Samuel Charap, John Drennan, and Pierre Noël, "Russia and China: A New Model of

Great-Power Relations," *Survival* 59, no. 1 (2017): 25–42, https://doi.org/10.1080/00396338.2017.1282670; and Michael Mastanduno, "Partner Politics: Russia, China, and the Challenge of Extending US Hegemony after the Cold War," *Security Studies* 28, no. 3 (2019): 479–504, https://doi.org/10.1080/09636412.2019.1604984.

37. Nina Hachigian, "China's Cyber-Strategy," *Foreign Affairs* 80, no. 2 (March/April 2001): 118–33; and William T. Dowell, "The Internet, Censorship, and China," *Georgetown Journal of International Affairs* 7, no. 2 (2006): 111–19.
38. According to the privacy-oriented site vpnMentor, China blocks more than 8,000 websites: Ariel Hochstadt, "The Complete List of Blocked Websites in China & How to Access Them," vpnMentor, 16 June 2020.
39. Niels N. Schia and Lars Gjesvik. "China's Cyber Sovereignty," Norwegian Institute of International Affairs, Policy Brief no. 2, 17 March 2017.
40. Amy Chang, *Warring State: China's Cybersecurity Strategy* (Washington, DC: Center for a New American Security, 2014).
41. "China Announces Cybersecurity Strategy," State Council, People's Republic of China, 27 December 2016; Jyh-An Lee, "Hacking Into China's Cybersecurity Law," *Wake Forest Law Review* 53, no. 1 (2018): 57; and Paul Rosenzweig, "China's National Cybersecurity Strategy," *Lawfare* (blog), 27 December 2016. For a detailed review of Chinese internet regulation, see Samm Sacks and Manyi K. Li, "How Chinese Cybersecurity Standards Impact Doing Business in China," Center for Strategic & International Studies, 2 August 2018.
42. Chang, "Warring State."
43. Gregory J. Moore, "Huawei, Cyber-sovereignty and Liberal Norms: China's Challenge to the West/Democracies," *Journal of Chinese Political Science*, no. 28 (2022): 1–17, https://doi.org/10.1007/s11366-022-09814-2.
44. Sacks and Li, "How Chinese Cybersecurity Standards Impact Doing Business in China."
45. VPN is a virtual private network; it extends a private network on top of a public network and enables users to use the web as if they were connected to the internet from a different location.
46. Mark A. Lemley, "Rationalizing Internet Safe Harbors," *Journal on Telecommunication & High Technology Law* 6, no. 1 (2007): 115–16; and Michael L. Rustad and Thomas H. Koenig, "Rebooting Cybertort Law," *Washington Law Review* 80, no. 2 (2005): 392.
47. Malcolm N. Shaw, *International Law*, 5th ed. (Cambridge, UK: Cambridge University Press, 2012), chaps. 1–2.
48. Harknett and Nye, "Deterrence and Dissuasion in Cyberspace," 47.
49. The right of self-defense is an inherent one under customary international law, as well as under the UN Charter. See United Nations, *Charter of the United Nations* (1945), art. 2(4), 51; and Shaw, *International Law*, 1,024–32.
50. An "armed attack" would include "the sending by or on behalf of a state of armed bands or groups which carry out acts of armed force of such gravity as to amount to an actual armed attack." The involvement of the state would be considered in order to hold it liable and legitimate an act of self-defense against it. See Shaw, *International Law*, 1,026–28. The consequences of a cyber act and its surrounding circumstances will determine whether it had crossed the "use of force" threshold. See Schmitt, *Tallinn Manual 2.0*, 328–56.
51. Robert Heinsch, "Conflict Classification in Ukraine: The Return of the 'Proxy War'?," *International Law Studies*, no. 91 (2015): 340.
52. Brittany Benowitz and Tommy Ross, "Time to Get a Handle on America's Conduct of Proxy Warfare," *Lawfare* (blog), 9 April 2020; and *The Legal Framework Regulating Proxy Warfare* (Chicago, IL: American Bar Association, 2019).
53. *The Legal Framework Regulating Proxy Warfare*; and International Law Commission, *Draft Articles on Responsibility of States for Internationally Wrongful Acts* (New York: United Nations, 2001), art. 8.
54. International Law Commission, *Draft Articles on Responsibility of States for Internation-*

ally Wrongful Acts. See also Harknett and Nye, "Deterrence and Dissuasion in Cyberspace," 44–71.

55. Deutsch, "External Involvement in Internal War"; and Mumford, *Proxy Warfare*.
56. General Assembly, United Nations, *Resolution 2625 (XXV), Declaration on Principles of International Law Concerning Friendly Relations and Co-Operation Among States in Accordance with the Charter of the United Nations* (24 October 1970); Clare Sullivan, "The 2014 Sony Hack and the Role of International Law," *Journal of National Security Law & Policy*, no. 8 (2016): 455–59; United Nations, *Charter of the United Nations*, art. 2(4), 51; and Sullivan, "The Sony Hack," 455 and onward. These would include, inter alia, just cause, proportionality, right intention and authority, last resort, and reasonable chance of success. See Anthony Pfaff, "Proxy War Ethics," *Journal of National Security Law and Policy*, no. 9 (2017): 308.
57. Sullivan, "The Sony Hack," 455 and onwards; and Schmitt, *Tallinn Manual 2.0*, 1–2.
58. Schmitt, *Tallinn Manual 2.0*, 1–2. See rule 69 that states that "a cyber operation constitutes a use of force when its scale and effects are comparable to non-cyber operations rising to the level of a use of force."
59. Schmitt, *Tallinn Manual 2.0*, 1–2, rules 66 (Prohibition of Intervention), 68–70 (Prohibition of the Use of Force), and 71 (Self-Defence Against Armed Attack).
60. United Nations, *Charter of the United Nations*, art. 2(4), 51; and Sullivan, "The Sony Hack," 455n128 and onward.
61. Drew Marvel, "Protecting the States from Electoral Invasions," *William & Mary Bill of Rights Journal* 28, no. 1 (October 2019): 216.
62. Schmitt, *Tallinn Manual 2.0*, rule 71.
63. Marvel, "Protecting the States," 220; and Sullivan, "The Sony Hack."
64. Anton Troianovski and David E. Sanger, "Putin Wants a Truce in Cyberspace—While Denying Russian Interference," *New York Times*, 25 September 2020.
65. These layers comprise of the physical, logic, information, and end-user. See Benkler, *The Wealth of Networks*.
66. See Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations* (1776; repr., Amsterdam, Netherlands: MetaLibri, 2007).
67. See Reno v. American Civil Liberties Union, 521 U.S. 844 (1997).
68. Communications Decency Act, 47 U.S.C. § 230 (2006).
69. For more on section 230, see Danielle K. Citron and Benjamin Wittes, "The Internet Will Not Break: Denying Bad Samaritans Sec. 230 Immunity," *Fordham Law Review* 86, no. 2 (November 2017): 401–24.
70. Jack M. Balkin, "Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society," *New York University Law Review* 79, no. 1 (2004): 34. The First Amendment currently protects speech, even if such speech is false. See United States v. Alvarez, 132 S. Ct. 2537, 2539, 2544 (2012).
71. Corey Omer, "Intermediary Liability for Harmful Speech: Lessons from Abroad," *Harvard Journal of Law & Technology* 28, no. 1 (Fall 2014): 315.
72. Cass R. Sunstein, "The First Amendment in Cyberspace," *Yale Law Journal*, no. 104 (1995); and Rebecca Tushnet, "Power Without Responsibility: Intermediaries and the First Amendment," *George Washington Law Review*, no. 76 (2008).
73. Sean Lawson, "The Law that Could Allow Trump to Shut Down the US Internet," *Forbes*, 2 December 2016.
74. "Twitter Tags Trump Tweet with Fact-Checking Warning," BBC, 27 May 2020.
75. John T. Bennet, "Trump Signs Controversial Executive Order that Could Allow Federal Officials to Target Twitter, Facebook and Google," *Independent*, 28 May 2020.
76. Kate Conger and Mike Isaac, "Twitter Permanently Bans Trump Capping Online Revolt," *New York Times*, 8 January 2021; and Adi Robertson, "Judge Dismisses Donald Trump's Twitter Ban Lawsuit," *Verge*, 6 May 2022.
77. Kim Lyons, "Biden Revokes Trump Executive Order that Targeted Section 230," *Verge*, 15 May 2021.
78. Editorial Board, "Opinion: Joe Biden—Former Vice President of the United States,"

New York Times, 17 January 2020; "Senate Panel Mulls Revoking Immunity, Citing COVID Scams Online," Reuters, 21 April 2021; and Oscar Gonzalez, "Bill Unveiled to Reduce Section 230 Protections for Social Media Companies," CNET, 5 February 2021.

79. Maggie Miller, "House Lawmakers Reintroduce Bipartisan Bill to Weed out Foreign Disinformation on Social Media," *Hill*, 22 January 2021.
80. Jonathan Zittrain, "A History of Online Gatekeeping," *Harvard Journal of Law & Technology*, no. 19 (2006).
81. For more on section 230, see Eric Goldman, "An Overview of the United States' Section 230 Internet Immunity," in *The Oxford Handbook of Online Intermediary Liability*, ed. Giancarlo Frosio (forthcoming).
82. Strict scrutiny requires demonstrating that the regulation is necessary to a compelling state interest; that it is narrowly tailored to achieve such purpose, and that the regulation uses the least restrictive means to achieve its purposes.
83. See 47 U.S.C. § 606(d) (1934).
84. Maura K. Perri, "Build the Fire-Wall: Potential Dangers to Internet Freedom under the Trump Administration," *Duquesne Law Review* 57, no. 1 (Winter 2019): 195.
85. Steven T. Dennis, "Senators Propose Social-Media Ad Rules after Months of Russia Probes," *Bloomberg*, 19 October 2017.
86. See, for instance, in Maryland: MD. Code Ann., Elec. Law § 13-405 (2020).
87. Eugene Volokh, "The Law of Compelled Speech," *Texas Law Review*, no. 95 (2018). Specifically, the Fourth Circuit recently held that Maryland law was unconstitutional on these grounds. For more on the debate of whether such move is constitutional under U.S. law, see "Notes: Two Models of the Right to Not Speak," *Harvard Law Review,* no. 113 (2020).
88. Adi Robertson, "The Big Legal Questions Behind Trump's TikTok and WeChat Bans," *Verge*, 10 August 2020.
89. Marcelo Thompson, "Beyond Gatekeeping: The Normative Responsibility of Internet Intermediaries," *Vanderbilt Journal of Entertainment and Technology Law* 18, no. 4 (Summer 2016).
90. Carole Cadwalladr and Emma Graham-Harrison, "Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach," *Guardian*, 17 March 2018; Jamie Condliffe, "How to Fix Social Media's Big Problems? Lawmakers Have Ideas," *New York Times*, 30 July 2018; and Lauren Feinr, "Senators Threaten to Regulate Encryption If Tech Companies Won't Do It Themselves," CNBC, 10 December 2019.
91. Niva Elkin-Koren and Eldar Haber, "Governance by Proxy: Cyber Challenges to Civil Liberties," *Brooklyn Law Review*, no. 82 (2017); Madeline Carr, "Public–Private Partnerships in National Cyber-Security Strategies," *International Affairs* 92, no. 1 (2016): 43–62; and Kristoffer K. Christensen and Karen L. Petersen, "Public–Private Partnerships on Cyber Security: A Practice of Loyalty," *International Affairs* 93, no. 6 (2017): 1435–52, https://doi.org/10.1093/ia/iix189.
92. Glenn Greenwald and Ewen MacAskill, "NSA Prism Program Taps into User Data of Apple, Google, and Others," *Guardian*, 7 June 2013.
93. Elkin-Koren and Haber, "Governance by Proxy."
94. See Consolidated Appropriations Act of 2016, Pub. L. No. 114-113, 129 Stat. 2242 (2016).
95. Kate Conger, "Twitter Removes Chinese Disinformation Campaign," *New York Times*, 11 June 2020; Mark Scott, "Facebook Slaps Labels on Chinese and Russian State-Controlled Media Amid Anger over Donald Trump's Posts," *Politico*, 4 June 2020; and Ellen Nakashima, "Biden Administration Imposes Significant Economic Sanctions on Russia over Cyberspying, Efforts to Influence Presidential Election," *Washington Post*, 15 April 2021.
96. Zachary Evans, "Trump Gave CIA Authorization to Increase Aggressive Cyber Attacks: Report," *National Review*, 15 July 2020.

97. Ben Westcott, "New US Bill Could Ban Imported Chinese Goods from Xinjiang Amid Forced Labor Concerns," CNN: Politics, 12 March 2020.
98. After first contemplating a nationwide ban on their operation, the president opted instead for a future ban on two Chinese-owned apps, unless they are sold to American hands. See Nikki Carvajal and Caroline Kelly, "Trump Issues Orders Banning TikTok and WeChat from Operating in 45 Days If They Are Not Sold by Chinese Parent Companies," CNN, 7 August 2020; and John Koetsier, "TikTok to Be Banned in USA, Trump Announces," *Forbes*, 1 August 2020.
99. Kate Klonick, "Creating Global Governance for Online Speech: The Development of Facebook's Oversight Board," *Yale Law Journal*, no. 129 (2020); and Nick Clegg, "Welcoming the Oversight Board," Facebook, 6 May 2020.
100. Elizabeth Culliford, "Facebook Removes Ukraine Political 'Influence-for-Hire' Network," Reuters, 6 May 2021.
101. Cameron Jenkins, "Twitter Removes Hundreds of Accounts Tied to Iran, Russia," *Hill*, 24 February 2021; Culliford, "Facebook Removes Ukraine Political 'Influence-for-Hire' Network"; Igor Bonifacic, "Twitter Bans 100 Accounts Linked to Russian Troll Farms," engadget, 23 February 2021; and Elena Chachko, "National Security by Platform," *Stanford Technology Law Review* 25, no. 1 (2021).
102. Anton Troianovski and Andrew E. Kramer, "Russia Says It Is Slowing Access to Twitter," *New York Times*, 10 March 2021.
103. Some might even argue that democracy is deteriorating, both online and offline, thus it might be somewhat inevitable that states will intervene in internet regulation more closely. See Nathaniel Persily, "Can Democracy Survive the Internet?," *Journal of Democracy* 28, no. 2 (2017): 74–75.
104. Lingling Wei, "China Declared Its Russia Friendship Had 'No Limits.' It's Having Second Thoughts," *Wall Street Journal*, 3 March 2022.

The Nationalization of Cybersecurity
The Potential Effects of the *Cyberspace Solarium Commission Report* on the Nation's Critical Infrastructure

H. Chris Tecklenburg, JD/PhD;
and José de Arimatéia da Cruz, PhD/MPH

Abstract: The United States is susceptible to cyberattacks. The *Cyberspace Solarium Commission Report* provides several recommendations to prevent and respond to such attacks. However, many of these recommendations attempt to nationalize cybersecurity. This article presents a historical overview involving the Department of Homeland Security, the Cybersecurity and Infrastructure Security Agency, and the Commerce Clause, which outlines nationalization and its effects. It will note a similar trend for cybersecurity. Finally, the positive and negative consequences of nationalization are presented.

Keywords: cybersecurity, nationalization, homeland security, critical infrastructure, commerce, *Cyberspace Solarium Commission Report*

Introduction

According to the *Cyberspace Solarium Commission Report* (the report hereafter), the United States is susceptible to cyberattacks.[1] Several countries and nonstate actors have been identified as presenting the most credible threats to the United States, including China, Russia, Iran, and North Korea. To combat these cyber threats, the report outlines several recommendations. Many of these require cooperation from the states and private sector with the federal government. However, based on the report, some may wonder whether cooperation is possible, what the result of such cooperation is,

Dr. H. Chris Tecklenburg, JD, is an associate professor of political science at Georgia Southern University, Savannah Campus. Dr. José de Arimatéia da Cruz, MPH, is a professor of political science at Georgia Southern University, Savannah Campus, and a research professor at the U.S. Army War College Center for Strategic Leadership, Homeland Defense and Security Studies, Strategic Landpower Futures Group.

Journal of Advanced Military Studies vol. 14, no. 1
Spring 2023
www.usmcu.edu/mcupress
https://doi.org/10.21140/mcuj.20231401007

and whether the cooperation is constitutional. This article will address each of these issues and focus on how the recommendations lead to the nationalization of cybersecurity. This article will briefly conclude by noting such nationalization's potential advantages and disadvantages.

The Old Mantra: Is Public-Private Cyber Cooperation Possible?

The report's recommendations hinge on cooperation. As the report makes explicit, layered cyber deterrence, which includes cooperation between the government, private sector, and citizens, is the strategy adopted. Cooperation is required since, as the report recognizes, many "devices and applications, as well as the communications infrastructure on which they rely, are overwhelmingly controlled by the private sector."[2] Thus, effective cybersecurity requires participation by the private sector with the government. In addition, cooperation between the state and the federal government is also required.

To accomplish this goal of cooperation, the report presents several proposals involving various topics. First, the report focuses primarily on cooperation between the private sector and the federal government. This is mainly seen with the recommendation of creating a congressionally funded grant, labeled the National Cybersecurity Assistance Fund, which provides funding for the mitigation of a clearly defined risk where there is no market-based solution and where there is a clear need for federal involvement.[3]

Another private-sector recommendation involves the executive branch and suggests that "Congress should direct the executive branch to develop and maintain continuity of the economic planning in consultation with the private sector to ensure the continuous operation of critical functions of the economy in the event of a significant cyber disruption."[4] This proposal includes private sector entities responsible for critical infrastructure, such as power and electric systems, gas pipelines, and other items comprising national and international financial exchanges and communication networks. According to the *National Response Plan*, critical infrastructure encompasses "systems and assets, whether physical or virtual, so vital to the United States that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters."[5]

The final group of proposals involving the private sector regards recommendations that include sharing information. For example, the report notes that "Congress should . . . direct and resource the federal government to establish a formal process to solicit and compile private-sector input to inform national intelligence priorities, collection requirements, and more focused U.S. Intelligence support to private-sector cyber security operations."[6] In addition, the report recommends the creation of a Joint Collaborative Environment in which information and relevant data can be shared across the federal government and between the public and private sectors.[7] Information sharing, especially within

the cyber environment, will be the most difficult challenge facing the United States as our near-peer competitors and adversaries continue improving their cyber capabilities. For example, our near-peer competitors and adversaries are using all the elements of mis-, dis-, and malinformation (MDM) to "elicit a strong emotional response from the consumer and bypass logical reasoning to incite action, whether the action is simply spreading the content further on social media or taking action in the real world, including acts or threats of violence."[8] Cybercriminals and nonstate actors are also migrating toward the "crime-as-a-service" model. Russia, one of our most important competitors, is heavily involved in the MDM business. A growing sector of Russia's MDM economy is the "Manipulation Service Providers," both in the national and international arenas.[9]

Regarding cooperation between the federal and state governments, the report recommends two proposals. The first proposal involves election security and assists in funding for maintaining election infrastructure. This assistance includes grants to the states for auditable voting systems, replacing outdated voting equipment, and providing for sufficient provisional ballots and post-election audits. Under this proposal, states must fund 30 percent of the cost.[10] The second proposal involves state and federal cooperation regarding the promotion of cyber insurance. Mainly, the report involves cybersecurity insurance products and provides for collaboration between federal officials and state insurance regulators.[11]

Based on these proposals, there is ample room for cooperation between the federal government, the private sector, and states. These recommendations make it clear that cooperation is required to protect the United States from cyber threats adequately. This requirement leads to the possibility of cooperation. However, while possible, it is still left to be explored what the result of such cooperation and is constitutional.

The Results of Governmental Cooperation with the Private Sector and States

Over time, there have been several instances in which the government has cooperated with the private sector or states. But unfortunately, the results of such cooperation typically lead to national government domination. Due to this domination, more trust may be needed between the private sector and the federal government or the states and the federal government. Such cooperation may therefore prove difficult as there may be a reluctance on the part of the private sector and states to participate willingly in such a cooperative approach as envisioned in the report. Regardless, this section will provide several examples of cooperation that led to such domination, which may account for the reluctance on behalf of the private sector and states.

Homeland Security and Defense after the 9/11 Attacks

The cyber domain is the new battlefield of the twenty-first century. The cyber

domain is no longer the domain of wannabe cyber hackers or script kids, who are unskilled computer users that use programs or scripts developed by others to carry out their nefarious activities online. Today, the domain is dominated by nation-states and their proxies, transnational criminal organizations (TCOs), and cyber criminals using sophisticated and malicious tactics to undermine our nation's critical infrastructure, steal intellectual property and innovation, engage in espionage, and threaten our democratic institutions. TCOs directly threaten the United States and its allies "through human trafficking, the production and trafficking of lethal illicit drugs, cybercrime, and financial crimes and money laundering schemes eroding the integrity of the international financial system."[12] According to the *Federal Bureau of Investigation Internet Crime Report* produced by the Internet Crime Complaint Center (IC3), in 2021, IC3 received 847,376 cyber complaints and reported a net loss of U.S. $6.9 billion. The top five crimes in 2021 were: extortion (39,360 cases), identity theft (51,629), a personal data breach (51,829 cases), nonpayment/nondelivery (82,478 cases), and phishing/vishing/smishing/pharming (323,972 cases).[13]

Another important group operating within the cyber domain carrying out its nefarious activities are digital influence mercenaries. Digital influence mercenaries are also called virtual mercenaries. They are highly skilled computer users available on the gray market to the highest bidder, be it a nation-state, nonstate actor, terrorist organization, or private individual. The digital influence of mercenaries' rise is also due to the simple economic forces of supply and demand.[14] Digital influence mercenaries claim "their services only focus on criminals and terrorists"; however, Meta's monthslong investigation concluded that targeting is indiscriminate and includes journalists, dissidents, critics of authoritarian governments, families of opposition, and human rights activists.[15] Digital information mercenaries are also responsible for spreading misinformation, disinformation, and malinformation.

The 9/11 attacks against the homeland showed how ill-prepared the United States was to protect the homeland. As a result, discussions ensued about what was needed after the attacks on the World Trade Center towers and the Pentagon. According to James Jay Carafano at the Heritage Foundation, post-9/11 "there was an effort to create a permanent and persistent federal structure to deal with the inside-outside enemy."[16] The solution to the 9/11 attacks on the homeland was the creation of the Department of Homeland Security (DHS). The DHS was created when President George W. Bush signed the Homeland Security Act of 2002 on 25 November 2022.[17] Former Pennsylvania governor Tom Ridge (R-PA) was appointed the first director of the Office of Homeland Security.

The DHS mission is to prevent attacks and protect Americans—on the land, in the sea, and in the air. Furthermore, DHS combines all or part of 22 different federal departments and agencies into a unified, more effective, integrated department, creating a strengthened homeland security enterprise and a more secure America that is better prepared to confront the range of threats the

United States faces. The DHS has three core values that all federal departments and agencies share under its overarching organizational structure. The first core value is integrity or "service before self." According to the DHS's website, members of the DHS family "will faithfully execute the duties and responsibilities entrusted to us, and we will maintain the highest ethical and professional standards." The second core value is vigilance or "guarding America." DHS professionals state, "we will constantly be on guard against threats, hazards, or dangers that threaten our values and our way of life." Finally, the third core value is respect or "honoring our Partners." According to the DHS's website, "We will value highly the relationships we build with our customers, partners, and stakeholders. We will honor concepts such as liberty and democracy, for which America stands." In summary, DHS's mission is:

> With honor and integrity,
> We will safeguard the American people,
> Our homeland, and our values.[18]

In addition to the three core values mentioned above, DHS is guided by five principles that shape its missions. The five guiding principles are to champion "Relentless Resilience" for all threats and hazards; reduce the nation's risk to homeland security dangers; promote citizen engagement and strengthen and expand trusted partnerships; uphold the privacy, transparency, civil rights, and civil liberties; and ensure mission-driven management and integration.[19]

According to the U.S. Department of Homeland Security's strategic plan for fiscal years 2020–24, DHS has six overarching homeland security missions that make up its strategic plan.[20] The six missions are counterterrorism and homeland security threats; secure U.S. borders and approaches; secure cyberspace and critical infrastructure; preserve and uphold the nation's prosperity and economic security; strengthen preparedness and resilience; and champion the DHS workforce and strengthen the department.[21] This article will primarily discuss how DHS's mission to secure cyberspace and critical infrastructure disproportionally favors the federal government. This potential federalization of the cyber domain may minimize the roles private entities can play in protecting the cyber domain and could hinder a partnership between the federal government and private entities vital to detecting, deterring, neutralizing, and protecting the cyber domain.

Recognizing that the cyber domain is a force multiplier within the operational environment in which nations compete for supremacy, thus rendering the threat landscape more challenging than ever, DHS has taken several steps to mitigate the potential cyber harm that could paralyze the nation's critical infrastructure. For example, in 2004, DHS created the National Cyber Security Division (NCSD). One of the primary functions of the NCSD is to "partner with government, industry, academia as well as the international community to make cybersecurity a national and shared priority."[22] Another vital DHS agency in the fight to protect our critical infrastructure and nefarious activities online

by criminal elements is the Cyber Crimes Center, composed of the following units: Cyber Crimes Unit, the Child Exploitation Investigations Unit, and the Computer Forensics Unit.

Furthermore, DHS's cybersecurity and critical infrastructure security responsibilities focus on four goals: securing federal civilian networks; strengthening the security and resilience of critical infrastructure; assessing and counter evolving cybersecurity risks; and combating cybercrime. According to DHS, "Serving as the designated federal lead for cybersecurity across the U.S. Government, DHS promotes the adoption of common policies and best practices that are risk-based and responsive to the ever-changing cyber threat environment."[23] Obviously, those so-called common policies and best practices sometimes conflict with the hardware and software used by federal agencies, primarily if they are owned and controlled by private investors. In fact, according to the *Department of Homeland Security's (DHS) Critical Infrastructure Protection Cost-Benefit Report*, the private sector owns most of the nation's critical infrastructure and key resources—roughly 85 percent.[24] However, the government has historically funded the construction and maintenance of specific infrastructure sectors such as transportation and water.[25]

A concern arises as more of the critical infrastructure used by the federal government is in the private sector's hands and controlled by private investors; the federal government may relinquish its traditional responsibility as caretaker of the nation's critical infrastructure and rely on the private sector to assume its traditional responsibilities instead of the federal government. As is pointed out in the *Solarium Report*, "businesses are often reluctant to let governments onto private, commercial networks without a clear understanding of their shared interests and responsibilities. Afraid of creating moral hazard, the federal government invests little in protecting the cybersecurity of commercial infrastructure or key systems controlled by states and local municipalities."[26] The distrust between the private sector and federal government and the lack of accountability on who is responsible for setting the nation's cybersecurity priorities produces dangerous security gaps. This gap occurs when "public- and private-sector responses are left uncoordinated, and the nation's critical infrastructure is left unprotected and vulnerable to adversaries who can, and will, exploit this opportunity."[27]

The Cybersecurity and Infrastructure Security Agency

The Cybersecurity and Infrastructure Security Agency Act of 2018 established the Cybersecurity and Infrastructure Security Agency (CISA).[28] Its director, Jen Easterly, leads CISA. CISA's Cybersecurity Division is led by Executive Assistant Director for Cybersecurity Eric Goldstein. CISA leads the nation's strategic and unified work to strengthen the cyber ecosystem's security, resilience, and workforce to protect critical services and the American way of life from cybercriminals, cyberterrorism, and adversaries. According to CISA's website, its primary mission is "lead[ing] efforts to protect the federal .gov domain of

civilian government networks and to collaborate with the private sector—the .com domain—to increase the security of critical networks." Protection of the government's .gov domain is accomplished through the following functions:

- Capability delivery
- Threat hunting
- Operational collaboration
- Vulnerability management
- Capacity building
- Strategy, resources, and performance
- Cyber defense education and training

It is often said that the only computer that has not been attacked is a computer that is not turned on. Recognizing the American way of life and its dependency on technology for almost everything, including but not limited to connecting with friends and relatives, banking, traveling, shopping, education, work, and romance, CISA "serves as both America's cyber defense agency and as the national coordinator for critical infrastructure security and resilience."[29] One such program where CISA takes a proactive approach to address our nation's infrastructure security and resilience is the ShieldsUp campaign introduced in late 2021.

ShieldsUp was launched in the aftermath of the Russian invasion of Ukraine. CISA's ShieldsUp campaign encourages "organizations of all sizes to take immediate steps to improve their cybersecurity and protect their critical assets" in the face of "potential spillover effects to the U.S. homeland" as the Russian-Ukraine conflict continues without a diplomatic solution or ceasefire.[30] CISA's position is that the increasing technological interconnectedness of the world and the American people's reliance on technology for almost every aspect of their daily life requires a "continuous, whole-of-government approach that spans all stakeholders."[31] In addition, specific sectors of the economy composing the National Critical Functions are essential to CISA's mission, which is "to lead the national effort to understand, manage, and reduce risk to our cyber and physical infrastructure."[32] According to the CISA's website, there are 16 critical sectors comprising the U.S. critical infrastructure. Those 16 essential sectors of infrastructure are crucial since their "assets, systems, and networks, whether physical or virtual, are considered so vital to the United States that their incapacitation or destruction would have a debilitating effect on security, national economic security, national public health or safety, or any combination thereof."[33]

Sectors of the U.S. economy considered part of the National Critical Functions are required by CISA to "put in place measures to detect, delay, and respond to physical and cyberattacks such as establishing security officials; creating barriers and access control measures; implementing intrusion detection capabilities; and developing incident reporting, response, and investigation programs for both physical and cyberattacks, among other measures."[34] That

is a tall order to accomplish given that the "majority of the critical infrastructure, hardware, and software that powers the information age resides in the private sector."[35] Furthermore, because private-sector companies are not "part of the defense industrial base, they have no legal obligation to report information technology system anomalies, increased traffic, or Information Technology (IT) security breaches."[36] Finally, businesses are often reluctant to "let government onto private, commercial networks without a clear understanding of their shared interests and responsibilities."[37]

To break down this cycle of distrust between the federal government and private sectors, the *CISA Strategic Plan 2023–2025* Goal 3 (Operational Collaboration) wishes to establish a culture of "trusted, sustained, and effective partnerships between the government and the private sector" as a foundation "to protect the nation's critical infrastructure."[38] CISA also states that under *Strategic Plan 2023–2025*, the organization will "approach every partnership with humility, transparency, gratitude, and a firm resolution to add value wherever possible."[39] This is an ambitious goal for an organization that has been institutionally limited in its ability to carry out its mission fully. As the report clearly states in its findings:

> CISA has been institutionally limited in its ability to fully carry out this mission, hindered by inadequate facilities, insufficient resources, lack of buy-in from other federal departments and agencies, ambiguity from Congress on its role and position about other agencies, and inconsistent support to and integration with private-sector.[40]

Commerce Clause

Finally, the best example of cooperation that led to national domination occurs between the federal and state governments in the realm of the Commerce Clause. Throughout the nineteenth and twentieth centuries, the federal government battled with conditions regarding interpreting this clause. The action was based on the type of federalism to which the Supreme Court should adhere to. First was dual federalism. The court held that there were two separate spheres in which certain rights fell under state authority, and the remaining rights were under the national government's authority. During the periods when the court adhered to dual federalism, it generally ruled in favor of the state's rights. This contrasts with cooperative federalism, where states and the federal government are supposed to cooperate. However, such cooperation inevitably led to national domination, with the Supreme Court consistently ruling in favor of the national government during this period. A brief historical analysis will help illustrate how the court enabled the strengthening and growth of the federal government at the expense of state sovereignty.[41]

Every Commerce Clause analysis should begin with Gibbons v. Ogden (1824), involving a dispute between an individual who had a state-granted monopoly on its waters, and Gibbons, who claimed the right to travel on interstate waters pursuant to a federal license.[42] At issue in the dispute was whether

the state had concurrent powers to regulate interstate navigation. The Supreme Court ultimately held that the federal government had the right to regulate interstate commerce, which included navigation. While the holding of the case is important, how the court arrived at its opinion is noteworthy. Justice John Marshall announced that states could regulate intrastate travel, but interstate travel was the federal government's responsibility.[43] The court's ruling was thus adhering to a form of dual federalism, but one that favored the federal government.

The next era also adhered to dual federalism but favored the state governments. Dual federalism eras are essential as they permitted the state governments to retain sovereignty. After all, at the root of dual federalism eras is a recognition of equal powers between the federal and state governments. This can be seen in the second era of Commerce Clause jurisprudence following Gibbons, lasting from 1836 to 1937, and showed state government dominance with the court ruling consistently in favor of the states.

For example, in United States v. E. C. Knight Company (1896), the American Sugar Refining Company acquired E. C. Knight Company, among others, giving American Sugar Company nearly 98 percent control of the country's sugar production.[44] The United States attempted to nullify the acquisition because the sale amounted to a monopoly, thus constituting a trust. Therefore, the Sherman Anti-Trust Act (1890), passed pursuant to the Commerce Clause, was applied to prevent the sale. The Supreme Court ultimately held that while the Sherman Anti-Trust Act was valid, it did not apply in this case since E. C. Knight was only involved in manufacturing and production, which did not constitute commerce during this era.[45] This case displays dual federalism, with the court attempting to create a test in which manufacturing and production fell under the auspices of state regulation. In contrast, distribution across state lines was a matter of federal regulation.

This test and application of dual federalism were also seen in Hammer v. Dagenhart (1916), which involved child labor.[46] More specifically, Congress had passed an act prohibiting such labor. The court held that Congress could not regulate child labor since such work was only involved in producing and manufacturing materials, which did not constitute commerce. Again, the Supreme Court's ruling was preserving the sovereignty of states at the expense of federal power.

The court's interpretation of the Commerce Clause and its use of dual federalism changed in 1937. This change was primarily due to the credible threat of court expansion from President Franklin D. Roosevelt in response to the court's continual rulings upholding state's rights and knocking down pieces of Roosevelt's New Deal legislation. Roosevelt's court-packing plan did not come to fruition, as one of the justices on the court began to switch his vote, ruling in favor of the New Deal legislation. This would usher in a period of cooperative federalism, in which the federal government was to "cooperate" with the state governments.[47] However, what resulted from the implementation of coopera-

tive federalism was national domination. This can mainly be seen in cases for the third era of the Commerce Clause, which lasted from 1937 to 1995. During this era, the Supreme Court would consistently rule in favor of the federal government.

Several cases are illustrative of federal domination during this era. The first is NLRB v. Jones and Laughlin Steel Corp. (1937), where the Supreme Court examined the constitutionality of the National Labor Relations Act passed pursuant to the Commerce Clause.[48] Recall that production and manufacturing did not constitute commerce in the prior era and hence could not be regulated by Congress. However, the court began examining the aggregate effect on commerce in this era. It therefore looked to the total impact of what was to be regulated and no longer just looked at production and manufacturing. Ultimately, the Supreme Court upheld the act and ruled in favor of the federal government.

This case was followed by United States v. Darby (1941), involving the constitutionality of the Fair Labor Standards Act of 1938 (FLSA).[49] Again, the issue involved goods produced in one state but shipped across state lines. While this issue appeared to have been resolved in the prior era, the court reexamined it and explicitly overruled Hammer v. Dagenhart. The court finally discarded the production/distribution rule it had utilized in the previous period and held that the FLSA was constitutional.

Two final cases show federal domination during the cooperative era. However, these cases are unique in that the link between what was being regulated and the Commerce Clause was arguably tenuous. For example, the first of these cases was Heart of Atlanta Motel v. United States (1964), involving the constitutionality of the 1964 Civil Rights Act.[50] The particular provision at issue, in this case, prohibited racial discrimination in areas affecting public accommodation. In Heart of Atlanta Motel, the motel essentially argued that they were not engaged in interstate commerce, even though they placed advertisements in national magazines and billboards and received most of their guests from out of state. Nevertheless, the court held that the act was valid and the Commerce Clause could be used to regulate racial discrimination. This was an expansive interpretation of the Commerce Clause.

However, the most expansive interpretation that favored the federal government during this era can be seen in the case of Wickard v. Filburn (1942).[51] In Wickard, the 1933 Agricultural Adjustment Act limited the wheat farmers could grow. Filburn grew several acres of wheat for consumption on his farm, more than the amount allowed under the act. The court held that the act was valid even though it seemingly regulated intrastate consumption. To justify its opinion, the court avoided analyzing the case by looking at Filburn's wheat consumption. Instead, the court considered the potential effects of all the individuals growing home wheat and how that could affect the overall market. Thus, the court considered the aggregate impact from all individuals violating the act and held that this constituted commerce.

The cases during this era make it clear that cooperative federalism leads

to national domination. In every case considered during this time frame, the Supreme Court upheld Congress's right to enact the legislation under the Commerce Clause. However, the results during the final Commerce Clause era (1995–present) are mixed at best. Some cases favor federal rights, while others favor states' rights. These mixed results are likely due to a court that grew more conservative and hence more in favor of states' rights. This revival of states' rights leads one to wonder whether the court reverted to dual federalism.

A few cases will illustrate this point. The first is perhaps the most important, as it was the first case where the conservative court overturned a congressional statute based on the Commerce Clause. The case, Lopez v. United States (1995), involved the Gun-Free School Zones Act of 1990, in which a high school student brought a gun to school and was charged with violating the act.[52] The student, Lopez, argued that Congress had exceeded its authority under the Commerce Clause in passing the act. The court agreed, holding that regulating guns on campuses did not amount to commerce.

A similar outcome was reached in United States v. Morrison (2000) involving the 1994 Violence Against Women Act.[53] Congress had again passed the act under the Commerce Clause, as it was argued that if you aggregate all the instances of domestic violence, many women would be unable to work during the year, which would impact the overall economy. The conservative court held that this was too tenuous of a connection and that crime cannot be aggregated to make commerce.

However, despite these two previous cases, the court did rule in favor of the government once during this era. This is mainly seen in Raich v. Gonzales (2005), in which individuals grew marijuana for their consumption.[54] In reaching their decision, the court cited Wickard and noted that home consumption of marijuana affected the overall economy of the primarily illegal product.

The mixed results obtained from the prior three cases are significant as they show the court needs help interpreting the Commerce Clause and justifying its decision. In other words, the court is trying to figure out which form of federalism it adopts and how to balance the delicate relationship between federal and state governments. For present purposes, what is important is that during the cooperative federalism phase, when the governments were supposed to cooperate, the federal government dominated the field.[55] This is like the expected outcome of the cooperation as envisioned in the *Solarium Report*. Based on the precedents established in the historical overview, one should be cautious in advocating cybersecurity changes that may alter the government's balance with the private sector. While there may be certain advantages to the federal government taking the lead in cybersecurity, the potential loser may be the private sector and the states.

Overall, as these precedents show, the relationship between the federal and state governments is tenuous. Nevertheless, the relationship between these governments is always in play, as the federal government consistently attempts to dominate the states. It is only during cooperative times that this becomes pos-

sible. Therefore, state governments and the private sector should be wary of any proposals of cooperation.

Is Cooperation Constitutional?

While the three previous sections outlined precisely how cooperation could lead to the potential nationalization of cybersecurity, some may wonder whether the proposals presented in the *Solarium Report* are constitutional. After all, constitutional questions can be expected when the balance of power between the governments and the private sector shifts.

The primary constitutional concern regarding the report involves the proposed cooperation between states and the federal government in election security. More particularly, the report recommends providing grants to states that require the states to match 30 percent of funds to protect federal elections from cyber threats. Brian T. Yeh examined the federal government's limitations to imposing conditions on grant funds.[56] These limitations are primarily found in South Dakota v. Dole (1987), which held that according to the Spending Clause, legislation must be in pursuit of the "general welfare."[57] In addition, Yeh noted that the Dole Court held that

> any conditions attached to the receipt of federal funds must: (1) be unambiguously established so that recipients can knowingly accept or reject them; (2) be germane to the federal interest in the particular national projects or programs to which the money is directed; (3) not violate other provisions of the Constitution such as the First Amendment or the Due Process or Takings Clauses of the Fifth Amendment; and (4) not cross the line from enticement to impermissible coercion, such that states have no real choice but to accept the funding and enact or administer a federal regulatory program.[58]

The fourth provision that could apply to the cybersecurity context involves coercion. The court has addressed coercion with the Taxing and Spending Clause in the Dole case and NFIB v. Sebelius (2012).[59]

In Dole, the federal government wanted to raise the drinking age to 21. However, states are typically in charge of such age requirements. Therefore, the federal government attached conditions to the receipt of federal highway money, prohibiting some of the funding from going to any state that failed to comply. The court ultimately held in favor of the federal government since it did not take away all funding but only threatened to take a small percentage of it.[60]

Dole can be contrasted with Sebelius, which involved the constitutionality of the 2010 Affordable Care Act. While there were multiple issues in the case, the most applicable one involved Medicaid expansion. More particularly, the federal government threatened the states with the complete loss of all Medicaid if they did not comply with the new proposed health care law. Unlike in Dole, the court ruled that the federal government had gone too far this time and that their actions amounted to coercion and were thus unconstitutional.[61]

Regarding the *Solarium Report*, it is recommended that the federal government help secure the state's election apparatus, with the states responsible for 30 percent of the costs. The constitutional issue is whether such a proposal is like Dole or Sebelius. In other words, is depriving states of a grant because of a 30 percent funding requirement constitute coercion like in Sebelius, or is it more consistent with Dole since it is not a complete threat of deprivation of federal funds? After all, only a small percentage was withheld in Dole, which was deemed constitutional, while a complete deprivation in Sebelius was deemed unconstitutional. This case is likely like Dole, but it should be noted that the constitutionality of such a provision is questionable.[62]

This section is essential in the report as it highlights a substantial constitutional question. While the constitutionality is questionable, nationalistic recommendations should not hinge on whether cooperation should be constitutional. While the report strives to maintain election security, there are other possible means to do so than forcing states to match federal funds that would be less constitutionally suspect. This includes matching funds at a lesser rate or true cooperation between the federal government and the states.

Implications

The fact that 85 percent of the critical infrastructure the government relies on is in private hands and controlled by private investors is a concern for a highly interconnected and wired nation. The United States' critical infrastructure "provides national critical functions that are so vital to the United States that their disruption, corruption, or dysfunction would have a debilitating effect on the Nation's security, economy, and public health and safety."[63] Critical infrastructure today faces increasingly new risks and challenges moving forward. Our modern way of life in an interconnected and highly wired world depends on confidentiality, integrity, and availability, also called the CIA Triad of data. Furthermore, the United States "is facing adversary nation-states, extremists, and criminals leveraging emerging technologies to an unprecedented degree. Authoritarian states seek to control every aspect of life in their societies and export this style of government, in which surveillance trumps liberty, to the rest of the world."[64] Finally, TOCs and cyber criminals are migrating toward the "crime-as-a-service" model in which threat groups purchase and exchange malicious code on the dark web.[65]

Recommendations

1. The U.S. government and the private sector must create a new social contract of shared responsibility to secure the nation's cyberspace, recognizing each other as partners to diminish the distrust between the private sector and the government.
2. Information sharing between the federal government and the private sector rather than operating in silos and keeping secrets from each other to diminish the distrust.

3. The matching amount states are required to give for election security should be reduced to ensure constitutionality; and
4. Increased workforce recruitment and talent acquisition and management

Workforce recruitment and talent acquisition will be a challenge to the U.S. government. The U.S. government needs to be able to protect its critical infrastructure with a crucial civilian workforce. Dr. Raj Iyer, Army chief information officer (CIO) at the U.S. Army Europe and Africa 2022 Cybersecurity Summit, held on 29 July 2022, pointed out that finding the right people is one of his biggest challenges as an Army CIO. He emphasized "the importance of filling the cyber talent gap and that the Army plans to address this perennial challenge by rolling out the Department of Defense Cyber Excepted Service, a new talent model for the civilian cyber workforce, this year. The service will take advantage of every available tool to recruit and retain the cyber workforce."[66] Unfortunately, one of the tools not available to the Army is the high compensation package provided by the private sector to cybersecurity professionals. For example, "positions includ[e] cybersecurity analyst, information security analyst, and penetration tester, and annual median salaries ranging from $75,000 to more than $100,000."[67]

Conclusion

Humans created the cyber ecosystem on which the nation relies; therefore, it is susceptible to vulnerabilities. Furthermore, this system is more than simply the technology that comprises it. It also comprises people, processes, and organizations that plug into the technology and the data they combine to produce complex products.[68] Overall, as reviewed throughout this article, the report has made several key recommendations that continue the trend of nationalizing cybersecurity. This trend was placed in a historical context through the creation of the Department of Homeland Security and CISA. In addition, the historical overview of the Commerce Clause explains how the federal government, through the lens of cooperative federalism, became dominant over states' rights. The report assumes the federal government's dominance is explicit, while it claims to seek cooperation and, in many instances, the recommendations arguably trample on state and individual rights. In other words, consistent with the Commerce Clause and the nationalization of our government, complete adoption of the report could lead toward the nationalization of cybersecurity.

In conclusion, one may wonder what the results of such nationalization are and whether it is positive or negative for the United States. One positive aspect of nationalization would be uniformity in cybersecurity. There is no need for the potential of 50 state responses to a particular cyber threat. In addition, a swifter response from the federal government may issue if nationalization occurred in the field. However, as noted throughout this article, nationalization has potentially adverse consequences. One example involves trampling states'

rights, especially in the election field. While the report has noble goals, forcing state participation may be unconstitutional. Furthermore, the relationship between the federal government and the private sector may need further analysis. Nationalization may also lead to forced participation in this relationship, which may lead to new constitutional challenges in the future.

Endnotes

1. *United States of America Cyberspace Solarium Commission Report* (Washington, DC: Cyberspace Solarium Commission, 2020), hereafter *Solarium Report*.
2. *Solarium Report*, 23.
3. *Solarium Report*, 58.
4. *Solarium Report*, 59.
5. *National Response Plan* (Washington, DC: U.S. Department of Homeland Security, 2004), 64.
6. *Solarium Report*, 100.
7. *Solarium Report*, 101.
8. "Election Infrastructure Insider Threat Mitigation Guide," Cybersecurity & Infrastructure Security Agency, accessed 17 April 2023.
9. *Solarium Report*, 68.
10. *Solarium Report*, 67–68.
11. *Solarium Report*, 80–82.
12. *Annual Threat Assessment of the U.S. Intelligence Community* (Washington, DC: Office of the Director of National Intelligence, 2022), 23.
13. *Internet Crime Report* (Washington, DC: Federal Bureau of Investigation, 2021). *Phishing* is a scam in which the perpetrator sends out legitimate-looking email to phish for personal and financial information from the recipient; *vishing* is a social engineering activity over the telephone system, most often using features facilitated by Voice over Internet Protocol (VoIP), to gain unauthorized access to sensitive data; *smishing* is the fraudulent practice of sending text messages purporting to be from reputable companies to induce individuals to reveal personal information, such as passwords or credit card numbers; and *pharming* is the fraudulent practice of directing internet users to a bogus website that mimics the appearance of a legitimate one to obtain personal information such as passwords, account numbers, etc. For more definitions of current cybersecurity terms, see Adam Gordon, ed., *Official ISC² Guide to the CISSP® CBK®*, 4th ed. (Boca Raton, FL: CRC Press, 2015).
14. James J. F. Forest, *Digital Influence Mercenaries: Profits and Power through Information Warfare* (Annapolis, MD: Naval Institute Press, 2022).
15. "Meta Bans 'Cyber-Mercenaries' that Targeted 50,000 People," Al Jazeera, 17 December 2021.
16. James Carafana, "Homeland Defense and Homeland Security: Distinctions and Difference," in *Introduction to Homeland Defense and Defense Support of Civil Authorities: The U.S. Military's Role to Support and Defend*, ed. Bert B. Tussing and Robert McCreight (Boca Raton, FL: CRC Press, 2015).
17. Homeland Security Act of 2002, Pub. L. No. 107-296 (2002).
18. "Creation of the Department of Homeland Security," Department of Homeland Security, 3 June 2022.
19. "Creation of the Department of Homeland Security."
20. *The DHS Strategic Plan: Fiscal Years 2020–2024* (Washington, DC: Department of Homeland Security, n.d.).
21. "Creation of the Department of Homeland Security."
22. William M. Oliver, Nancy E. Marion, and Joshua B. Hill, *Introduction to Homeland Security: Policy, Organization and Administration*, 2d ed. (Burlington, MA: Jones & Barlett Learning, 2021), 75.

23. "Secure Cyberspace and Critical Infrastructure," Department of Homeland Security, accessed 12 May 2023.
24. *The Department of Homeland Security's (DHS) Critical Infrastructure Protection Cost-Benefit Report* (Washington, DC: Government Accountability Office, 2009).
25. *Critical Infrastructure: Long-term Trends and Drivers and Their Implications for Emergency Management* (Washington, DC: Federal Emergency Management Agency, 2011).
26. *Solarium Report*, 16.
27. *Solarium Report*, 17.
28. The Cybersecurity and Infrastructure Security Agency Act of 2018, Pub. L. No. 115-278 (2018).
29. *CISA Strategic Plan, 2023–2025* (Washington, DC: Cybersecurity and Infrastructure Security Agency, 2022), 3.
30. *CISA Strategic Plan 2023–2025*, 5.
31. *CISA Strategic Plan 2023–2025*, 5.
32. *CISA Strategic Plan 2023–2025*, 6.
33. "Critical Infrastructure Sectors," Cybersecurity and Infrastructure Security Agency, accessed 18 April 2023. The 16 critical infrastructure sectors are: chemical sector; communication sector; dams sector; emergency services sector; financial services sector; government facilities sector; information technology sector; transportation systems sector; commercial facilities sector; critical manufacturing sector; defense industrial base sector; energy sector; food and agriculture sector; health care and public health sector; nuclear reactors, materials, and waster sector; and water and wastewater systems sector.
34. *CISA Strategic Plan 2023–2025*, 17.
35. *Solarium Report*, 16.
36. Bert B. Tussing et al., *Contested Deployment: A US Army War College Center for Strategic Leadership Integrated Research Project* (Carlisle, PA: Army War College Press, 2022), 114.
37. *Solarium Report*, 16.
38. *CISA Strategic Plan 2023–2025*.
39. *CISA Strategic Plan 2023–2025*, 23.
40. *Solarium Report*, 105.
41. For more discussion regarding the different types of federalism and resulting power if the federal government, see Theodore J. Lowi et al., *American Government: Power and Purpose*, 16th ed. (New York: W. W. Norton, 2021), 81–93.
42. Gibbons v. Ogden, 22 U.S. 1 (1824).
43. David M. O'Brien, *Constitutional Law and Politics*, vol. 1, *Struggles for Power and Governmental Accountability*, 10th ed. (New York: W. W. Norton, 2017), 519, 533.
44. United States v. E. C. Knight Co., 156 U.S. 1 (1895).
45. O'Brien, *Constitutional Law and Politics*, vol. 1, 539, 543–45.
46. Hammer v. Dagenhart, 247 U.S. 251 (1918).
47. For more discussion regarding the change in federalism, see Lowi et al., *American*, 81–93. For more background and discussion regarding 1937 and the threats toward the court, see O'Brien, *Constitutional Law and Politics*, vol. 1, 553–54.
48. NLRB v. Jones & Laughlin Steel Corp., 301 U.S. 1 (1937).
49. United States v. Darby, 312 U.S. 100 (1941).
50. Heart of Atlanta Motel, Inc. v. United States, 379 U.S. 241 (1964).
51. Wickard v. Filburn, 317 U.S. 111 (1942).
52. United States v. Lopez, 514 U.S. 549 (1995).
53. United States v. Morrison, 529 U.S. 598 (2000).
54. Gonzales v. Raich, 545 U.S. 1 (2005).
55. For more discussion regarding the Commerce Clause and different types of federalism, see Lowi et al., *American Government*, 81–93. For more background and discussion of the cases previously cited, see David M. O'Brien and Gordon Silverstein *Constitutional Law and Politics*, vol. 1, *Struggles for Power and Governmental Accountability*, 11th ed. (New York: W. W. Norton, 2020), 549–645.

56. Brian T. Yeh, *The Federal Government's Authority to Impose Conditions on Grant Funds* (Washington, DC: Congressional Research Service, 2017).
57. South Dakota v. Dole, 483 U.S. 203 (1987).
58. Yeh, *The Federal Government's Authority to Impose Conditions on Grant Funds*.
59. National Federation of Independent Business v. Sebelius, 567 U.S. 519 (2012).
60. O'Brien, *Constitutional Law and Politics*, vol. 1, 636–38.
61. O'Brien, *Constitutional Law and Politics*, vol. 1, 639–48.
62. One may attempt to distinguish Dole and Sebelius from our current situation, since in those cases the states had a preexisting grant that was threatened to be reduced or extinguished. In this case, states merely have an option to participate in the election security program and arguably will not necessarily suffer a reduction or extinguishment of funds. Yet, the states are still suffering from a potential loss of participation in a critical program based on not matching federal funds. Thus, not participating in a grant program is still viewed as a loss or extinguishment of said program.
63. "National Critical Functions," Cybersecurity and Infrastructure Security Agency, accessed 12 May 2023.
64. *Solarium Report*, 19.
65. Keman Huang, Michael Siegel, and Stuart Madnick, "Cybercrime-as-a-Service: Identifying Control Points to Disrupt" (working paper CISL#2017-17, MIT Sloan School of Management, November 2017), 13.
66. Sun Vega, "Army CIO speaks at Army Europe and Africa 2022 Cybersecurity Summit," Army.mil, 10 August 2022.
67. Steve Morgan, "Cyber Jobs Report: 3.5 Million Openings by 2025," *Cybercrimes Magazine*, 14 April 2023.
68. *Solarium Report*, 71.

Including Africa Threat Analysis in *Force Design 2030*

Glen Segell, PhD

Abstract: This article examines the threat analysis across Africa that should be included in *Force Design 2030* for the United States Marine Corps to be deployed landward to Africa or seaward of the continent. It is a strategic guidance document examined from a threat analysis of China, Russia, Korea, Iran, and violent extremist organizations. Africa is not mentioned, and this is a notable omission given that high level interventions in the past to Africa have not been overtly successful. Given geostrategic significances and hot spots it is inevitable that the Marines will be deployed there again. This article examines lessons learned from failures in Somalia, Libya, and Lebanon and successes in Syria and Iraq as well as the experiences of others—France in Mali and Burkina Faso and United States Africa Command. Great power competition, violent extremist organizations, and the gray zone phenomena across Africa are examined as are security, intelligence, counterintelligence, and hybrid warfare.

Keywords: *Force Design 2030*, United States Marine Corps, Africa, great power competition, gray zone, violent extremist organizations, security, counterinsurgency, intelligence, counterintelligence

Introduction

The roles and deployment of the United States Marine Corps are dynamic. *Force Design 2030* (FD2030) has been written as a strategic guidance documented for a modernization program that aims to ensure that the U.S. Marine Corps remains relevant to the current and future battlespace, has

Dr. Glen Segell is professor at the University of Cambridge and visiting professor and research fellow in the Department of Political Studies and Governance, University of the Free State, South Africa. He is also a research fellow at the Ezri Center for Iran and Gulf States Research, University of Haifa, Israel. He holds the rank of brigadier general (reserves), where he also consults as an expert for the North Atlantic Treaty Organization (NATO). He worked in active intelligence and offense operations in Iraq, Kuwait, Sudan, and Libya.

Journal of Advanced Military Studies vol. 14, no. 1
Spring 2023
www.usmcu.edu/mcupress
https://doi.org/10.21140/mcuj.20231401008

adapted to do so, and so can outmaneuver any potential adversaries. FD2030's purpose is to transform the Marine Corps' existing force design to contend with the character of war that will include precision strike regimes, gray zone strategies, and an emphasis on maritime campaigns.[1]

The threat analysis of the version of FD2030 that is available on its website examines China, Russia, Korea, Iran, and violent extremist organizations. The problem statement that this article examines is that a notable omission in FD2030 is Africa. It is not included in the threat analysis nor mentioned anywhere in the strategic guidance document. The complete spectrum of the character of future war in different regions is unique and has therefore not been fully examined. The Marine Corps has deployed to Africa and will deploy again as examined in this article. By not including a threat analysis that includes Africa, the Marines will be vulnerable to failures.

In tackling the problem statement, this article examines issues and topics that should be included in FD2030 as strategic guidance to enable a better operationalization of the Marine Corps in Africa. To this end, this article recommends that FD2030 should include lessons learned from the past, both failures and successes, lessons from allies, and strategic guidance based on threat analysis on great power competition, violent extremist organizations, and the gray zone in Africa. It should also include means such as hybrid warfare, partnerships, security, intelligence, and counterintelligence.

The methodology of this article is to examine these issues and topics in different sections in a step-by-step process enouncing the concepts and conceptualizations, giving examples of these from both primary and secondary sources and directly quoting the advice of others. This will lead to the conclusions justifying why it is essential to include Africa as a threat analysis in FD2030.

The Africa Threat Environment

At the fore of the threat analysis for the United States in Africa is great power competition by more than those noted in the FD2030 threat analysis (China, Russia, Korea, and Iran). To this should be added India and Turkey. This great power competition is manifest between each other and with the U.S. and European countries, predominately Britain, France, Germany, Italy, and Belgium. The purpose of this great power competition is to gain as many African partners as possible to attain local and regional strategic influence. The tangible benefits of such influence are gains in economic, political, informational, and military interests where the African leaders are willing partners for national or personal interests.[2]

The U.S. concentration on Africa could be said to be ongoing since the Cold War struggle for influence in the postcolonial proliferation of sovereign states across the continent. However, the characteristics of the great power competition have changed. The nature of the great power competition in Africa differs from other regions in the world, so FD2030 should include an examination of this. It differs for the U.S. and European countries that deploy forces (boots

on the ground), sometimes within the context of the North Atlantic Treaty Organization (NATO) or the European Union, while the others such as China, Saudi Arabia, and Japan do not deploy forces even though they have bases, for example in Djibouti in the Horn of Africa on the Indian Ocean.[3]

Other countries, such as Russia, prefer to use mercenaries while Iran uses a combination of its own Islamic Revolutionary Guard Corps (IRGC) and proxy insurgents such as Hezbollah. In line with American strategic culture and strategy of engagement, the use of mercenaries and nonstate proxies are not an option. It would be unthinkable for American policy makers to use these. American strategic culture favors its own conventional armed forces, sometimes by Special Forces, as observed in past deployments in Somalia and Libya. Consequently, the United States and its allies do not engage other non-African states in direct great power competition military confrontation.[4]

If Africa involved solely great power competition in direct military confrontations, then the force design called for by FD2030 might have value, for most of the threat analysis therein emanates from great power competition threat analysis. However, the growth of violent extremist organizations, both local and ideological extensions of Middle East Islamic fundamentalism, has added to the threat environment and endangers U.S. interests in Africa. These destabilize and threaten locals and have escalated regionally to threaten U.S. geostrategic interests, for example al-Shabaab, Boko Haram, Hezbollah, Islamic State in Iraq and Syria (ISIS), and al-Qaeda. They recruit from the local population and merge with them, equating any Marine Corps deployment tasked with the impossible mission of "looking for a needle in a haystack."[5]

Despite the difficulties to be overcome on the tactical level, violent extremist organizations have been included in FD2030 as tackling them is a cornerstone of U.S. strategic policy. The catalyst for direct U.S. military action against violent extremist organizations leading to deployment to engage them in a preemptive and preventive manner and to assist African state partners to do so emerged with kinetic diplomacy. This was a definition used to describe the policy of President George W. Bush after 11 September 2001. This "war on terrorism" has been applied by the Pentagon, resulting in a shift in U.S. military strategy. That shift has been from containing threats such as applied by Cold War deterrence to deploying forces to engage the threats abroad preemptively, akin to taking the battle to the territory of the enemy.[6] An example of such a shift was that in the Cold War period American forces were required to defend a line such as the Rhine River against an invading Soviet force. Their presence in Germany was to serve as a deterrent to a Soviet offense and they never entered combat with Soviet forces. In the contemporary modern context, the tactics of forward deployment require more logistics capability, for example the liberation of Kuwait in 1991, the deployment to Afghanistan in 2001, and the Iraq War in 2023. In these conflicts, American forces engaged in combat.

The political direction from the White House for such a shift from passive defense to active offense has been echoed and outlined in military doctrinal

documents from the Pentagon. This shift has been discussed within the parameters of the evolution in Marine Corps thinking that led to FD2030. For example, there is a paradigm shift detailed in the open-source web version of FD2030. It informs of a paradigm shift from the 1990s to 2015 quoting from official documents such as *Insurgencies and Countering Insurgencies*, Field Manual 3-24, that is also Marine Corps Warfighting Publication No. 3-33.5.[7] During that period the United States was the sole superpower and enjoyed air, land, and sea supremacy. However, there was a growing need to deploy forces globally in counterinsurgent operations and this influenced the paradigm shift.[8]

Following this sound strategy logically of fighting the adversary away from U.S. soil and given the existence of great power competition and violent extremist organization adversaries, it would have been assumed that FD2030 would have included Africa as a threat analysis. However, it is a notable omission. Moreover, it should be included given the clear identification of great power competition and violent extremist organization threats where it is fair to state that Marine Corps deployment to Africa is an inevitability.[9]

Despite this inevitability, the immediacy of any such deployment is characterized by hesitations that could lead to escalations and even failures. In the past, U.S. policy makers have authorized Marine Corps deployments where and when there has been a clear distinction of when conflict or war exists compared to peace. This is not always the case anymore in Africa. The combination of great power competition and violent extremist organizations is exacerbated by a relatively new phenomena and terminology, mainly since 2016, to describe the changing nature of adversaries and warfare. The multitude and diversity of adversaries in the African threat environment has been summed up by a former commander of United States Special Operations Command Africa, retired Army brigadier general Donald C. Bolduc. He explains that Africa is the best example of a gray zone environment that U.S. forces encounter.[10]

A definition of the gray zone is provided by a National Security Information team:

> A conceptual space that describes a set of activities that occur between peace (or cooperation) and war (or armed conflict) occurring when actors purposefully use single or multiple elements of power to achieve political-security objectives with activities that are typically ambiguous or cloud attribution and exceed the threshold of ordinary competition, yet intentionally fall below the level of large-scale direct military conflict.[11]

Such gray zones are to be found in many countries and regions and have become a military, political, and academic buzzword. The gray zone definition and identification is significant to Africa for the adversary is not solely due to great power competition and violent extremist organizations but could also be a blend of tribal insurgents, jihadists, and criminals operating in failed states. Sometimes these could be the same people as in violent extremist organizations

and great power competition. For example, Hezbollah, which is a violent extremist Islamic Shia organization first seen in Lebanon in the 1980s as a rival to the Shia Amal Movement. Hezbollah is now prevalent in many countries worldwide and is a direct proxy of Iran. It is supported financially and militarily aided by the IRGC that are part of the armed forces of Iran. Hezbollah and the IRGC have been seen operating in unison, for example as first seen in 2010 in Nigeria.[12] Another example of great power competition and the gray zone is seen in Russian mercenaries, exemplified by the Wagner Group, supporting the regime in Mali.[13]

While combating great power competition and violent extremist organizations have a clearer statement in FD2030, the gray zone in Africa does not. This is because the conflict manifests in struggles of resources versus political/religious/ideological domination and is both physical and in cyberspace.[14] The contest in the gray zone has ambiguous characteristics somewhere between peace (or cooperation) and war (or armed conflict). It does not cross the threshold to the point where the U.S. president can clearly declare a state of war. Any conventional Marine Corps force deployment to Africa has become a precarious domain for U.S. strategic culture and policy makers and military elites.[15]

Both the political and military elites might hesitate as deployment traditionally has been when conflict or war is clearly identifiable, and this is not the case even against jihadist movements in the gray zone when they are not posing an immediate threat. There might be cells of jihadists within the local population widely dispersed and in rural areas far from U.S. geostrategic interests. They might also be in cells of two or three in a village of 10,000 and not openly definable as combatants (e.g., not wearing identifiable uniforms). Even if a decision is taken to deploy the Marine Corps against them, the contest might be asymmetrical. A more suitable security force would be police to make arrests, if available.[16]

Adding to such hesitation has been previous Marine Corps deployments to Africa that have been met with varying degrees of success and failure in 1992.[17] Further hesitation might stem from lessons learned from hazards of other direct high-profile interventions, for example the Marine Corps deployment to Lebanon in the Middle East, when there was no clearly defined exit strategy.[18]

A dichotomy prevails for while there might be hesitancy for deployment arising from previous experiences that have not been overtly successful, for example Somalia, there are also clear geostrategic interests and military hot spots necessitating U.S. Marine Corps deployment. A clear threat analysis needs to be included in FD2030 to ameliorate this dichotomy. The geostrategic interests and military hot spots have been identified in a document released by the White House in August 2022, where President Joseph R. Biden spoke of the U.S. strategy toward Sub-Saharan Africa. From this it can be construed the specific geostrategic locations where the Marine Corps might be deployed to support

such a strategy. These would include the location of minerals such as uranium in Niger, the maritime sea routes and choke points of the Cape Route, the Suez Canal, and the Straits of Gibraltar, the southern flank of NATO that is North Africa, especially Libya and off the coast of East Africa in the Indian Ocean to protect shipping and trade.[19]

Among the specific military hot spots that can be identified for Marine Corps deployment are against violent extremist organizations like al-Qaeda, ISIS, and Hezbollah branches that operate in Somalia, the Sahel, the Maghreb, Lake Chad, and most recently in Congo and Mozambique. Intertwined is the gray zone that causes instability and conflict and is creating illegal migration to Europe that is a cause for deep concern. Also, a potential focus is the ongoing conflict in Yemen and the IRGC forces operating in the Red Sea off the African east coast—a region that has seen naval piracy. [20] This piracy has declined but other maritime crimes have increased. Illegal fishing along with smuggling and trafficking of people and illicit items such as narcotics are all on the rise. These and the protection of the ships delivering humanitarian aid, for example by the United Nations World Food Programme, is also seen as a priority in these hot spots.[21]

At the time of the writing of this article in January 2023, it is not known how many military missions that the United States has undertaken or is currently undertaking in Africa or how many troops are deployed. That fact remains top secret. However, the White House strategy also stresses joint and combined operations and support other than direct military intervention. This should also be included in FD2030 to determine how the Marine Corps could work with other branches, for instance United States Africa Command (AFRICOM) and the National Guard.

AFRICOM operates with African states and European allies and NATO in the mission to counter transnational threats and malign actors. This is accomplished by training and equipping local and regional security forces, in the provision of economic, education, and environmental assistance and expertise and overall advancing U.S. national interests through assistance, development, education, and training programs.[22] AFRICOM informs that it is engaged in West Africa and the Sahel, North Africa, Central Africa, and East Africa.[23] It is also reported that U.S. National Guard troops have or are still deployed to the Horn of Africa. This includes Djibouti, Kenya, and Somalia.[24]

Lessons from the Past

Lessons from the Marine Corps deployment to Somalia in Africa (1992–94) are not that dissimilar to lessons from its deployment to Lebanon in the Middle East (1982–83). A large force deployment with high-level intervention into violent urban areas where there is no local stable governance could better be achieved by shorter deployment and precision strikes. In both instances, the Marine Corps withdrew after facing unacceptable causalities.[25]

Such specific examples and lessons are not mentioned in FD2030 but an overall objective of FD2030 calls for the Marine Corps to be restructured for

just that type of deployment. That is, FD2030 calls for the Corps to be a lighter, faster, and more lethal service—one that can perhaps integrate Marines and sailors into versatile "stand-in forces" that can respond to an array of crises. However, just having an appropriate force structure would not guarantee that the Marine Corps would be more successful than before. Specific lessons from previous Marine Corps deployments should be included to enhance planning and preparation.

An example that should be included is the Marine Corps deployment to Somalia. On 9 December 1992, President George H. W. Bush ordered 1,800 Marines to Mogadishu, Somalia, to spearhead a multinational force aimed at restoring order. Their role was part of a larger United Nations humanitarian effort after the collapse of the Somali government. There were some successes by the U.S. troops; international aid workers were soon able to restore some food distribution and other humanitarian aid operations.[26]

However, without law and order, rival factions and militia groups emerged. The Marines found themselves in roles they had not undertaken before and were not prepared for it. They were in the crossfire of the militia groups, operating in violent urban environments for protracted periods with many patrols, and could not easily defend themselves. Also debatable were the effectiveness of rules of engagement. The Marines progressively found their main mission was their own protection. More failures than successes, especially the downing of a Black Hawk helicopter, led President William J. "Bill" Clinton to order all U.S. troops to withdraw from Somalia by 31 March 1994.[27]

Both on the tactical and strategic levels, the Marine Corps was not prepared and with a mismatch between force structure and objectives it did not achieve the objectives. On the tactical level, the intervention in Somalia was considered a failure due to the daily mayhem in the streets of the capital city of Mogadishu, which bedeviled the security operation. On the strategic level, when the Marine Corps arrived there was a lack of a national Somali leadership and when they departed there was still no functioning government. The Marines had no mission capability, nor were they tasked with the role to establish stable governance in Somalia that would have been a prerequisite for the objective of any sustainable humanitarian effort.

Examining the experience of U.S. allies is also a valuable tool. France has launched many expeditionary missions in Africa, especially in the Sahel and Chad. It has learned similar lessons to the United States in Somalia—a lack of sustainable, stable local governance coupled with the inability of African regional forces to support them. For instance, the Economic Community of West African States (ECOWAS) and the G5 Sahel have been catalysts to determine that small forces should only be deployed for short precision-type missions.[28] Large forces deployed for a long period spend more time defending themselves than anything else. Such lessons have also been learned from United Nations and NATO deployed to Sudan, for example.[29]

Most recently in 2022 and early 2023, France has announced that French

forces in Mali and Burkina Faso would withdraw after nearly 10 years of fighting insurgents and jihadists. This was due to Mali's military junta's cooperation with Russian mercenaries from the Wagner Group and Burkina Faso's request for France to do so. France has even withdrawn its ambassador from the latter in January 2023.[30]

Others including Germany, the UK, and the European Union force contingent have followed suit to withdraw forces from both countries. They have noted that leadership in Mali and Burkina Faso that faced a coup have not been cooperative and so their own presence is seen as foreign intervention rather than foreign assistance.[31]

Such contestation are elements of great power competition that should be included in the threat analysis of FD2030. China and Russia have apparently gained the advantage as this withdrawal has opened the door for them to enter these countries, and they have done so as advisors and trainers, while reaping numerous economic and mineral deals.[32]

Learning from such deployments should be included in FD2030 to signal that there are instances where the Marine Corps does not need to be deployed landward as a large force. Adding to this is another example that highlights that the correct force needs to be chosen for the mission. However, unless suitably trained and equipped, the Marine Corps, even as a small precision strike force, is not the correct one. One example is the attack on the U.S. consulate in Benghazi, Libya, on 11 September 2012. The terror group Ansar al-Sharia undertook a premeditated attack that resulted in the deaths of both the U.S. ambassador and a U.S. foreign service information management officer as well as two Central Intelligence Agency (CIA) contractors. Due to transportation challenges, it was not even possible to deploy U.S. Marine Corps Fleet Antiterrorism Security Teams (FAST) or even unmanned, unarmed surveillance drones.[33]

In both Somalia and Libya, lessons have been learned and the Marine Corps has not been subsequently deployed landward. An example is the most recent Marine Corps deployment to Africa in December 2020 as part of Operation Octave Quartz. The Makin Island Amphibious Ready Group consisting of the 15th Marine Expeditionary Unit was available offshore Somalia and not landward. Their mission was the protection of the withdrawal of American forces from that country where their presence offshore was aimed to serve as a viable force multiplier and as a deterrent to escalation.[34]

While learning from past failures and from the experiences of allies, it is also important to learn from where the United States has had the greatest success in working with local partners and to include this in FD2030. Lessons applied from Libya are an impetus to liaisons and work with locals and have more viable rapid reaction forces. This was applied working with People's Protection Units (*Yekîneyên Parastina Gel* or YPG) in 2014 in Syria and the use of the Counter Terrorism Service (CTS) in Iraq to combat ISIS jihadists. The United States provided air support to the YPG during the siege of Kobanî and during later campaigns. It helped the YPG defend territory against attacks by ISIS.

The Syrian side in the civil war was supported by Russia, and so this conflict involved both great power competition and violent extremist organizations.[35] Another example is when the United States worked with the CTS in Iraq deploying special forces.[36] In both cases the specific context of the threat analysis determined the force design. [37] Such specific contextual assessment should be included in FD2030 with a specific threat analysis of Africa.

Decluttering the Gray Zone across Africa

Learning from the past—both failures and successes—and from others is a valid methodology for strategic guidance. Just as valid is identifying the adversary and its capability and preparing and planning a Marine Corps force design and structure with appropriate weaponry. The gray zone as described is a cluttered battlespace given that it is urban warfare, where in the crowded environment it is difficult and problematic to easily distinguish between civilians and combatants as, for example, the latter might not wear uniforms. Such a battlespace with potentially ambiguous targets makes it hard to acquire, understand, track, and to apply military effects and forces with precision. It is also a cluttered battlespace in that there are many different types of adversaries, sometimes with different goals and sometimes with overlapping intentions. These different types include the local state's security forces, great power competition using proxy forces, violent extremist organizations, local militias, and local and international organized criminal networks. Therefore, it is cluttered because there are multiple adversaries in multiple guises presenting multiple threats that require multiple scenarios and probability analysis to be included in FD2030. The methodology of preparing and planning for these has been learned from previous insurgent and terrorist events such as the Madrid bombings in 2004. This requires precise and valid evaluations and implementation of security, intelligence, and counterintelligence.[38] The value of these will be to de-clutter the gray zone, namely to identify and to provide the Marine Corps with a precise adversary, its location, and its threat capability thereby enabling the appropriate size and shape of any deployment with the necessary preparation and planning.

The common thread running through all such strategic guidance is that the gray zone inevitably must be decluttered for U.S. policy makers to be confident when deploying a Marine Corps force. This is easier said than done, for gray zone adversarial activities are not cataclysmic but tend to be gradual. There is not a clear condition of war. This is evident both by activities of states in great power competition, nonstate actors in violent extremist organizations, and others, for example, organized crime. When their gradual adversarial activities are classified by U.S. criteria as lower than the threshold of armed conflict, U.S. policy makers will not be assured that a Marine Corps deployment would not become the cause for an escalation to war and thus the United States would be blamed for such foreign intervention—so they will not deploy. That might result in a "too little too late" syndrome emerging.[39]

Decluttering the gray zone can be systematic by segregating the known

from the unknown. For example, the location of states is known with borders on a map. The governance of states are identifiable people and includes the bureaucratic organization (e.g., political and military elites). Data can be gathered about their intentions and their state's military capability in both manpower and equipment. The effectiveness and readiness of these can be observed during exercises and so they become a known quantity and quality should the need arise to engage them in combat. It is possible to determine to what degree they are aligned to U.S. interests.[40]

More challenging is the nonstate-based threat environment, as the activities occur between war (or armed conflict) and peace (or cooperation). Many activities fall into this turbid situation in the gray zone. For example, organized crime, including narcotics and weapon smuggling, insurgent movements, lone-wolf terrorists, religiously motivated social movements and fanatics, cyber threats, and illegal migration patterns.[41] There could also be multiple overlaps where local actors could be acting with great power competition support or as a proxy to them that would mean that each has their own objectives. Deterring or dissuading one element might be effective on one level but not another.[42]

An example is to be found in Iranian proxy Hezbollah operations in West Africa that have also been identified as being linked to money laundering activities. Arrests and breaking the latter activity have not ended the former's destabilizing presence or Hezbollah's recruitment of locals for Iran's global Shia Islamic revolutionary movement.[43] Another example is al-Qaeda in the Islamic Maghreb (AQIM) that is both an Islamic fundamentalist group and is also engaged in drug smuggling in Mali and Niger.[44]

An example of the entanglement of great power competition and violent extremist organizations in the AQIM situation also highlights the cluttered gray zone in Mali. France intervened in Mali and worked with local and Chadian forces to upend the AQIM in north Mali. The French special forces operated as light infantry in armored personnel carriers, but it should be noted that 59 French soldiers were killed.[45] The success was short-lived and since late 2022 France has begun withdrawal of forces from Mali and Burkina Faso, where in addition to great power competition with increased Russia presence in Mali, there is also a growth of violent extremist organizations and gray zone activities. For example, AQIM has an increasingly active presence in Mali as an Islamic fundamentalist movement and is also engaging in smuggling.[46]

These examples are just the tip of the iceberg that justify why Africa should be included in the threat analysis of FD2030 looking at great power competition, violent extremist organizations, and gray zone conflict. Africa is a diverse and unstable environment with multiple numbers of adversaries engaged in overlapping activities and connections.[47] Caution should be taken that such a complex environment does not lead to organizational complexity of overly prescriptive force design in the strategic guidance in the form of multiple bureaucratic levels, which would inhibit fluid operational-organizational inertia on the tactical level.[48]

To forestall such an eventuality, strategic guidance could suggest that the most appropriate tactics would be a short period precision strike deployment of the Marine Corps to a hot spot. That would require security and intelligence analysis on an ongoing basis to determine the right moment to deploy and to withdraw for the greatest operational effect. This entails long-term planning and preparation that needs to include extreme options; for example, deploying the Marine Corps for less than 24 hours at less than 6 hours' notice. As already noted in FD2030, drones can play a greater role. At the same time caution needs to be applied as too much data from surveillance and reconnaissance without accurate analysis would not enable decision makers to be more efficacious.[49]

Therefore, a diverse force design should be included in FD2030 to cover multiple options given the cluttered gray zone in Africa. Hybrid warfare is a way that could be applied to suit such a diverse force design. It has been defined as a fusion of different tools and instruments. Options that should be included are a blend of the conventional force of the Marine Corps, drones, irregular warfare by special forces, partnerships with other U.S. military branches such as AFRICOM and the National Guard—and including cyberwarfare.[50]

Security

Decluttering the gray zone would be dependent on having accurate information and analysis. This will require security, intelligence, and counterintelligence. While there is a strong link between them, they are sometimes at odds with each other. For instance, there can be organizational competition and reluctance to share data and analysis.[51]

Unless this is overcome, decluttering the gray zone in Africa will be compounded. Each require clear definition and role and task assignment that should be included in the strategic guidance of FD2030 supporting the specific force design. For example, security and protection of U.S. interests is an existential rationale for the Marine Corps. Security objectives could be to establish a short period foothold in a hot spot or to deter an escalation while other means are employed such as diplomacy. Protection could be supplied to vital installations such as ports or as a force multiplier when other forces withdraw as seen in the deployment offshore Somalia in 2020.[52]

To improve threat analysis and regulate tackling nonstate adversaries below the threshold of war, the suitable security factors should be aligned with intelligence. At the top of the list would be to ascertain when the gradual escalation by great power competition and violent extremist organizations has reached the point that would require the Marine Corps to move from a passive offshore presence to that of an active landward deployment.[53]

The nature of the intelligence product on the multiplicity of local actors, violent extremist organizations, overlapping gray zone activities, and links to great power competition would serve to classify the descriptor of security needs and whether the Marine Corps would partner with others. These partners could

include the Navy, AFRICOM, European allies, and any form of hybrid warfare such as cyber. An integral element of the intelligence product would also need to determine the required logistics. One of the challenges is the size of Africa, the second largest continent after Asia. Africa is three times the size of Europe with the terrain that is diverse and includes both deserts and jungles.[54] Every Marine Corps deployment would be unique and complicated as Africa has 54 sovereign states of which 38 are coastal and several island nation-states.[55]

From lessons learned from Somalia, for the security of the Marine Corps force, the intelligence product would also need to identify who and where the adversary is. Nonstate adversaries might not wear uniforms and as they were recruited from the local population could conceal themselves therein. They could receive housing and food support from it; additionally, in a failed state they have ungoverned territories that can provide safe haven for them to hide in.[56]

Intelligence

The intelligence product therefore requires data and analysis on all aspects of great power competition in Africa, violent extremist organizations, and gray zone actors.[57] The takeaway from this intelligence product to be included in FD2030 would be the warning signs in threat analysis that would trigger a Marine Corps deployment. As gray zone activities include denial and deception efforts and stealth, the strategic guidance needs to register unexpected outcomes rather than cataclysmic changes.[58]

Identifying such unexpected outcomes for the threat analysis can be classified into two categories: puzzles and mysteries. To be effective for both, intelligence gathering will need to penetrate certain specific communities within the overall society, namely human intelligence (HUMINT), especially those that have been identified as recruitment grounds for violent extremist organizations.[59]

Puzzles have a definite answer and intelligence needs to find it. Puzzle type intelligence can be applied to various partnerships between the Marine Corps, the Navy, AFRICOM, the National Guard, and African states. Examples are to build maritime safety and security, to counter illicit trafficking, to address humanitarian needs, to promote regional stability and security, to strengthen local, regional, United Nations, and African Union combined operations, and to encourage sustainable development. The role of the Marine Corps in these could range from active and passive protection and deterrence to escalations.[60]

Mysteries have no definite answer where any answer could be contingent on other factors. Mystery type intelligence tends toward analysis that offers a best forecast or probable scenario. The intelligence product tends toward sense-making for responses as a different combination of the same factors could lead to a different outcome. This was the case with Somalia and Libya.[61] Mysteries are the context where the Marine Corps is best not deployed. For example, a lone-wolf terrorist is better left to counterintelligence efforts to trick him rather

than sending boots on the ground to intercept them. The strategic guidance of FD2030 should also note this as a limitation for Marine Corps deployment.

Counterintelligence

Without an excellent intelligence product, no Marine Corps force could deploy successfully. Counterintelligence also has a part to play in U.S. tactics in Africa though this is a different role from intelligence. It can be used when the Marine Corps cannot be deployed or in lieu of it or to supplement and complement a deployment. It does not necessarily aim to offer security or to protect people, physical territory, or even information in cyberspace. Whereas intelligence to support Marine Corps operations may struggle on puzzles and mysteries in threat analysis, counterintelligence can operate and be successful with less uncertainty.[62]

Examples of intelligence and counterintelligence cannot be released due to secrecy, but the National Intelligence Council has described their significance in a memorandum updated on 4 August 2022.[63] From this memorandum it is possible to ascertain that counterintelligence has a role in hybrid warfare with the Marine Corps in the gray zone in Africa to trick adversaries into responding to classified information released about Marine Corps exercises or obsolete Marine Corps plans. In this way it is sometimes contrary to security-driven deployment and could lead to a dispute with intelligence over the release of such material. Another tactic is to restrict normally unclassified or "open source" information (OSINT), for example Marine Corps collaboration with the African Union, if it is known that adversaries were using it to further their own purposes for purposes such as extortion.[64] In both instances such tactics could serve to disorientate and trick a violent extremist organization/gray zone adversary to reveal its intentions and location.[65]

Nonetheless, both counterintelligence and intelligence using HUMINT and OSINT can be a double-edged sword and counterproductive to security and objectives. They might rely heavily on monitoring and gathering data from the technologies and services provided by mobile/cell telecommunications, the internet, and social media. However, any violent extremist organization or gray zone adversary can also do so, especially if they have the support of other countries' counterintelligence and intelligence services in great power competition.

Dictatorial totalitarian states that typify Africa can also use these same services and technologies to monitor, censor, and subjugate their population.[66] Various violent extremist organizations such as ISIS, Boko Haram, and al-Shabaab have also used them for recruitment and psychological influence purposes.[67] A vivid example was live commentary with photos by al-Shabaab of its attack on the Westgate shopping Mall in Kenya in 2013 on Twitter.[68] But this can be used by U.S. cyber teams to ascertain the physical location of such violent extremist organizations when they broadcast and so determine where, when, and how large a Marine Corps precision strike force to deploy—and this should be included in the strategic guidance of FD2030.[69]

Conclusions

The 2018 U.S. *National Defense Strategy* was the impetus for FD2030 in that it called for changes in American forces after evaluating the threat environment and finding that there was a need to build a more lethal force and implement reforms for greater performance.[70] While FD2030 contains many positive elements as a strategic guidance to meet this call, there are also certain elements that been criticized while there are clear omissions.

For example, in being critical of FD2030, three senior retired Marine Corps officers (Colonel Gary Wilson, Lieutenant Colonel William A. Woods, and Colonel Michael D. Wyly) started their article by noting that the words "Send in the Marines! The situation is serious. We need to fix it—fast!" have a special meaning.[71] They truthfully inform that the Marine Corps has for centuries proven themselves in battle as a reliable force. However, these retired officers have also spoken out over their concerns that FD2030 abandons the principles of maneuver warfare and has an overreliance on technology. In their view, the threat analysis of FD2030 and therefore its strategic guidance for restructuring is oriented toward the Marine Corps fighting units operating in the Indo-Pacific region.[72]

The author of this article concurs that the threat analysis of FD2030 is too specific and should also include great power competition, violent extremist organizations, and gray zone in Africa as examined here. As it presently stands using the strategic guidance of FD2030 and its suggested force design, deploying Marine Corps "boots on the ground" will not necessarily bring success and victory in Africa. It is fair to state that despite such deficiencies the White House together with the Pentagon will continue to look to the Marine Corps to be deployed to Africa. There is no other U.S. military branch, together with the Navy, that could defend the geostrategic concerns and tackle the hot spots mentioned in this article.

As a matter of priority to ameliorate the deficiencies outlined above, FD2030 needs to include lessons learned from the failures in Somalia, Libya, and Lebanon. These clearly show that the U.S. Marine Corps at the time of its deployment was not fit for the intended purpose of policy makers. At present, it is also not fit for purpose in a deployment to Africa. Lessons learned from successes elsewhere, for instance in Syria and Iraq, should also be included in FD2030 as well as the experiences of others in Africa like the French in Mali and Burkina Faso.

Lessons from the past and from others are only one item. Other entries to be included are those examined in this article on great power competition, violent extremist organizations, and gray zone contests and competition in Africa that are challenging. The threat environment has multiple actors engaged in multiple overlapping activities. The overriding concerns noted in this article conclude that without decluttering the great power competition/violent extremist organizations and gray zone in Africa, the Marine Corps is not going to be able to engage and easily combat the adversary and cause of the threat,

especially nonstate actors and insurgent movements who are an extension of Middle East fundamentalism, nor others such as the Russian Wagner group and local Sub-Saharan African terrorist groups.

Options for strategic guidance and a force design that have been suggested in this article are, for a short period, precision strike deployment of the Marine Corps dependent on precise intelligence on the location, size, strength, and intentions of the adversary. Counterintelligence can play a role to trick an adversary into revealing these details. Victory could also be attained through hybrid warfare dissuading and deterring using information or psychological operations. Here also drones could be used as recommended by FD2030 and software (algorithms) could play a role in ensuring mission success.

With a restructured force design, the Marine Corps could still deploy in its traditional role offshore as a formidable deterrent force and as a force multiplier partner with the Navy, special operation forces, AFRICOM, the National Guard, African states, and European allies, both in precision strikes landward and beyond. The Marine Corps could collaborate in security cooperation including training and education, humanitarian assistance, medical readiness, development strategies, and interdiction of illicit activities. In doing so, the policy of containment and its strategies that prevailed prior to 9/11 would also be furthered. The domino effect of reducing or eliminating great power competition, violent extremist organizations, and gray zone conflicts in Africa would enable U.S. forces to be concentrated elsewhere.

Endnotes

1. *Force Design 2030* (Washington, DC: Headquarters Marine Corps, 2020).
2. Donald C. Bolduc, Richard V. Puglisi, and Randall Kaailau, "The Gray Zone in Africa," *Small Wars Journal*, 29 May 2017.
3. Selcan Karabektas, "When the Powers of the Middle East Export Their Rivalries to the Horn of Africa," *Beyond the Horizon*, 25 April 2022.
4. Thomas G. Mahnken, *United States Strategic Culture* (Fort Belvoir, VA: Defense Threat Reduction Agency Advanced Systems and Concepts Office, 2006).
5. "Violent Extremism," United States Institute of Peace, accessed 11 April 2023.
6. Karl P. Mueller et al., *Striking First: Preemptive and Preventive Attack in U.S. National Security Policy* (Santa Monica, CA: Rand, 2007).
7. *Insurgencies and Countering Insurgencies*, Marine Corps Warfighting Publication 3-33.5 (Washington, DC: Department of Defense, 2014).
8. *Force Design 2030*. The open-source web version shows graphics of paradigm shifts.
9. Paul Collier, "Security Threats Facing Africa and its Capacity to Respond," *Prism* 5, no. 2 (2015): 31–34.
10. Bolduc, Puglisi, and Kaailau, "The Gray Zone in Africa."
11. George Popp and Sarah Canna, "The Characterization and Conditions of the Gray Zone," *NSI* (Winter 2016).
12. "Nigerian Court Charges Iran Revolutionary Guard Member over Arms Seizure," France 24, 25 November 2010.
13. Ian Davis, "Armed conflict, and Peace Processes in Sub-Saharan Africa," in *SIPRI Yearbook 2022* (Oxford, UK: Oxford University Press, 2022).
14. Julie L. Marble et al., "The Human Factor in Cybersecurity: Robust & Intelligent Defense," in *Cyber Warfare: Building the Scientific Foundation*, ed. Sushil Jajodia et al.,

Advances in Information Security 56 (Cham, Switzerland: Springer, 2015), https://doi.org/10.1007/978-3-319-14039-1_9.
15. Colin S. Gray, "Strategic Culture as Context: The First Generation of Theory Strikes Back," *Review of International Studies* 25, no. 1 (January 1999): 49–69, https://doi.org/10.1017/S0260210599000492.
16. Ido Levy, "Fighting Jihadists By, With, and Through U.S. Partners: Lessons Learned and Future Prospects," Washington Institute Policy Notes 124, 20 October 2022.
17. Robert F. Baumann, Lawrence A. Yates, and Versalle F. Washington, "*My Clan Against the World": U.S. and Coalition Forces in Somalia, 1992–1994* (Fort Leavenworth, KS: Combat Studies Institute Press, 2004).
18. Benis M. Frank, *U.S. Marines in Lebanon, 1982–1984* (Washington, DC: Headquarters Marine Corps, 1987).
19. *US Strategy Toward Sub-Saharan Africa* (Washington, DC: White House, 2022).
20. Jon Gambrell, "US Navy Says New Force to Patrol Red Sea Amid Attacks by Iran-backed Yemen Rebels," *Times of Israel*, 14 April 2022.
21. Details of the United Nations World Food Programme and where they work can be found at https://www.wfp.org/.
22. "About the Command," U.S. Africa Command, accessed 12 April 2023.
23. "What We Do," U.S. Africa Command, accessed 13 April 2023.
24. Eric Durr and SSgt Alexander Rector, "New York Army National Guard Troops Deploy to Africa," U.S. Army, 9 September 2022.
25. Kenneth Allard, *Somalia Operations: Lessons Learned* (Washington, DC: CCRP Publications, 1995).
26. Chester A. Crocker, "The Lessons of Somalia: Not Everything Went Wrong," *Foreign Affairs* 74, no. 3 (May–June 1995): 2–21, https://doi.org/10.2307/20047117.
27. F. M. Lorenz, "Rules of Engagement in Somalia, Were They Effective?," *Naval Law Review*, no. 42 (1995): 62–87.
28. "West Africa and the Sahel," Security Council Report, January 2023.
29. Glen Segell, "The First NATO Mission to Africa: Darfur," *Scientia Militaria* 36, no. 2 (2011): 1–18, https://doi.org/10.5787/36-2-49; and Glen Segell, "The United Nations Africa Union Mission in Darfur—UNAMID," *Strategic Insights* 7, no. 1 (February 2008).
30. "France Recalls Its Ambassador from Burkina Faso after Demands to Pull Out Troops," France 24, 26 January 2022.
31. Glen Segell, "The External Actors: What Is the Way Forward for the Sahel?" (conference paper presented at the Security Institute for Governance and Leadership in Africa, University of Stellenbosch, Exploring the Interface Between Coups, Violent Extremism, and Poor Governance in the Sahel, 26 January 2023).
32. Marie Sandnesm and Ilaria Carrozza, "Russia, China and New Power Dynamics in the Sahel Region," *PRIO* (blog), 16 January 2023.
33. Mike Bailey and Dan Yurkovich, "Benghazi Consulate Attack," *Marine Corps Gazette*, November 2016.
34. Richard Sisk, "A US Naval Armada and 2,500 Marines Are Off Somalia to Cover Troop Withdrawal," Military.com, 22 December 2020.
35. Doruk Ergun, "External Actors and VNSAs: An Analysis of the United States, Russia, ISIS, and PYD/YPG," in *Violent Non-state Actors and the Syrian Civil War*, ed. Özden Zeynep Oktav, Emel Parlar Dal, and Ali Kurşun (Cham, Switzerland: Springer, 2017), 149–72.
36. David Witty, *The Iraqi Counter Terrorism Service* (Washington, DC: Center for Middle East Policy at Brookings, 2016).
37. "Special Operators in Syria Must be Viewed in Context, Dunford Says," DOD News, 4 November 2015.
38. Glen Segell, "Intelligence Methodologies Applicable to the Madrid Train Bombings," *International Journal of Intelligence and CounterIntelligence* 18, no. 2 (2005): 221–38, https://doi.org/10.1080/08850600590882119.

39. Robert J. Giesler, "Today's Wars Are Fought in the 'Gray Zone'," Atlantic Council, 23 February 2022.
40. John Patrick Finnegan and Romana Danysh, *Military Intelligence* (Washington, DC: U.S. Army Center of Military History,1998).
41. Clementine G. Starling, "Today's Wars Are Fought in the 'Gray Zone'," Atlantic Council, 23 February 2022.
42. Devin Ellis et al., "Gray Zone Crises in MENA and Eastern Europe," START, March 2017.
43. Ibe Okegbe Ifeakandu and Habila Ardzard, "The Role of Institutional Framework in Entrenching Effective Anti-Money Laundering/Combating Terrorist Financing in West Africa (GIABA) in Perspective," *Beijing Law Review* 13, no. 3 (September 2022): 575–93, https:/doi.org/10.4236/blr.2022.133037.
44. Luca Raineri and Francesco Strazzari, "Drug Smuggling and the Stability of Fragile States. The Diverging Trajectories of Mali and Niger," *Journal of Intervention and State-building* 16, no. 2 (2022): 222–39, https://doi.org/10.1080/17502977.2021.1896207.
45. Giovanni Carbone and Camillo Casola, eds., *Sahel: 10 Years of Instability: Local, Regional, and International Dynamics* (Milan, Italy: Ledizioni LediPublishing, 2022).
46. "Russian Troops Deploy to Mali's Timbuktu after French Exit," Al Jazeera, 7 January 2022.
47. David Carment and Dani Belo, "Gray-zone Conflict Management: Theory, Evidence, and Challenges," *Air Force Journal of European, Middle Eastern, and African Affairs* 1, no. 1 (Spring 2019).
48. John Raine, "War or Peace?: Understanding the Grey Zone," *ISS Analysis*, 3 April 2019.
49. Ian Williams, "Eyes Everywhere: Intelligence and Strategic Decision-making in the Gray Zone," *CSIS Analysis*, 23 November 2021.
50. Frank G. Hoffman, *Conflict in the 21st Century: The Rise of Hybrid Wars* (Arlington, VA: Potomac Institute for Policy Studies, 2007).
51. Example provided by Jennifer Sims, "The Future of Counter-intelligence: The Twenty-first-Century Challenge," in *The Future of Intelligence: Challenges in the Twenty-first Century* (London: Routledge, 2014), 60–61.
52. Jahara W. Matisek, "Shades of Gray Deterrence: Issues of Fighting in the Gray Zone," *Journal of Strategic Security* 10, no. 3 (Fall 2017): 1–26, https://doi.org/10.5038/1944-0472.10.3.1589.
53. Heather M. Bothwell, "Gray Is the New Black: A Framework to Counter Gray Zone Conflicts," *Joint Force Quarterly*, no. 101 (2021).
54. "Africa, History, People, Countries, Regions, Map, & Facts," Encyclopedia Britannica, 14 April 2023.
55. "Africa," Encyclopedia Britannica, 13 April 2023.
56. James A. Russell and Aaron May, "Tomorrow's Proliferation Pathways: Weak States, Rogues, and Non-States" (paper presented at Naval Postgraduate School Conference Report, Belfast, ME, 17–18 July 2008).
57. *Army Futures Command Concept for Intelligence 2028* (Austin, TX: Army Futures Command, 2020).
58. Emily Harding, McKenzie Richardson, and Matthew Strohmeyer, "From Data to Insight: Making Sense out of Data Collected in the Gray Zone," CSIS, 20 October 2021.
59. *Human Intelligence Collector Operations* (Washington DC: Department of the Army, 2006).
60. "Security Cooperation," U.S. Africa Command, accessed 13 April 2023.
61. Framework provided by Gregory F. Treverton, "The Future of Intelligence: Changing Threats, Evolving Methods," *The Future of Intelligence*, ed. Isabelle Duyvesteyn, Ben de Jong, and Joop van Reijn (London: Routledge, 2014), 27–48.
62. Hank Prunckun, *Counterintelligence Theory and Practice* (Lanham, MD: Rowman & Littlefield, 2019).

63. *Africa: Increasing Weight in the Global Arena* (Washington, DC: Office of the Director of National Intelligence, 2022).
64. Gašper Hribar, Iztok Podbregar, and Teodora Ivanuša, "OSINT: A 'Grey Zone?'," *International Journal of Intelligence and CounterIntelligence* 27, no. 3 (2014): 529–49, https://doi.org/10.1080/08850607.2014.900295.
65. Peter F. Kalitka, "Counterintelligence: A Law Enforcement Function," *American Intelligence Journal* 8, no. 2 (May 1987): 11–13.
66. Chino Takahiro, "The Modern State and Future Society: Gramsci's Two Conceptions of the 'Ethical State'," *European Legacy* 27, no. 2 (2022): 125–42, https://doi.org/10.1080/10848770.2021.2001888.
67. Kate Cox et al., *Social Media in Africa Presents Double-edged Sword for Security and Development* (Santa Monica, CA: Rand, 2018).
68. Karen Allen, "Terrorists' Use of Tech in West Africa Must be Contained," *ISS Today*, 15 September 2022.
69. *The European Union and the African Union: A Statistical Portrait* (Luxembourg City: Eurostat, 2016).
70. *Summary of the 2018 National Defense Strategy of the United States of America: Sharpening the American Military's Competitive Edge* (Washington, DC: Department of Defense, 2018).
71. Gary Wilson, William A. Woods, and Michael D. Wyly, "Send in the Marines?: Reconsider Force Design 2030 Beforehand," DefenseNews, 4 August 2022.
72. Wilson, Woods, and Wyly, "Send in the Marines?"

The Deficiency of Disparity
The Limits of Systemic Theory and the Need for Strategic Studies in Power Transition Theory

Athahn Steinback and Steven Childs, PhD

Abstract: This article synthesizes power transition theory (PTT) at the grand strategic scale with military studies methods at lower levels of analysis. We analyze the Russo-Japanese War, the recent Afghan War, and the ongoing war in Ukraine as conflicts where political-military specificities enabled outmatched powers to win or force a stalemate. These cases demonstrate the decisive influence of power projection, doctrine, geopolitical constraints, and readiness on conflict outcomes. Finally, the authors operationalize PTT at the grand strategic scale alongside military studies methods at the operational level to propose U.S. responses to Chinese regional revisionism.

Keywords: power transition, Russo-Japanese War, Afghanistan War, war in Ukraine, China-Taiwan crisis

Introduction

Power transition theory (PTT) offers an effective systemic theory to explain competition between states but struggles to predict the outcome of specific conflicts due to reliance on broad metrics of national power. PTT primarily estimates national power by comparing the Composite Index of National Capability (CINC) scores or gross domestic products (GDP) of rival states. By focusing on CINC and GDP, PTT implicitly assumes total economic

Athahn Steinback is an alumnus of the national security studies and social sciences and globalization graduate programs at California State University, San Bernardino. As a simulation designer, he specializes in modeling combined arms warfare. As an independent researcher, he analyzes factionalism within autocracies. Steven Childs is an associate professor in the Department of Political Science at California State University, San Bernardino, where he teaches in the national security studies graduate program. His research interests include conventional arms proliferation, nuclear deterrence, and the security politics of the Middle East and Asia regions. His scholarship has appeared in *Comparative Strategy*, *Defense & Security Analysis*, and the *Journal of Advanced Military Studies*.

Journal of Advanced Military Studies vol. 14, no. 1
Spring 2023
www.usmcu.edu/mcupress
https://doi.org/10.21140/mcuj.20231401009

mobilization while omitting case-specific influences including power projection and readiness. The authors explore the Russo-Japanese War, Afghan War (2001–21), and ongoing war in Ukraine as cases where military studies methods provide more compelling explanations of a conflict's course. The authors then use the potential case of a U.S.-China war over Taiwan to synthesize PTT's grand strategic level with a military studies approach at the operational level to demonstrate how the theory can better guide policy makers. From a policy perspective, PTT can be employed at the grand strategic level to detect emerging challengers and identify which states to mollify or isolate. Meanwhile, military studies approaches should be used in conjunction with PTT at lower levels of analysis to determine how to respond to threats. Strategic considerations such as power projection, readiness, and foreign intervention shape conflict outcomes more decisively than abstract measures of national power as wars are not fought on spreadsheets.

Merits and Limits of PTT as a Systemic Theory

Power Transition Theory serves as a leading theoretical lens for the study of conflict at a systemic level. PTT rests on two pillars. First, it assumes that the distribution of power within the system reflects a hierarchy of states akin to a pyramid with a single state at the top. Second, PTT argues all states in the system are either satisfied or dissatisfied with this dominant power's order. The dominant state constructs an order that reflects its own preferences, and the order persists as long as the majority of power within the international system remains in the hands of the dominant state and its satisfied supporters.[1] The preferences that underlie the dominant state's order are shaped by any number of factors, including history, territory, ideology, religion, culture, and so forth and the dominant state establishes institutions and norms that reflect these preferences.[2] Descending the pyramid from the few great powers at the top to the slightly more numerous middle powers, and then down to the plethora of minor states with little influence, the degree of satisfaction diminishes. Within PTT, the world consists of numerous weaker states dissatisfied with the dominant order and a small number of satisfied states wielding the majority of power perpetuating the dominant order.

Using this pyramid of power and satisfaction as a basis, PTT defines the mechanics of conflict in the international system. PTT predicts power transition conflicts frequently occur when a rising dissatisfied state approaches power parity with a dominant state, leading to either the challenger initiating a power transition conflict, or the declining dominant power striking preemptively to protect its position. Peaceful transitions can occur when a satisfied state supplants the dominant state as the United States did with the UK, but dissatisfied challengers often resort to war to impose their own preferences, as demonstrated by both World Wars.[3] The closer a dissatisfied challenger comes to power parity with the dominant state, the more likely a power transition conflict becomes. Overwhelming power deters challenges, while parity invites them.[4]

While much PTT scholarship largely revolves around global power transition conflicts, the theory applies equally to regional and subregional power structures as well.[5] Within PTT, a dominant regional power can simultaneously be a revisionist state against the current international order. Moreover, a rising global revisionist such as China may attempt a regional transition challenge against the dominant global power before it initiates a bid for global dominance. Despite PTT's limitations, it does reflect the general dynamics of systemic conflict at a highly aggregated level from the late modern period onward.[6]

Power transition theory ably draws on the most cogent elements of its major theoretical rivals to model international politics at a systemic level. From realism, PTT draws the importance of power as a central component of international relations but provides clear conditions under which conflict erupts. PTT's prescription that an imbalance of power deters conflict is more empirically sound than realism's embrace of parity as stabilizing in conflict dynamics.[7] PTT draws on liberalism to explain the persistence of hierarchy and international institutions created by dominant powers.[8] Moreover, PTT's concept that dominant powers typically establish their order in negotiation with satisfied partners, instead of unilaterally imposing them, also draws on liberal concepts of interstate cooperation.[9] PTT's recognition of hierarchy forms a solid basis to understand the persistence of peace between transition conflicts in contrast to realism's unrealistic tenet of perpetual anarchy. Finally, PTT implicitly draws on constructivism in recognizing the influence of identity and ideology in shaping preferences and animating satisfaction.[10] Through preferences and satisfaction, PTT more effectively grasps *why* individual states support the status quo or become revisionist than any exclusive focus on power itself. Collectively, PTT's blend of realist, liberal, and constructivist concepts allow it to robustly explain how power leads to conflict, why hierarchy and peace reign between wars of transition, and why some states resort to violence while others support the status quo.

Limitations of PTT as a Guide to Policy

While PTT enjoys numerous advantages over its peer systemic theories, it still suffers from inherent limitations that inhibit its ability to inform policy regarding conflict emergence and outcomes. Due to PTT's focus on systemic understanding of total power, the theory overlooks the political-military realities that characterize individual conflicts. For policy makers attempting to operationalize PTT, failure to understand these details may literally mean the difference between victory or defeat in a power transition conflict.

PTT suffers from several key limitations that have already been ably critiqued. Measuring national power by CINC or GDP sometimes creates contradictory predictions of conflict within the theory. CINC scores may suggest a power transition conflict, but GDP indicates the rising power remains outside the 80 percent power threshold necessary to initiate a challenge.[11] Efforts to base the theory's entire operation on readily quantifiable data such as GDP

or CINC have eliminated concepts including national morale and geography present in the theory's original conceptualization.[12] Finally, no consensus exists within PTT regarding *how* to measure a state's satisfaction.[13]

Beyond these critiques, the authors focus on two limitations that most severely inhibit PTT's utility to policy makers. First, sound policy making requires considering case-specific political-military variables absent in PTT's system-level approach. The theory's two levers affecting the likelihood of war, power, and satisfaction are used deterministically when applying theory to policy. According to PTT logic, to avoid war states must either increase their power versus their challengers to prevent a challenge at all or encourage their rising challengers to become satisfied to facilitate a smooth transition.[14] These pathways are not realistic. Total power does not easily lend itself to manipulation by policy levers. GDP growth and key factors within CINC such as total population, urban population, and energy consumption may take decades of bottom-up processes to meaningfully improve. Second, the concept that absent a power advantage, policy makers hoping to avoid war should rely on socializing a rising adversary into adopting the norms of a system alien to its own is excessively optimistic. The prospect of successfully socializing a dissatisfied and rising revisionist power such as China is questionable, and attempting to do so threatens to enhance the revisionist's leverage to subvert the dominant order. Dominant powers must be prepared to fight to retain their position, even if a rising adversary surpasses them in total power. Moreover, proper exploitation of political-military details can enable weaker states to fight and defeat stronger rivals. PTT helps predict when conflicts may arise, but when rivalry erupts into a war of power transition, case-specific political-military realities determine how the war unfolds.

In short, PTT focuses on the wrong levers of national policy to confront (or initiate) a transition challenge by emphasizing material elements absent strategic bearing. By focusing excessively on total measures of national power, PTT struggles to predict conflict outcomes in favor of weaker powers buoyed by strategic advantages not captured in the narrow economic logic of GDP or CINC. This critique applies to all theories that operate at such a generalized level. However, the authors believe PTT possesses great merit as a systemic theory and seek to help the theory understand political-military specificities that shape conflicts once they emerge to make PTT more useful to policy makers. Policy makers must define their interests and devise strategies to safeguard them in the complex international system. Conceptually *how* to harness a state's available power in its many manifestations matters far more than *how much* power a state is thought to have.

Deficiencies of Measuring Total National Power

Measures of total national power based on economic or material measures omits key strategic constraints necessary to understand specific conflict outcomes. GDP- and CINC-based measures of national power can be used at the grand

strategic level to estimate a state's total potential power within the international system, but these metrics should never be treated mechanistically as total power superiority guarantees little. CINC- and GDP-based power models do not consider power projection, implicitly assume total economic mobilization, and struggle to predict the impact of foreign intervention in regional conflicts. Excessive reliance on these metrics harms PTT's ability to help policy makers predict conflict outcomes unless the theory pairs itself with a military studies approach to handle analysis below the grand strategic scope.

Measuring total national power by GDP provides a generalized estimate of a state's power. GDP alone does not ensure that a state invests in its military capabilities. High GDP states can possess dysfunctional militaries insufficient to protect their foreign interests as exemplified by Germany today.[15] Likewise, the implicit technological advantages afforded by a higher GDP do not automatically equate to insurmountable military superiority. No military can equally fill every niche, and competent combatants focus their efforts on procuring technologies that exploit a rivals' weaknesses. For instance, China possesses advanced antishipping missile capabilities that largely nullify America's powerful aircraft carriers within 600 kilometers of China's coastline.[16] Thus, a key American military advantage can be mitigated by an opponent boasting lower GDP. Measuring national power by GDP also neglects the use of intellectual property theft by a lower GDP state to close technological gaps with wealthier rivals. For the purposes of PTT, GDP provides a rough measure of the total *theoretical* power of a state, but it does not capture a state's actual military capabilities.

Composite Index of National Capability (CINC) scores estimate total national power based on early twentieth-century measurements of economic and military might that fail to grasp the complexity of modern warfare or globalized economics. CINC measures each state's military expenditure, military personnel, energy consumption, iron/steel production, urban population, and total population ratios to estimate the state's total share of power in the international system.[17] Military expenditure offers only a surface-level estimate of potential force composition or capabilities. An advanced military can still suffer from poor power projection. Likewise, well-funded militaries, such as those of many Arab states, can still chronically underperform due to ineffective, politicized command structures.[18] Total personnel presents an anachronism because militaries typically become *smaller* as they professionalize and technologically advance. CINC further disregards nuclear weapons, thus conflicting with PTT's concepts that nuclear arms provide leverage, and the threat of mutually assured destruction does not intrinsically deter aggression by nuclear-armed revisionists.[19]

CINC's estimates of economic power through iron/steel production, energy consumption, and urban/total population neglect the complexities of a globalized postindustrial world economy. Iron and steel output do not measure productivity in economies powered by microprocessors and manufactured composites.[20] Second, energy consumption has never been a reliable indica-

tor of productivity, as it conflates inefficiency with output in the case of states such as the USSR that consumed more energy to produce less output.[21] Finally, CINC exaggerates the benefits of large populations by failing to account for the social, economic, and military costs of maintaining them. Middle income states with large populations such as Brazil or Indonesia consume large volumes of economic output supporting their existing population without substantially contributing to economic growth or mustering military might.[22] CINC provides accessible but imperfect insight into total national power; policy makers operationalizing PTT should only use CINC with full awareness of its limitations. Bearing these core limitations of GDP and CINC in mind, deeper challenges of measuring total national power, including power projection, limited warfare, and foreign assistance can be fully explored.

The primacy of power projection cannot be understated in interstate conflict. Power projection represents a state's ability to project military force to achieve political ends beyond its own borders.[23] The degradation of power projection across distance is nonlinear and efforts to use distance and travel time to a prospective warzone as a proxy for power projection fail to capture the complexities of power projection in modern warfare.[24] Assets such as aircraft carriers, aerial tankers, cargo aircraft, and forward bases disproportionately amplify a state's power projection capabilities at distance as demonstrated in both Gulf Wars.[25] However, antiaccess capabilities such as antishipping and antiaircraft missiles employed by a regional adversary can limit the utility of these advantages.[26] For example, even though the United States possesses unmatched *global* power projection, a weaker competitor such as China may still gain a regional power projection advantage through capabilities that deny access to American power projection assets. One such example is the Chinese militarization of artificial islands in the South China Sea to extend forward basing directly into potential combat zones. Moreover, states such as Germany and Japan that appear strong according to GDP and CINC may suffer from abysmal power projection capabilities even in their own home region.[27] Power projection is essential in predicting the possible outcomes of military escalation, but it requires case-specific analysis that PTT omits without the assistance of military studies.

Attempting to measure total national power also implicitly assumes total economic mobilization for a war effort. While PTT theorists correctly observe that economic power has become more fungible into military power over the past century, they underestimate the speed of modern warfare. State combat during the Second Gulf Wars lasted little more than a month. Likewise, the Russo-Georgian War lasted less than two weeks and Russia's forceful occupation of Crimea was complete in slightly more than one month.[28] Modern mechanized warfare moves with such speed that a political or territorial fait accompli is often reached before any meaningful level of economic mobilization can be achieved. While protracted conflicts and extensive economic mobilization still occur as demonstrated by the ongoing Russian invasion of Ukraine, protracted state conflicts are an exception, not the norm. Policy makers would do well to

remember that most wars are limited wars and total economic mobilization should not be automatically assumed.

Competing strategic interests or concerns play an equally decisive role in determining the outcome of individual conflicts. Even in cases of severe total power imbalance, such as the United States versus China in the Korean War, policy makers may deliberately temper escalation due to strategic threats posed by other states. Likewise, all the power in the world counts for little if policy makers in status quo states faced with rising revisionist adversaries decline to act due to domestic political concerns. Alternatively, a rising state such as the United States in the late 1800s may simply decline to initiate a transition challenge and focus on its own internal affairs.[29] National power is a tool directed according to the priorities of policy makers, not the iron laws of theory.

The extent and efficacy of foreign intervention in regional conflicts also eludes prediction within PTT based on GDP or CINC scores. As the United States demonstrated in Afghanistan, even decades of security assistance and billions of dollars in equipment and training does not guarantee a positive military outcome for the recipient state. Conversely, by employing limited military force and highly selective material aid, France was able to play a decisive role in ejecting Libya from Chad during the Toyota War (1987).[30] Likewise, substantial Western equipment, training, and intelligence aid to Ukraine following the February invasion has lent Kyiv key qualitative advantages over Russia.[31] Foreign intervention in regional conflicts ties so closely to case-specific political and military considerations that efforts to predict outcomes of interventions through measures of total national power becomes futile.

Due to the inherent limitations of measuring total national power, PTT struggles to predict the outcome of specific conflicts. PTT serves as a threat radar to detect likely conflicts. Using PTT, policy makers can identity which states are powerful enough to warrant mollification if satisfied or isolation if revisionist. Without accounting for power projection, PTT lacks a method to measure how much of a state's total power can realistically deploy in a specific conflict. By assuming total mobilization, the theory overemphasizes national power available for any conflict short of total war. By omitting competing strategic pressures limiting deployment of state power, PTT overestimates the power of most states in any given conflict. Finally, total measures of power are wholly inappropriate to predict the impact of foreign intervention. Consequently, despite its merits at the systemic level, PTT currently lends little useful guidance to policy makers confronting specific potential conflicts. PTT will be more useful to policy makers if theorists supplement PTT's strategic level of analysis with a military studies approach at the operational level.

Cases

Each of the following cases highlights modern conflicts where political-military nuances enabled a weaker power to militarily defeat or outlast a stronger opponent contrary to the predictions of PTT. In each of these cases, the defeated

party possessed vastly superior total national power as measured by both CINC and GDP. In the Russo-Japanese War, superior Japanese power projection and qualitative superiority enabled Tokyo to reshape East Asia's regional power hierarchy in its favor. In the Afghan War, the Taliban survived, owing to America's inability to finish them off inside Pakistan due to competing strategic concerns and later returned to overthrow the weak Afghan government. Finally, the 2022 invasion of Ukraine revealed the Russian military's poor readiness, low morale, and obsolete doctrine rendered them vastly inferior to their smaller Ukrainian neighbor and shattered illusions of Russian great-power status. While the details of each case are unique, trends and themes within them such as the primacy of power projection, influence of conflicting strategic goals, and impact of readiness and morale are common to wars across history and at all levels of the international system. By exploring cases wherein political-military specificities shaped the outcome of conflicts contrary to the expectations of PTT, the authors aim to demonstrate the value to policy makers of supplementing PTT's utility at the grand strategic level with military studies approaches at lower levels of analysis. PTT enjoys many merits, but for the purposes of informing policy it needs to be paired with a military studies approach.

While it may be tempting to disregard these conflicts as limited wars and thus of little interest to great-power politics, in PTT regional hierarchies matter because they alter regional power structures within the broader global hierarchy.[32] Japanese victory over Russia transferred regional dominance to Japan, thus enabling its subsequent revisionist actions against the United States and the UK. Russia's invasion of Ukraine explicitly sought to undermine American regional dominance by eliminating what Moscow believed was a vulnerable American partner. Even the U.S. war in Afghanistan, despite not being a traditional power transition conflict, was still relevant to PTT and worthy of study. This is because of its two-decade long opportunity cost it inflicted on American resources and political attention that could have been better spent containing resurgent revisionist Russia or rising China. These limited wars are critical in international relations, and policy makers and PTT theorists alike should not discount their lessons.

Case 1: The Russo-Japanese War, 1904–1905

The 1904–5 Russo-Japanese War exemplifies the supremacy of case-specific strategic considerations such as power projection and doctrine over PTT's adherence to total power in actual warfare. Russia's CINC score of 0.11 was more than double Japan's 0.05 in 1904, and during the conflict Russia's CINC score *grew*, while victorious Japan's score declined.[33] Likewise, Russia enjoyed a 2.5 to 1 GDP advantage over Japan in 1904.[34] According to the logic of calculating national power through CINC and GDP, Russia should have possessed clear superiority, and yet Moscow was soundly defeated. The keys to that defeat rested in Russia's deficient power projection ability in the East Asian theater, adherence to obsolete military doctrine, and abysmal military morale. These

case-specific constraints rendered total Russian power irrelevant and enabled Japan to execute a regional power transition challenge without reaching parity.

Overview of the Russo-Japanese War

Following the failure of negotiations to demarcate separate Russian and Japanese spheres of influence in Manchuria and Korea in 1903, Tokyo opted to seize its territorial claims by force.[35] The Russo-Japanese war did not result from miscalculation; it was a deliberate gamble by a weaker power to leverage its regional military advantages to force a territorial fait accompli against a stronger adversary. Heading into the war, Tokyo understood that Russia's strategic position was undercut by three factors. First, Russia's overwhelming military might could not be concentrated in theater due to the great distances involved and poor logistical capabilities.[36] Second, Japan enjoyed the benefit of surprise as Russian leadership believed their total power advantages deterred Japanese aggression.[37] Third, the Anglo-Japanese Alliance deterred intervention by third parties.[38] Consequently, Japanese leadership estimated that Russia's regional vulnerability created a window of opportunity to gain a foothold on the continent and assert Japanese great-power status.

Japan initiated hostilities on 8 February 1904, with a surprise attack on the Russian fleet anchored in Port Arthur, followed by a series of unopposed naval landings over the next 10 days. Throughout 1904, Japan enjoyed the freedom to strike when and where it pleased, while Russian commanders possessed limited intelligence regarding Japanese force concentrations or movements, and even less means to inhibit their advance. Russian forces were left immobilized by Japanese naval superiority, isolated by the advancing Japanese army, and crippled by underestimating Japanese capabilities.[39] By May 1904, Japan conquered Korea and much of Russia's ground forces, and all its battleships in theater were trapped in Port Arthur. During the siege of Port Arthur, fighting continued throughout Manchuria, but despite growing Russian force levels, outnumbered Japanese forces consistently repelled Russian counterattacks attempting to relieve Port Arthur. In January 1905, after seven months of grueling siege warfare, and several indecisive naval battles, Russian forces in Port Arthur surrendered and all surviving battleships of the Pacific fleet were scuttled to avoid capture.[40] Although Japan suffered high casualties capturing Port Arthur, seizing the port secured a crucial supply hub for reinforcements and freed Japanese forces to advance farther into Manchuria. Moreover, the loss of Russia's Pacific fleet battleships ended credible Russian naval resistance in theater.

As the war dragged into 1905, Japan reached the limit of its logistical capacity to support a major land war. Japan was running out of ammunition in theater, trained reservists at home, and the Japanese army faced severe difficulties getting supplies forward to combat units.[41] Japanese morale, however remained high, and its soldiers continued to outperform their more numerous Russian adversaries, but Japan now ran a real risk of simply being ground to death in a war of attrition. In February–March 1905, Japan inflicted approx-

imately 89,000 Russian casualties during the Battle of Mukden, but Japanese forces were unable to pursue the retreating Russians and secure a decisive victory.[42] Japan's increasingly fragile logistics simply could not support further exploitation inland. Conversely, Russian forces afflicted by lethargic leadership, poor morale among the rank and file, their own logistical woes, and civil unrest in Russia meant it was unable to regain initiative and drive the Japanese back.[43] Russia's final gambit rested on the prospect that the Russian Baltic Fleet could credibly oppose the Japanese Combined Fleet after trekking halfway around the globe without access to proper port facilities and maintenance en route. Unsurprisingly, the Baltic Fleet was annihilated in detail when intercepted at the Battle of Tsushima Strait in May 1905. With the Russian navy defeated, the mounting costs of the land war, and growing popular unrest, Moscow sued for peace through negotiations.[44] Russia subsequently ceded Korea, Manchuria, and South Sakhalin to Japan. To understand how Japan achieved this seemingly impossible victory, we need to explore the crucial roles of power projection and doctrine in the Russo-Japanese War that made Tokyo's victory possible.

Twin Failures of Russian Naval Power Projection and Force Posture

On paper, Russia enjoyed overwhelming naval superiority over Japan, but in practice Japan enjoyed a regional naval advantage. The Russian Pacific Fleet's battleships were based in the easily blockaded Yellow Sea at Port Arthur, while the bulk of the Pacific Fleet's armored cruisers and torpedo boats were based in Vladivostok, separated by the entire Korean Peninsula.[45] Individually, neither of these flotillas possessed the firepower to face the Japanese Combined Fleet; Russia's battleships were ill-suited to combat light torpedo boats, whereas the cruisers lacked the heavy ordnance to duel with Japanese battleships. To achieve parity, Russia's Pacific Fleet needed to sail its squadrons through Japanese-controlled waters and link up, at the risk of their own annihilation in transit. Japan's surprise torpedo boat attack on Port Arthur on the opening day of the war damaged several Russian capital ships, sharply reducing their odds of successfully evading a Japanese blockade to reach Vladivostok.[46] Advancing Japanese ground forces further jeopardized Russia's naval position by threatening to eject them from Port Arthur into the waiting guns of the Japanese fleet or bring them under constant artillery fire from land.[47] From the moment the war started, Japan enjoyed a dominant position on the high seas, despite its numerical disadvantage in total warships.

Throughout 1904 the Russian Pacific Fleet's battleships remained trapped in Port Arthur and incapable of interdicting the flow of Japanese troops and equipment to the continent. In this way, Russia's inability to contest the high seas also contributed to its difficulties on land. By nature of the region's geography and Russia's dispersed force posture, the Russian Pacific Fleet was compelled to face a stronger enemy who had already dealt a surprise blow on unfavorable terms. The Pacific Fleet was ground down through a series of skirmishes during

Table 1. Comparative Russian and Japanese fleet strengths in East Asia, February 1904

Russian Pacific Fleet		**Japanese Combined Fleet***	
battleships	7	battleships	6
armored cruisers	4	armored cruisers	6
cruisers	7	cruisers	12
destroyers	21	destroyers	22
torpedo boats	22	torpedo boats	28

*Omits Japanese warships assigned to auxiliary or coastal defense duties and unavailable for offensive combat operations.
Source: Yoji Koda, "The Russo-Japanese War: Primary Causes of Japanese Success," Naval War College Review *58, no. 2 (Spring 2005): 10–44.*

Table 2. Comparative Russian and Japanese fleet strengths in East Asia, May 1905

Baltic Fleet		**Japanese Combined Fleet***	
battleships	8	battleships	4
coastal battleships	3	coastal battleships	1
armored cruisers	3	armored cruisers	8
cruisers	6	cruisers	18
destroyers	9	destroyers	21
torpedo/gunboats	0	torpedo/gunboats	34

*Omits Japanese warships assigned to auxiliary or coastal defense duties and unavailable for offensive combat operations.
Source: Piotr Olender, Russo-Japanese Naval War 1905*, vol. 2 (Sandomierz, Poland: Stratus, 2010).*

summer 1904 and the scarcity of naval repair yards in the Far East prevented Russia from mitigating attrition in theater. When the Pacific Fleet's battleship squadron was finally forced to sea by Japanese artillery in August 1904, the squadron failed to evade the blockade and its vessels were interned in neutral harbors or forced to return to Port Arthur and was subsequently scuttled. Likewise, the Vladivostok cruiser squadron was crippled beyond repair while attempting to slip through Korean waters to rendezvous with the battleships.[48] By the time the Russian Baltic Fleet resupplied in Madagascar in January 1905, the Pacific Fleet was already defeated.[49] In the ensuing four months, it took the Baltic Fleet to reach the combat zone, the Japanese enjoyed ample time to repair damage and replace lost vessels.[50] Conversely, the Russian Baltic Fleet was forced to transit halfway around the world, without access to friendly forward bases or repair facilities, with some vessels that were never designed for high-seas service.[51] The Baltic Fleet's numerical superiority in capital ships was undermined by demoralized and undersupplied crews and the fleet was annihilated by Japan's navy at the May 1905 Battle of Tsushima Strait.

Russia's theoretical naval superiority was meaningless in the actual conduct of the Russo-Japanese War. The separation of its fleets by half a globe diminished Russia's naval might, while the country's poor global power projection capabilities and the vulnerable posture of existing forces in theater further exacerbated this weakness. Russia's lack of overseas repair and refueling assets to support the Baltic Fleet's global transit and the unreadiness of its equipment for long-range redeployment rendered it combat ineffective by the point it arrived in theater. Likewise, the Pacific Fleet's disposition of force at the start of the conflict undermined its ability to counterbalance the Japanese Combined Fleet. Japan never faced the combined might of the Russian navy; instead, it faced two weaker Russian fleets and defeated them separately. Consequently, a strategic situation that appeared to assure Russian victory on the macroscale, in fact, favored Japan.

Failure of Russian Land Power Projection and Military Doctrine

Russia's apparent superiority on land proved equally illusory due to poor power projection, archaic military doctrines, and abysmal morale. Russia's poor infrastructure in the Far East prevented it from projecting overwhelming force against Japan and allowed Japan to fight the Russian army on roughly equal terms. Likewise, obsolete military doctrines combined with the low morale and poor training of the army prevented Moscow from gaining qualitative superiority. Consequently, revisionist Japan gained the upper hand on both land and sea and forcefully reshaped the power dynamics of East Asia. While the particulars of every war vary, Russia's defeat in 1905 serves as a stern warning that total power does not guarantee military victory, even against weaker regional revisionists.

Russia's entire land war effort hinged on the single-track Trans-Siberian Railroad, wholly inadequate for the logistical burden of high-intensity warfare. Transit times across the Trans-Siberian rail line averaged 40–50 days, thus planners in Moscow were compelled to plan resupply and reinforcement far in advance of actual events at the front and errors required months to correct.[52] Weather hazards, incomplete rail sections, and chronic derailments further compounded the Trans-Siberian Railroad's logistical difficulties.[53] Consequently, Russia's land power projection capabilities were undermined throughout the conflict, because it could not move men and material into theater as quickly as the Japanese could by sea. During the conflict, Japanese forces generally remained at full strength as fresh replacements arrived from the home islands, while Russian forces hovered around 70 percent of their paper strength due to shortages of replacements in theater.[54] Despite the numerical supremacy of the Russian army in its entirety, Japan enjoyed a significant regional force advantage at the beginning of hostilities.[55] Due to these logistical constraints, Moscow could not overwhelm the Japanese through sheer force of numbers. Moreover, the threat of invasion by European rivals further reduced Russian power pro-

jection capabilities by tying down Russia's finest troops on its western borders.[56] Logistical deficiencies severely undermined Russian land power projection and strategic uncertainty further exacerbated these problems.

Russia enjoyed key technological advantages on land that promised to offset its logistical failures, but these advantages were undermined by obsolete doctrine and poor morale. When the war began, Russia had already begun issuing machine guns to its combat divisions, while Japan was only just beginning to embrace the new weapon.[57] Likewise, modern Model 1900 field guns made up a full one-third of Russian artillery, in contrast to the archaic Type 31 mountain gun used by Japan.[58] Russia further utilized modern entrenchment techniques, barbed wire, and minefields to funnel Japanese forces into the killing fields of its machine guns. From a purely technical perspective, the Russian military was well-equipped to fortify and defend its far-flung Eastern holdings. However, archaic doctrines and abysmal morale undermined qualitative advantages offered by these technological advances.

Russian infantry doctrine continued to embrace nineteenth-century massed volley fire tactics followed by a bayonet charge.[59] While Russian training emphasized archaic practices of unaimed massed fire, their Japanese adversaries embraced modern concepts of individual marksmanship and initiative.[60] Likewise, Russian doctrine did not foresee the possibility of night combat and infiltration, which the Japanese explicitly trained for and exploited to great effect throughout the conflict.[61] Russia's obsolete practices proved disastrously ineffective on war waged around the clock with battlefields filled with trenches, machine guns, and bolt-action rifles. While Russian forces offered effective resistance when fighting from strong defensive fortifications, their obsolete doctrines diminished their ability to retake lost territory. Before the war had even begun, poorer Japan already fielded a better trained military that embraced new doctrines suited to modern warfare, despite its qualitative inferiority in equipment. Technological advantages alone do not ensure a military is prepared for a modern conflict; sometimes the less affluent combatant more accurately predicts and exploits the conditions of future warfare.

Second, obsolete Russian artillery doctrine allowed Japan's technologically inferior artillery to outperform Russian rivals. Russian field artillery doctrine did not use indirect fire to engage targets beyond line of sight, despite their new field guns possessing that capacity. Russian doctrine further failed to anticipate the dangers of hostile indirect fire and called for artillery to be deployed in concentrated groups on high hilltops, without entrenchment or camouflage. Finally, Russian artillery operated autonomously and chronically failed to coordinate with adjoining infantry to protect valuable artillery assets.[62] The Japanese exploited these doctrinal deficiencies by employing their own artillery pieces in dispersed groups, firing from beyond line of sight, coordinated by field telephone wires and forward observers, to destroy Russian artillery with little fear of reprisal.[63] Throughout the war Japan continued to improve combined arms cooperation by ultimately co-deploying artillery in the trenches with the infan-

try while Russia continued to keep these forces separated.[64] Thus, Russia's technological artillery advantage was nullified by doctrinal incompetence against the innovative thinking of their less advanced rival.

Finally, poor morale and unmotivated soldiers exacerbated Russia's doctrinal deficiencies. Russian field commanders consistently noted "complete apathy, almost an indifference toward the war" among rank-and-file soldiers.[65] Likewise, postwar Russian military reformers identified the lack of national sentiment or investment in the conflict's outcome as a major contributing factor to defeat. Military mutinies and civilian riots in Russia's core territories erupted during the conflict in the east, precipitating its transition to constitutional monarchy after the war.[66] In a society where 70 percent of the army's conscripts were impoverished peasants living under constant repression, the average soldier had little incentive to sacrifice themselves for the tzar's interests halfway across the world.[67] Conversely, Japan's better trained soldiers also displayed greater commitment to the conflict. Throughout the war, Japanese units continued to fight without breaking, even as their commanders repeatedly threw them at entrenched fortifications without concern for their survival.[68] Despite suffering significant casualties, Japan reinforced its logistical and doctrinal advantages by maintaining discipline among its rank-and-file soldiers throughout the conflict.

Inadequate power projection precluded Russia from gaining quantitative superiority and forced Moscow to rely on qualitative advantages that it chronically misused. Russia's technological edge in artillery was nullified by deploying these assets in ways that made them easy targets for less advanced Japanese artillery. Likewise, despite acceptable defensive performance, Russian infantry doctrine was poorly suited to retake lost ground. Finally, pervasive poor morale stemming from the tsarist regime's declining legitimacy lent the average Russian soldier little reason to sacrifice themselves to defend far-flung outposts of the empire. Despite impressive power according to the logic of GDP or CINC, Russia could not concentrate its full might in East Asia. Bereft of overwhelming numbers, Moscow badly misused the forces it did possess in theater and received a humiliating defeat at the hands of a regional rival.

Underestimating the Weaker Party—Japan Triumphant

Moscow's defeat in the Russo-Japanese War presents a stark reminder that GDP and CINC reveals little about the likely outcome of a conflict to policy makers. Russia's total power did not prevent weaker Japan from seizing regional dominance due to its superior power projection in theater. Likewise, possessing technological advantages does not guarantee that a military will exploit these advantages in combat. Predicting military outcomes from total power alone courts the same hubris that delivered Russia an ignominious defeat in 1905 and created a new political order in East Asia. American policy makers confronting China today must carefully consider the military aspects of the rivalry to avoid falling victim to a similar gambit.

Case 2: The U.S. War in Afghanistan, 2001–2021

The recent war in Afghanistan further demonstrates the problems of PTT's GDP/CINC-based predictions of conflict outcomes divorced from strategic and broader geopolitical factors. The Afghan War simultaneously reveals the difficulties PTT faces predicting outcomes in asymmetric warfare, coupled with its omission of competing priorities that impede the use of national power. Compared to the Russo-Japanese case, the conventional power imbalance between the belligerents in Afghanistan was significantly wider. Table 3 notes the CINC scores and gross domestic product measures before the onset of hostilities between the U.S. and Taliban governments up to the most recently available data. The government of Pakistan is further included given the prominent role that elements of its security apparatus played in supporting the Taliban.

At the beginning of the war, in late 2001, the United States maintained a nearly 35 to 1 advantage in power over the Taliban as measured by CINC score. This ratio narrows to a factor of 9 to 1, including the full government of Pakistan on the Taliban's side of the ledger.[69] With GDP alone, the imbalance skyrockets to a factor of more than 1,000 to 1, down to 86 to 1, if including Pakistan. This phase of the conflict fully aligns with PTT's predictions considering the huge power advantage of the United States and its allies over the Taliban. Such calculations also assume that all the energies of the government of Pakistan were devoted to the effort, which is clearly not the case. Consequently, the practical disparities in capability should be even greater than these portrayals.

The fall of Kabul to the Taliban in 2021, however, contradicts the theory's expectations. As a nonstate actor there are no concrete datasets that provide a hard power measure of the Taliban's capabilities; however, comparing the U.S. and Afghan governments' combined scores to that of the government of Pakistan yields preponderance factors of 56 to 1 for GDP and a CINC power imbalance by a factor of 9 to 1. Moreover, the U.S. and Afghan governments operated with the support of North Atlantic Treaty Organization (NATO) partners in the International Security Assistance Force, further bolstering the power brought to bear against the Taliban insurgency.

Security assistance data demonstrates that the U.S. government provided nearly $73 billion in military aid to Afghanistan between 2001 and 2020, which was close to 20 times the government of Afghanistan's military expenditures.[70] These monies were directed to paying the salaries of Afghan security personnel and included extensive training and equipping efforts for the Afghan National Police and the Afghan National Army. Despite these significant investments, there was no return in terms of security performance. Former Ambassador Ryan C. Crocker argued that Afghan personnel were "useless as a security force because they are corrupt down to the patrol level."[71] Such was the extent of this corruption that as many as 18 percent of security personnel on record were "ghost" soldiers who existed for the purposes of commanders skimming their paychecks.[72] U.S. forces and logistical support was required to keep Afghan security forces operating. However, the Taliban continued to make gains.

The Afghan government could count on more than 250,000 of its own security personnel, with some scholars estimating that the daily on-call number was 180,000.[73] By comparison, estimates of the Taliban hover around 60,000 full-time fighters.[74] These figures do not include the thousands of Western troops or contractors in theater at any given point during the conflict and the extensive support they brought in terms of artillery and airpower. Notably, during the course of the war most of the Afghan forces operated with qualitatively superior weaponry and in a largely defensive role while Taliban fighters maintained an offensive orientation without access to air support or armored vehicles. Per raw troop count and the disposition of forces the traditional military logic of a 3:1 advantage, albeit at the theater level, was inverted in favor of the defense. Yet, the Afghan National Army decisively lost the war and failed to defend its capital city.

How did such a lopsided case in terms of power distribution yield a decisive victory for the weaker side? The key to U.S./Afghan defeat in 2021 rests in the broader geopolitical situation in South-Central Asia. Despite an overwhelming advantage in on-call firepower in favor of the U.S.-aligned Afghan government, the Taliban benefited from a safe haven and sponsorship by elements within the Pakistani government. The Taliban enjoyed relative freedom to reorganize and recruit inside Pakistan, protected from most elements of U.S. power by Pakistani sovereignty. America's understandable unwillingness to expand the war into the Taliban's safe havens inside Pakistan and risk throwing the entire region into chaos by destabilizing a nuclear-armed power strongly contributed to the Taliban's victory. The Taliban could afford to fight for years, even sustain heavy losses, because losses could be recovered inside Pakistan, aided by direct support from elements of the Pakistani government. Thus, America was never fully able to defeat the Taliban because it never gained full control over the Taliban's bases of operation or neutralized the Pakistani government's role in resupplying the Taliban. Geopolitical realities precluded military victory, and the United States was never capable of fully defeating the Taliban before handing control over to the fatally flawed and doomed to fail Afghan government.

The Taliban's Long Road (Back) to Kabul

Eventual Taliban victory in August 2021 was aided by regional political trends established before the war even began. As part of their efforts to achieve strategic depth against India, Pakistani leadership long sought to establish a friendly regime on their western flank in Afghanistan, while directly supporting Kashmiri militant groups against India.[75] For the former, Pakistan facilitated aid to the Taliban, while in the latter it enlisted the assistance of international jihadists, including Osama bin Laden. Pakistani ties to the Taliban also extended beyond simple matters of state policy as Inter-Services Intelligence (ISI) personnel operating in Afghanistan predominantly shared both Pashtun tribal identity and fundamentalist Islamist ideology with the Taliban they assisted.

During the Afghan warlord period in the early- to mid-1990s, Pakistan

Table 3. Comparative U.S., Afghan, and Pakistani CINC and GDP (in millions constant 2015 USD) values

	CINC (2000)	CINC (2016)	GDP (2000)	GDP (2020)
United States	0.1426877	0.1330576	$13,754,300	$19,247,056
Afghanistan	0.0040984	0.0028154	$11,900	$20,621
Pakistan	0.0132468	0.0151497	$146,487	$320,098

Sources: CINC data per "National Material Capabilities v6.0 Dataset"; GDP data for Afghanistan in 2000 extrapolated from Maddison project dataset and converted to constant 2015 USD; Jutta Bolt and Jan Luiten van Zanden, "Maddison Style Estimates of the Evolution of the World Economy. A New 2020 Update" (Maddison Project Working Paper WP-15, October 2020); and GDP drawn from World Bank, "GDP (constant 2015 US$)—United States, Pakistan, Afghanistan," World Bank Group, accessed 31 October 2022.

extended its influence inside Afghanistan by harnessing cross-border Pashtuns tribal connections. First, Pakistan attempted to bring Afghanistan's Pashtun population under their influence through pro-Pakistani mujahideen commander Gulbuddin Hekmatyar and then, more successfully, by supporting the Kandahar-based warlords who subsequently formed the Taliban.[76] Extending Pakistani influence, although not control, over the Afghan Taliban proved possible since many Taliban fighters and leaders alike were strongly shaped by and connected to Pakistan. As a prominent expert on the Taliban, Ahmed Rashid observed:

> The Taliban were born in Pakistani refugee camps, educated in Pakistani madrassas and learnt their fighting skills from Mujaheddin parties based in Pakistan. Their families carried Pakistani identity cards. The Taliban's deep connections to Pakistani state institutions, political parties, Islamic groups, the madrassa network, the drugs mafia and business and transport groups came at a time when Pakistan's power structure was unravelling and fragmented.[77]

ISI support for the Taliban and al-Qaeda persisted following the 11 September 2001 terrorist attacks. After the United States and the Northern Alliance defeated the Taliban at Kabul in November 2001, Taliban and al-Qaeda combatants coalesced in their last major bastion in the northern city of Kunduz. Among them were Pakistani military advisors and intelligence officials embedded with the movement. With Northern Alliance forces closing in and fearing the embarrassment of its agents being captured, Pakistani president Pervez Musharraf orchestrated a quid pro quo with President George W. Bush seeking to evacuate his military advisors in exchange for helping the United States gain access to the region for military operations.[78] The evacuation commenced in a series of secret Pakistani flights dubbed the "Kunduz Airlift." Whether through

these evacuations or in the successive weeks via the porous border, al-Qaeda's senior leadership also escaped capture. That Osama bin Laden was ultimately found in a compound within a mile of the Pakistan Military Academy suggests some degree of complicity within elements of Pakistan's security services.

Following the success of the Northern Alliance and U.S. forces in deposing the Taliban, the security situation allowed officials to create a constitution and hold elections in 2004. After these milestones, the U.S. government in 2005 elevated Afghanistan to the level of "strategic partner."[79] For Pakistani strategists, the potential rise of a pro-Indian government in Afghanistan backed by the United States directly threatened their aspirations for strategic depth. Reestablishing a pro-Pakistani government inside Afghanistan became a matter of Pakistani national interest, and the pro-Islamabad Pashtun Taliban continued to present ideal proxies for the task.

The security situation in Afghanistan quickly destabilized in subsequent years as Pashtun fighters from Pakistan flowed across the border to support a Taliban insurgency. The central government based in Kabul faced the difficult task of creating a strong centralized government in a society that is ethnically fractured, tribal, and dramatically underdeveloped. Demographically the largest ethnic bloc in the country are Pashtuns (42 percent) who are joined by various other ethnic minorities.[80] Persistent grievances between concentrations of minorities in the north and the majority Pashtun populations in the south and east continued to sow mistrust and conflict. Although the nation's Human Development Index climbed in the years following the Taliban's ouster in 2001, in 2021 Afghanistan still ranked 180th out of 191 countries, even before the government's collapse.[81] Moreover, the country ranked 174th out of 180 countries on Transparency International's Corruption Index in the last year where data was available.[82] The Taliban exploited the national government's rampant corruption and weakness and enjoyed a steady stream of fighters from across the border supported by camps in Pakistan's then Federally Administered Tribal Areas.

ISI support for the Taliban remained significant despite Pakistan's alleged support for the U.S.-led Global War on Terrorism. ISI support efforts coincided with major Taliban offensives such as in 2006 and the extent of assistance extended beyond sanctuary across the border to include actively providing training and furnishing equipment, fuel, and ammunition.[83] ISI training provided to the Taliban included instruction in creating suicide bombs and improvised explosive devices, both crucial to Taliban combat operations. Some estimates note that as many as 80 percent of Taliban fighters in some sectors were trained in Pakistan.[84] Much of this training and recruitment took place in madrassas in the border region, which are ideologically aligned with the Taliban.

Despite shared interests and ideological similarities, the Taliban were not completely aligned with the Pakistani government. In fall 2007, the Pakistani Taliban launched an offensive against the Pakistani government to seize the city of Swat in 2006. A Pakistani counteroffensive subsequently recaptured the

Swat valley, but soon the Pakistani Taliban returned, forcing the government into a cease-fire in February 2009. By spring 2009 the government launched a renewed offensive to reclaim Swat and pursue the group's leadership specifically. However, the offensive did not seek to assert Pakistani control over the entirety of the border region or seek to undermine the broader Taliban movement. As the Swat valley confrontation indicates, while the ISI consistently aided the Taliban, they never controlled them. The Taliban was simultaneously a partner and challenger to the Pakistani government, a weapon the Pakistani government unleashed against Afghanistan at its own risk, but never a puppet. As part of the equilibrium, Islamabad never attempted to drive the Taliban from Pakistani territory entirely.

Pakistan's role as the Taliban's primary state supporter and as a Taliban base of operations was actively understood by U.S. leadership at the time. The theater of operations was even routinely referenced as AfPak in recognition of Pakistan's persistent role during the conflict, both as a Taliban base of operations and active supporter. The U.S. special representative to Afghanistan and Pakistan at the time, Richard Holbrooke, noted the critical role that the territory of Pakistan played when he said that "it is on the eastern side of this ill-defined border that the international terrorist movement is located."[85] U.S. leaders were fully aware of Pakistan's key role in the conflict, but larger concerns precluded serious action against/inside Pakistan. At an operational level, the United States relied on Pakistan as the most direct route of resupply into Afghanistan. Meanwhile, at a strategic level, fear of creating a worse crisis by destabilizing Pakistan loomed large over American decision making.

The border region haven and continued assistance by the ISI worsened the security efforts to stabilize Afghanistan. The year 2010 marked the height of the insurgency, following the announcement of a troop surge in 2009 by incoming President Barack H. Obama. This effort sought to "disrupt, dismantle, and defeat al Qaeda and its safe havens in Pakistan and Afghanistan, and to prevent their return to either country in the future."[86] The surge in troops was matched by a surge in American foreign aid, which doubled from approximately $1 billion in 2009 to an average of $2 billion in 2010 and 2011. After security gains in 2013, the U.S. government officially handed over security to the Afghan government, and in 2014 President Obama announced a schedule of the U.S. withdrawal. However, the Taliban's bases in Pakistan remained active and thus the cornerstone of the entire Afghan insurgency survived the surge. In the successive seven years, the Taliban gradually reasserted control of regional provinces in the south and east, which later became the basis for their campaign of national conquest.

In the final stages of U.S. involvement, in 2017 President Donald J. Trump adopted a policy of expanding military operations and delineating more decision making to military officers in theater. Trump simultaneously directed diplomats to negotiate with the Taliban while remaining security operations confronted growing Islamic State elements in Afghanistan. The following year

Trump restricted aid to Pakistan and announced a renewed offensive against the Taliban. By 2020, U.S. and Taliban representatives signed a peace deal shifting diplomacy to the Taliban and Afghan government and the United States announced a major drawdown to 2,500 personnel remaining in country. In April 2021, incoming President Joseph R. Biden announced a full withdrawal of U.S. forces by 11 September. On 6 August, the Taliban captured their first regional capital at Zaranj, and within a little more than a week the Afghan national government collapsed and the national capital at Kabul fell without significant fighting on 15 August. With the protective shield of U.S. forces removed, even after years of training and material support the Afghan government proved utterly unable to defend itself.

Competing Strategic Goals Preclude U.S. Victory

In the end, the Afghan government proved unable to defend itself, but the foundation of U.S./Afghan defeat stemmed from the failure to neutralize Pakistan's role as a base for and supplier of the Taliban. If the Taliban never enjoyed the luxury of regrouping and recruiting inside the relative safety of Pakistan, the weak Afghan government may have never faced an opponent strong enough to overcome its limited power. Larger geopolitical considerations precluded America from crippling the Taliban insurgency by eliminating its Pakistani bases of operation. On the contrary, the United States provided Pakistan more than $91 billion in foreign assistance since 2001, even as Pakistan's security services supported the Taliban.[87] Pakistan was simultaneously an American adversary and partner.

U.S. policy makers used foreign assistance to Pakistan at an operational level to maintain a key supply route through Pakistani territory and gain easier access to the eastern portions of the Afghan theater. Intervening inside Pakistan with the level of U.S./allied force would have required control of both sides of the border and necessitated toppling the Islamabad regime and turning the crucial U.S. logistics route through the country into an active insurgency zone. In short, from an operational perspective, intervening inside Pakistan was prohibitively dangerous, even ignoring the more pressing strategic implications of invading a nuclear power. In a strategic sense, propping up the Islamabad regime was preferable to the risk of allowing the regime to fall and risk its nuclear arms falling into the hands of anti-American terror groups. Islamabad's support for the Taliban was dangerous and prevented U.S. victory in the Afghan War, but Pakistani regime survival was still preferable to the hazards posed by its potential collapse. The United States could either win the Afghan War but risk creating a larger regional catastrophe with global ramifications, or it could support the same state that kept the Taliban insurgency alive in the hopes of averting a wider crisis.

Predicting outcomes based on CINC or GDP is futile in conflicts such as the Afghan War, because in the end American defeat stemmed from competing strategic concerns that precluded the military steps necessary to win. The Unit-

ed States could contain the Taliban by beating their advances back annually using the extraordinary military power available to America and its allies, but it could never destroy the Taliban completely. Conversely, the infinitely weaker Taliban merely needed to prevent the Afghan national government from creating a stable civil society, replace its losses behind the shield of Pakistan, and then topple the weak Afghan government as soon as the United States ceased combat operations. The Afghan government's rapid defeat was the epilogue of U.S. strategy at odds with itself, and Washington wasted two decades in a holding pattern until the conflict was abandoned and allowed to run its course. Victory or defeat in Afghanistan had little to do with actual power and everything to do with larger strategic constraints on the use of power.

War Cannot Escape Politics—The Taliban Victory

The United States spent two decades pursuing illusory victory in Afghanistan, while any chance for lasting victory remained firmly out of reach behind Pakistani borders. Total measures of power favored by PTT cannot capture case-specific limitations on the use force wrought by competing strategic or political priorities. Simply because a state can project overwhelming power as the United States did inside Afghanistan matters little if it cannot fully defeat its opponent due to competing political necessities, such as the American unwillingness to expand the war inside Pakistan. Instead, the United States frittered away manpower, resources, and its national image chasing victory that could never come. When Afghanistan finally collapsed in 2021, the shock to American power reverberated around the globe. Revisionist adversaries seized on apparent American weakness to test the edges of what they perceived to be Washington's declining imperium. Just 20 days after Kabul's fall, China dramatically increased the number of incursions into Taiwan's air defense identification zone, and these increased sortie rates expanded dramatically in the subsequent month.[88] A little more than six months after Kabul's fall, Russian president Vladimir Putin launched a full-scale invasion of Ukraine in direct opposition to the Pax Americana that has dominated European politics since the end of the Cold War. After two decades, thousands killed, and more than $2 trillion spent during the conflict, the Afghan War clearly demonstrated that no amount of material preponderance can guarantee victory if overarching political conditions prevent operational level military success.[89]

Case 3: The War in Ukraine, 2022–2023

The ongoing Russian invasion of Ukraine illustrates the pitfalls of obsolete military doctrine, inadequate modernization, and poor training or morale in the face of a determined combatant strengthened by foreign material aid. GDP and CINC have failed spectacularly to predict the course of the war thus far. Ukraine stands at approximately 20 percent of Russia's CINC score, 7 percent of its GDP, and despite apparently overwhelming odds, continues to push Russia back.[90] According to the logic of measuring power by CINC and GDP

Russian victory *should* have been swift and decisive and yet the war leans in Kyiv's favor as Ukrainian forces have delivered stunning defeats to the Russian military. Understanding the war in Ukraine requires engaging with concepts that cannot be readily captured by sweeping assessments of broad national capabilities. Simply stated, Ukrainian forces outfight their Russian opponents due to deeply engrained Russian doctrinal and procurement deficiencies. For the broader field of PTT, the war in Ukraine proves finite military details matter in predicting conflict outcomes. Even seemingly minute details such as equipment modernization, doctrine, and training can profoundly reshape a conflict in favor of a seemingly hopelessly outmatched state.

Opening Moves, 24 February–8 April

On 24 February 2022, after months of preparation, Russia invaded Ukraine intent on toppling the Ukrainian government. Instead of a swift blitzkrieg, Moscow found itself trapped in a quagmire with its forces overextended and vulnerable. Russian forces initially attempted to drive through Ukrainian positions guided by the dubious assumption that Ukrainians would not resist. Entire Russian units were annihilated with little resistance as they wandered blindly into Ukrainian defenses without support or preparation.[91] Russian units that did fight frequently advanced deep into Ukrainian territory without protecting their flanks, thereby exposing themselves to encirclement and their logistics to ambush.[92] While attempting advances on four primary axes (Kyiv, Sumy/Kharkiv, Donbas, and Kherson), Russia demonstrated an inability to support all four lines of advance effectively.[93] Crucially, most Russian axes of advance lacked follow-on forces needed to secure lines of supply and neutralize defenders bypassed in the initial breakthrough.[94] Deprived of fuel, necessary supplies, or immediate reinforcements, Russian spearhead units lost momentum and operational initiative passed to the defenders in northern Ukraine. Meanwhile, Russia's air forces remained ineffective both due to their frequent absence and failure to secure air supremacy—necessary to prevent movement of Ukrainian forces reacting to Russian breakthroughs. Where Russian forces encountered strong resistance near Kyiv, Kharkiv, Chernihiv, and along the Donbas line of contact, they quickly defaulted to costly direct assaults with little maneuver or finesse in efforts to dislodge the defenders through sheer weight of numbers. While Russian forces made incremental gains in the Donbas and overran Ukrainian territory south of the Dnieper, their main efforts against Kyiv and Kharkiv ended in humiliating defeat that saw Russian forces withdraw from much of northern Ukraine by 8 April.[95] Even Russia's deepest breakthrough into southern Ukraine was soon halted and forced back on the defensive.[96] Despite possessing nearly every material advantage on a spreadsheet, Russia utterly failed to capitalize on those advantages in real combat.

Donbas Offensive, 18 April–25 June

After a brief period of reorganization and redeployment, Russia renewed its

offensive, intent on eliminating Ukrainian forces in the Donbas region. Unlike the opening phase, Russia concentrated its forces on a single-front advance, supplemented by small operations elsewhere. While Russia made better use of its substantial artillery advantage during the Donbas phase of the war, it still replicated the same pattern of attempted maneuver, failure, and default to costly frontal assault. Russia launched the Donbas offensive with a classic pincer movement aimed at shattering the defenders' flanks near Izium and Popasna to encircle large portions of Ukraine's most experienced combat units near Sievierodonetsk.[97]

Despite some initial success, Russia's pincer lost momentum and degenerated into a series of bloody frontal assaults culminating in the capture of Sievierodonetsk on 25 June.[98] Ukrainian forces evaded encirclement, retreated in good order, and continued to contain the Russian advance. Despite gaining a local territorial victory, Russia failed to either unhinge the Ukrainian defensive line and achieve broad territorial gains or inflict the kind of catastrophic casualties needed to irreparably damage Ukraine's military capabilities. Meanwhile, Ukrainian counterattacks regained ground near Kharkiv and held Russian forces in check along the southern front toward Kherson.[99] The failed Donbas phase of the conflict served as a bloody interlude between the humiliating Russian failures of the initial invasion and Ukraine's counteroffensive.

Ukrainian Counteroffensives, 6 September–11 November

After grinding Russian forces down for six months, Ukraine gained operational initiative and launched a wildly successful counteroffensive in Kharkiv and Luhansk Oblasts starting on 6 September. Within a week, Ukrainian forces liberated the logistical hubs of Izium and Kupiansk, severing Russia's railway lifelines into northeastern Ukraine.[100] Likewise, Ukrainian forces regained near complete control of Kharkiv Oblast to the Oskil River and ejected Russian forces across the northern border. Within a month Ukraine conducted a second encirclement of Russian forces near Lyman, pushed into Luhansk Oblast, and threatened to unhinge Russia's northern flank above Sievierodonetsk.[101] As fighting for Lyman subsided, Ukraine launched a second counteroffensive in the south, creating another localized rout, and ultimately culminating in the liberation of Kherson, the only major city captured by Russia, on 11 November. In two months, Ukraine regained more territory than Russia conquered during the entire summer Donbas offensive and dealt deep material and personnel blows to Russian forces in the field. Throughout both counteroffensives, Russian forces continued to fight poorly, rout frequently, and prove unable to wage mobile warfare. Likewise, Russian modernization and equipment readiness continued to backslide as modern combat vehicles were replaced by older models from reserve stocks.[102]

In response to the twin shocks of the Kharkiv and Kherson counteroffensives, Russia declared partial mobilization and began forced conscription across

the country to replace losses. However, little attention was given to training or equipping mobilized forces, and many fresh soldiers were deployed to combat without even receiving basic training.[103] Attacks of recruitment facilities escalated throughout Russia, soon joined by reports of fratricide inside mobilized units, and more than a million Russians fled the country following mobilization.[104] Partial mobilization provided the Kremlin manpower to bolster numbers at the front, but the lack of training provided to mobilized personnel combined increased reliance on obsolete equipment, perpetuating the qualitative decline of Russia's armed forces. The same key issues of modernization, readiness, and morale that facilitated Russia's shocking failures throughout its invasion of Ukraine have intensified. If the war in Ukraine remains a conventional conflict, Russia will fail to achieve its objectives of regime change and Ukraine will continue to liberate lost territory.

Inadequate Modernization and Poor Material Readiness

Although Russia's CINC and GDP overmatch versus Ukraine seems formidable on paper, inadequate modernization of equipment and poor material readiness undermined Russia's military capabilities. Russia continued to modernize its arms since the end of the Cold War, however, production rates fall far short of demand. Analysis of Russian combat losses reveals that many maneuver units operate obsolete equipment.

Of Russia's most modern vehicles and aircraft, only 67 T-90M main battle tanks, 9 BMP-T tank support vehicles, 133 Kamov KA-52 attack helicopters, and 97 Sukhoi SU-35 4.5 generation fighters were in service at the war's onset.[105] Due to years of delays, Russia's next generation of armored fighting vehicles have not entered active service.[106] Whatever value Russia's modern hardware offers is of limited utility in such small numbers. In truth, Russia's land forces have received low priority in modernization budgets since 2020, despite their paramount importance to Moscow's revisionist ambitions. As an example, 26 percent of the 2020 State Armament Program's funding was directed to Russia's vestigial navy, compared to 14 percent for its gargantuan army.[107] New vehicles have been developed primarily for export, with little regard to improving the capabilities of the Russian army itself. Moreover, Russian military industry remained dependent on technology imported from the West to produce its most modern equipment right up to the day of the invasion.[108] Sanctions imposed after the invasion of Ukraine have subsequently impeded production of Russia's newest war machines, forcing the country to adapt to resume production or produce older, less sophisticated weapons instead. The months of production lost have further hindered efforts to replace losses, let alone modernize its forces. Moscow's military-industrial complex is well-suited for internal security and developing new arms exports, but it lacks reserves of modern equipment necessary to wage sustained high-intensity warfare.

Shortages of modern equipment have forced Russia to rely heavily on Soviet-

era arms, but due to a combination of poor maintenance and corruption, actual equipment stocks have fallen far short of on-paper estimates. Many Russian reserve vehicles lack critical equipment, such as engines.[109] In late March, the Russian military remobilized long-term vehicle reserves in Boguchar, near the Ukrainian border, but found 40 percent of equipment stored there was inoperable.[110] Within days after the invasion, antiquated Soviet-era vehicles such as the T-72A (1970), T-72B (1984), and BMP-1P (1979) began to enter combat. Many of these Soviet relics were captured in incredibly poor states of repair, having been visibly neglected for years, further degrading their combat capabilities. More importantly, many of these obsolete vehicles lack crucial add-on explosive reactive armor that they are *supposed* to have, to protect against modern antitank weapons. Moscow's reliance on Soviet-era equipment did not abate as the war dragged on; in June, Russia began deploying 50-year-old T-62s to Ukraine in response to mounting losses among more advanced tanks.[111] With Russia suffering more than1,600 visually confirmed tank losses alone, Russian forces fall further behind in the modernization of equipment.[112] Russia has failed or is unable to mitigate deficiencies with equipment such as explosive reactive armor.[113] Consequently, Moscow is trapped in a downward spiral of equipment quality due to limited industrial incapacity versus the large volumes of force it must support in combat.

Failure of VKS

The inability of Russian Aerospace Forces, the *Vozdushno-kosmicheskiye sily* (VKS), to gain air superiority over Ukraine stands as one of the war's colossal failures. The VKS's impotence stems from a combination of material deficiencies and an obsolete air warfare doctrine that views air forces as auxiliary support for ground units. The VKS lacks both an effective doctrine to utilize aircraft on a modern battlefield and the material reserves to wage high-intensity state war.

The core of the VKS's failure in Ukraine stems from a World War II-era doctrine rooted in the idea that air forces should be "flying artillery," intended to support ground forces, instead of a fully developed warfighting tool.[114] This flying artillery doctrine hobbles Russian military operations because it does not envision the use of air forces independent of ground operations or conduct air campaigns. Properly executed air campaigns aim to lock down airspace, decapitate command control, and interdict enemy ground movements.[115] The idea of maintaining air control, relentlessly striking hostile air defense systems, and expending thousands of tons of ordnance to render airfields useless is alien to Russian doctrine. Instead, the VKS focuses on providing flying artillery support to ground forces, spotting for ground-based artillery, and terror bombing against civilian populations.[116] Russia's employment of air power during the war in Ukraine reflects these ideas. The VKS embarked on a few indecisive days of strikes against Ukrainian air defenses and airfields at the start of the invasion before reverting to its traditional role in tactical air support, interspersed with terror bombings using valuable precision-guided munitions.[117] Even more tell-

ing, as Russia commits soldiers and vehicles into Ukraine by the thousands, the VKS typically deploys one to four aircraft per strike package with little coordination between strikes. The VKS appears incapable of coordinating large-scale airstrikes or air operations more sophisticated than localized air support and has never attempted to establish air superiority over any part of Ukraine.[118] Russia has squandered its overwhelming aerial advantage on paper through flawed doctrine exacerbated by material deficiencies.

The VKS suffers from equipment modernization pitfalls detailed earlier, specifically limited availability of its newest aircraft. However, the failure to mass produce modern aircraft *should* have mattered less for the VKS considering Ukraine's reliance on Soviet-era air and antiaircraft systems when the war began.[119] The VKS's most serious material pitfalls flow from insufficient reserves of precision-guided munitions. The VKS rapidly depleted its stockpiles of these munitions against infrastructure and civilian targets in a terror bombing campaign, compatible with Russian doctrine but divorced from military needs on the ground.[120] Doing so forces fighter-bombers to resort to less accurate unguided munitions, delivered at lower altitudes, placing aircraft in additional danger from low-altitude air defense systems.[121] Consequently, Russia's air attacks have become less effective as the war progresses while Ukraine has enhanced its air-defense capabilities with more advanced Western equipment.

Beyond shortages of precision-guided munitions, Russia failed to modernize its drone arsenal to supplement traditional air power. While Ukraine responded to similar difficulties of using traditional air power due to Russian antiaircraft systems through increased reliance on drones, Russian drones have not yet played a significant role in the conflict.[122] Ukraine uses a sizable arsenal of Western drones to strike supply depots, antiaircraft systems, and facilities beyond reach of traditional air power. Conversely, Russian drone development and production was sidelined in favor of more vulnerable attack helicopters.[123] Ukraine has leveraged the advantages of drones to great effect whereas Russia struggles to catch up due to its frail electronics industry and reliance on imported Western components.[124]

After the first days of the war, the VKS failed to suppress or destroy Ukrainian medium- and high-altitude air defenses. This pressured VKS sorties to remain close to the ground where they are vulnerable to low-altitude air defenses. This failure was matched by a parallel failure to eliminate the Ukrainian air force in the opening days of the war. Russia attacked Ukrainian airfields at the beginning of the invasion but overall failed to cripple the Ukrainian air force. Ukrainian aircraft face the same air-defense threats as their Russian adversaries, but the Russians still failed to remove them entirely. Moreover, the Ukrainian air force actively trains its pilots for low-altitude operations, and Russian pilots found themselves forced into a low-altitude battlefield for which they were inadequately trained.[125] The Ukrainian air force remains a threat-in-being, helping deter Russian air strikes against crucial interior lines of communication in western/central Ukraine and providing occasional support to

Ukrainian ground forces. Without air superiority, Russia cannot meaningfully interdict the movement of Ukrainian ground forces from the skies, lending Kyiv freedom to move its forces between hot spots.

Deficient Training and Poor Morale

Much like its Russo-Japanese War ancestor, the modern Russian military suffers from inadequate training levels and low readiness rates combined with competing security concerns outside of theater, as well as poor morale. Russian forces are not well trained for combat. Relatively few units are actually available for action in Ukraine, and poor morale saps the capabilities of the units that are deployed.

The Russian military suffers from low training levels. Roughly 30 percent of the Russian military, including its elite paratrooper units, consists of conscripts and even the vaunted Spetsnaz commandos still employ some conscripts.[126] While conscripts *can* fight effectively when properly motivated by crises such as foreign invasion, they present a liability to armies that already suffer from poor morale or lack a professional core to supplement. Many of Russia's personnel are one-year contract soldiers or conscripts.[127] Half of Russia's military personnel are contract soldiers, of which 30 percent are professionals who have completed more than one contract, and the remaining 20 percent are first-year contract soldiers.[128] Consequently, between first-year contracts and conscripts, a staggering 50 percent of the Russian military consisted of soldiers with less than a year of military experience at the beginning of the invasion. Russian forces admit this time frame is insufficient to train soldiers for combat.[129] By contrast, the standard enlistment period for American or British soldiers is eight and four years, respectively. Russian units in the field unsurprisingly display signs of low morale and poor discipline, including looting, atrocities, failure to execute orders, and occasional mutinies. Stopgap measures like mobilization and the indefinite extension of military contracts or forcible remobilization of reservists and conscripts threatens to undermine morale further. As casualties mount, Russia has transferred internal security forces out of Chechnya and garrison units from its foreign military bases in Tajikistan and occupied Georgia to replace losses in Ukraine. Likewise, it levies conscripts from occupied Ukraine itself under the banner of the Donetsk and Luhansk People's Republics (DNR/LNR puppet republics), to provide additional cannon fodder. Finally, Moscow extensively utilizes its private military contractors to stiffen the resolve of its proxy forces.[130] As Russia continues to deploy even lower-grade forces, it will further debase the quality of its troops.

While Russia must rely on quantity in the conflict, low readiness rates and competing security concerns undermine its ability to overwhelm Ukraine with sheer volume of force. Before Russia launched its invasion of Ukraine, it was already 150,000 personnel short of its nearly million-person target.[131] Standard prewar practice in the Russian army required brigades to maintain a single battalion battlegroup in a combat-ready status, less than 30 percent of each brigade's total force.[132] The rest of each brigade must be hastily assembled in a crisis

from conscripts and contract soldiers. Hasty mobilization leaves units little time to learn to operate as a coherent unit before entering combat.[133] Practical experience in Ukraine reveals many Russian units routinely fail to coordinate internally, let alone with other units, and fight with little tactical finesse.[134] Russia's need to garrison its extensive borders and requirements of internal security in southern/eastern Russia further diminishes the amount of force it can realistically deploy to Ukraine. Partial, or even general mobilization, will not resolve Russia's manpower deficiencies.

In truth, Russia has little choice but to deploy conscripts, mercenaries, and other expendable cannon fodder en masse. In 2019, the Russian army maintained a mere 4,000–5,000 reservists by Western standards.[135] On paper, it could reactivate the two million former conscript and contract soldiers available in deep reserve to compensate for losses, but remobilizing veterans proves problematic. First, only 10 percent of former soldiers receive any refresher training after completing their initial service. Second, the Russian Ministry of Defence admits that it does not effectively track ex-soldiers, frustrating remobilization efforts.[136] Russia's recent "partial" mobilization amounted to random conscription regardless of prior military experience to provide untrained bodies for the war in Ukraine.[137] In 2021, the Russian Army trialed a new reserve program aimed at creating a three-year contract active reserve, but the effort fell far short of stated goals and did not bolster Russian reserves.[138]

Russia was prepared for defense and deterrence, not initiating the largest war in Europe since 1945.[139] Russia's professional reserve deficiencies were further exacerbated by disproportionately high casualties among elite units during the initial invasion and heavy casualties among the officer corps and long-serving contract soldiers.[140] Russian forces may grow more numerous, but they will not grow more competent as the war progresses. Relying on poorly trained, understaffed, and low morale forces undermines Russian prospects for victory against Ukraine. Even if Russia somehow resolves its modernization deficiencies and the VKS embraces a modern air warfare doctrine, the poor quality of Russian forces will continue to cripple performance.

A Hollow Bear—When Power Is an Illusion

The ongoing course of the Ukrainian war defies the mechanistic logic of PTT, but it is unsurprising when viewed in terms of political-military specificities. Russia enjoys overwhelming material superiority on paper, but its actual power against Ukraine is limited in ways PTT's preferred power metrics of GDP and CINC cannot capture. Inadequate modernization forces Russia to rely on obsolete equipment against an opponent with access to vast quantities of modern Western equipment. Due to material shortages and archaic doctrine, the VKS has proven itself incapable of conducting a sustained air campaign against Ukraine, lending Kyiv freedom to maneuver largely unharmed behind the front lines. Finally, deficient training, readiness, and morale degrades Russian effectiveness in combat, in the face of a determined and skillful opponent. From

the grand-strategic perspective where PTT prospers, these fine military details may seem like minutiae, but they are decisive in actual military campaigns. If policy makers operationalize PTT based solely on the theory's measures of total national power independent of case-specific military considerations, opportunities for victory against revisionist rivals, as in Ukraine against Russia, will be overlooked and lost. Russia's own failure to appreciate the importance of military details has mired it in a war it cannot win, left its reputation as a great power in tatters, and strengthened the U.S.-led dominant order in Europe. Wars are not fought on spreadsheets. Details matter. Moscow failed to see this, and PTT scholarship must avoid the same pitfall as it seeks to guide policy.

Operationalizing PTT in Defense of Taiwan

Power transition theory can aid policy makers by serving as a threat radar at the grand strategic level, while leaving predicting theater strategic and operational outcomes of specific conflicts to more appropriate military studies methods. PTT helps policy makers understand the wider grand strategic stage, while military studies provides essential case-specific understanding of military and political specificities to predict likely conflict outcomes. In light of this concept, the authors will use the case of China's regional power transition challenge against the United States, with Taiwan as its likely first target as an example of how to better operationalize PTT for Western policy makers.

Shaping the Conflict through PTT

Following the logic of PTT, China is a dissatisfied revisionist great power that presently lacks the power or allies to confront the United States on a global scale. Thus, a regional power transition struggle in East Asia initiated by Beijing against the United States presents a more credible threat. Given Taiwan's proximity and its political importance to the Chinese Communist Party, the island republic presents a likely first target for Chinese revisionism. Consequently, this article will apply PTT's methods to determine which states in the region matter at a grand strategic scale. Table 4 displays the rough total power disparity between China and the U.S.-aligned Quadrilateral Security Dialogue (QSD) states surrounding China that are the most likely to take direct action against Beijing's revisionist ambitions. When measured by CINC, China reaches 90 percent of the total strength of the QSD states, easily within the 20 percent power parity danger zone advocated by PTT theorists. However, as measured by GDP, China only reaches 54 percent of the strength of the QSD status quo states. While CINC and GDP's propensity for exaggerating national power should be kept in mind, it remains clear that China poses a regional threat to U.S. interests. Beijing outmatches all U.S. partners in the region combined in total power as measured by both CINC and GDP. American involvement is essential to confront China, and with U.S. power included the situation becomes far more favorable.

Through PTT's total measures of national power combined with its con-

Table 4. Comparative CINC and GDP values of China and Quadrilateral Security Dialogue states

	CINC (2016)	GDP (2021) millions 2015 USD
United States	0.1330575	$20,338,578
Japan	0.0329674	$4,433,848
India	0.0868413	$2,733,062
Australia	0.0018544	$1,512,962
People's Republic of China	0.2306177	$15,801,911

Sources: CINC data per "National Material Capabilities v6.0 Dataset." GDP from "GDP (constant 2015 US$)—United States, Japan, India, Australia, China," World Bank Group, accessed 29 September 2022.

cept of satisfaction, policy makers can identify which states to integrate further into security arrangements, which to satisfy purely to keep them out of revisionist coalitions, and which to ignore as irrelevant. To confront China, the United States must deepen the Quadrilateral Security Dialogue toward a formal alliance and ensure the continued satisfaction of its members as the basis of regional opposition to Beijing. The QSD already includes both of South and East Asia's leading non-Chinese local powers, India and Japan, as measured by CINC and GDP. Care should be taken to accommodate the satisfaction of these states, particularly rising India, by further investing them in American-led institutions to discourage defection from Washington's coalition. Moreover, steps should be taken to improve the satisfaction of relatively powerful U.S.-aligned or neutral regional states such as South Korea and Indonesia that have shown little interest in directly confronting China.[141] While it may be unrealistic to expect either state to take direct action against Beijing, mollifying them promises to prevent defection of these major regional actors to any Chinese-led revisionist bloc. Beyond the region, the United States should also seek to invest other major strategic partners in opposition to China such as Britain and Germany, which have indicated interest in containing China's rise.[142] Even if little direct military assistance is expected from these states, involving them in opposing China helps invest them in potential economic pressure or embargo schemes necessary in a major war. Thus, America's objective should be to isolate China by keeping powerful regional actors satisfied with the American-led order to either deter China entirely or force it to oppose U.S. regional dominance without major partners.

This application of PTT promises to help policy makers understand regional power dynamics in East Asia, conceptualize what states possess power, and what states are worth going out of America's way to satisfy. These considerations are essential at the grand strategic level because they help inform American diplomacy, but they predict little about the likely course of any conflict. PTT sets the stage for conflict. It does not predict how the actors will perform.

Understanding the Specifics

China's territorial disputes with its neighbors and efforts to build competing international institutions such as the Asian Infrastructure Investment Bank, signal Beijing's dissatisfaction and revisionist ambitions. Taiwan constitutes a logical first target of Beijing's regional power transition challenge, but only military studies can model the specifics of such a conflict. PTT sets the stage for the conflict, now it hands off to military studies to predict the political-military details that threaten to shape a war for Taiwan. The United States and its allies enjoy advantages in Taiwan's highly defensible geography and command of the global commons. Conversely, China possesses advantages in antiaccess/area-denial (A2/AD) weapons and proximity of its industrial heartland to the combat zone. Taiwan may be hopelessly outmatched according to total power calculations of CINC and GDP, but there are compelling arguments that predict a war over the island could go either way.

Taiwan's geography gives the island republic a fighting chance. First and foremost, a direct invasion of Taiwan requires the largest opposed amphibious landing in human history.[143] Amphibious landings are notoriously difficult even for the most experienced militaries, and they require both excellent interservice cooperation and extreme logistical support to execute successfully. Moreover, China's interservice cooperation required to pull off such a complex operation is untested. The People's Republic of China has never staged a sizable opposed amphibious landing and has not fought a major state war since it invaded Vietnam in 1979. Even after clearing the substantial amphibious hurdle at the start of an invasion, most of Taiwan's interior consists of rugged mountains, suburban sprawl, and metropolises, all of which substantially favor the defender. The forested mountains of Taiwan's interior provide innumerable points of concealment for missile systems. Likewise, modern metropolises constitute substantial fortifications in and of themselves, and Taiwan's well-developed metro network promises shelter to move personnel and equipment around combat zones without attracting attention.[144] Even if the Chinese military proves itself highly competent, a direct invasion of Taiwan promises a meat grinder for Chinese forces. To defeat a Chinese invasion, Taiwan does not need to defeat every landing; instead, it needs to inflict sufficient damage on Chinese logistics and amphibious assets to render resupply of invasion forces impossible and eliminate them via attrition. Swift Chinese victory over Taiwan remains unlikely, and short of extraordinary strategic surprise and decapitation, there are few reasons to believe Taiwan will be a cheap acquisition for Beijing. Using the present war in Ukraine as an example, the longer Beijing takes to secure its objectives and the more collateral damage inflicted on the civil population, the greater international support for economic and political consequences for China becomes.

Beyond Taiwan, America and its allies command the global commons through unmatched power projection capacities and possess the capability to blockade China and liquidate its overseas bases. China possesses a relatively

weak nuclear-powered attack submarine fleet, and its surface ships as well as bases abroad are vulnerable to American retaliation in a major escalation.[145] Even if China assembles a formidable navy with potent power-projection capabilities, it does little to help the fact that China is largely surrounded by pro-American states. American allies can and should be provided with anti-shipping capabilities to target any Chinese ships passing through their waters in a conflict. Much like the German High Seas fleet during World War I, all the advanced warships in the world matter little if they cannot safely leave home. Moreover, a general blockade of naval trade to China conducted beyond China's antishipping missile range threatens to impose a heavy economic price on Beijing. China may be able to confront the United States and its allies within East Asia, but its lack of global partners and reach presents a major liability if the West is willing to accept the economic costs of imposing a blockade. America controls the global commons. China does not. If Taiwan can inflict substantial losses on Chinese forces, and if its allies can punish China with severe economic sanctions, then a potential pathway opens to destabilize the Chinese Communist Party and force it to either negotiate or risk collapse.

Conversely, China boasts strong A2/AD capabilities that curtail American power projection in East Asia's littoral waters. Chinese ballistic missiles threaten American access to the region by holding key forward island bases on Okinawa, Guam, and Saipan at risk in the event of a conflict.[146] Likewise, antishipping missiles in continental China project a roughly 600-kilometer area denial radius against American surface naval assets operating in the East and South China seas.[147] Consequently, China threatens simultaneously to impede resupply of Taiwan and sharply degrade American power projection inside the combat zone surrounding the island. In the event of a Chinese invasion or blockade of Taiwan, the United States and its allies will not be able to flood the island with equipment as has been done in Ukraine. Disabling China's antishipping and ballistic missile capabilities requires extensive strikes inside continental China itself. Such strikes remain implausible due to heavy concentrations of Chinese antiaircraft systems and the high mobility of the targeted launcher units.[148] In summary, the United States cannot credibly eliminate China's A2/AD capabilities. It must plan around them in any war over Taiwan. Only a small part of total U.S. power will be available to directly defend Taiwan, and most equipment aid will have to arrive *before* a conflict starts.

Beyond capacity to curtail American power projection, China enjoys its own power projection advantage against Taiwan due to proximity. Whereas American equipment will have to travel to Japan, Australia, or even back to the United States for repair and reinforcement, China can replace losses in theater. Damaged but not destroyed assets will be much easier for China to replace than the United States. China's power projection capabilities fall far short of the United States, but they do not need power projection parity to cross a strait less than 200-kilometers wide. Whereas the United States and its allies must operate aircraft from either vulnerable carriers or forward bases on Okinawa, China

can operate virtually its entire air force in the combat zone from well-defended bases in its own homeland.[149]

Still, China faces key naval and air power projection challenges in any proposed invasion of Taiwan. First, Chinese air forces need to successfully suppress Taiwanese air defense networks and establish air superiority over the island. Taiwan's rugged geography complicates this task, and executing sustained suppression/destruction of enemy air defense operations requires a talented air force coupled with excellent planning. Second, China possesses an overwhelming ground force, but it must transport those forces by air and sea into Taiwan to present a credible threat. China still deploys relatively few amphibious assault ships and will likely rely heavily on civilian vehicle transports to project land forces into Taiwan.[150] If China fails to either control the air or protect its amphibious assets, its invasion risks failure. Consequently, China commands a major regional power projection advantage due to proximity, but it still has to use these capabilities wisely and avoid undermining its whole war effort by wasting them.

Synthesizing PTT with Strategic Specificities in Taiwan

Taken at face value, applying PTT to a Taiwan escalation without an additional military studies perspective suggests overwhelming Chinese victory. However, even a brief overview of the political-military specifics of a potential Taiwan conflict indicates that outcome is far from assured. Both sides enjoy strong advantages and crucial weaknesses. Taiwan's geography and the amphibious nature of any invasion grants substantial advantages to the defender. Moreover, American global power threatens to place China under severe international economic pressure that it is poorly positioned to resist. Conversely, China's A2/AD capabilities negate projection of substantial American military power around Taiwan. China's innate advantages from proximity further strengthen Beijing's hand by shortening its logistical tail and allowing it to operate from the relative safety of its own home territories. War for Taiwan could realistically go either way; in fact, the situation strays dangerously into the territory of parity that PTT argues incentivizes conflict in the first place.

Fortunately, by synthesizing PTT at the grand strategic level with military studies at theater-strategic and operational levels, the steps for U.S. policy makers to defend American dominance in East Asia become clear. At the grand strategic level, China will initiate a regional power transition challenge against the United States as it moves closer to regional power parity. Taiwan represents the logical first target due to its proximity to China, strategic significance inside the first island chain, and political import to the Chinese Communist Party. Thus, we can predict where but not when the first blow of China's power transition challenge will land. To confront this challenge at the grand-strategic level, U.S. policy makers need to:

1. Satisfy and further invest members of the Quadrilateral Security Dialogue in a security arrangement designed to contain China.

2. Satisfy neutral regional actors to deter defection to a Chinese revisionist bloc. Make China face the U.S. regional order alone.
3. Engage major powers outside the region to build support for severe economic consequences, up to and including blockade, if China strikes Taiwan.

Meanwhile, at the theater level, the United States needs to prepare for the war PTT indicates is coming. This article divides these suggestions into those focused specifically on strengthening Taiwan and those intended to improve American or allied capacities to confront Chinese revisionism across the region.

Defending Taiwan

1. Provide Taiwan with its own antishipping, ballistic missile, and unmanned submersible vehicle capabilities to threaten China's amphibious assets, break blockades, and hold China's airfields and ports at risk.
2. Strengthen Taiwan's air-defense capabilities to prevent total Chinese control of airspace to facilitate at least limited U.S. aerial resupply.
3. Provide Taiwan with large volumes of infantry weapons, ammunition, and communication equipment suited for urban combat. Taiwan will largely have to fight with the tools it starts the war with; Ukrainian-scale resupply is not an option.
4. Sponsor Taiwanese creation and training of a territorial defense force along Ukrainian or Polish lines to maximize deployable Taiwanese force against Chinese invasion.

Confronting Chinese Regional Revisionism

1. Ensure forward deployed U.S. air and logistical assets in Japan are sufficient for high-intensity warfare to create a ready reserve of personnel and equipment already in theater.
2. Support comprehensive modernization and expansion of Japanese air, air defense, and naval forces. Japan must be able to protect its airspace and defeat Chinese intrusions into the East China Sea.
3. Prioritize modernization and production of American submersible combat assets suited to hunt Chinese amphibious landing ships and submarines.
4. Enhance American train and equip cooperation with India to make the Indian military a more substantive threat to tie-down Chinese forces outside of the main East Asian combat zone.

Collectively, these policy recommendations seek to strengthen U.S. dominance in East Asia and isolate China. China will challenge the United States as a dissatisfied revisionist, driven by its opposition to America's rules-based order, and Washington must prepare accordingly for war. The authors base the theater-level recommendations on the concept established throughout this ar-

ticle that the state with weaker total power can defeat a stronger state given the right tools, conditions, and suitably limited war goals. This principle applies equally to the United States and China. Global power dominance of the United States combined with its allies does not assure Chinese defeat in Taiwan, and American policy makers should not become complacent by assuming it will. Likewise, China's overwhelming on-paper overmatch against Taiwan does not guarantee a direct invasion will succeed. The authors have chosen the theater strategic policy recommendations on the basis that Taiwanese victory is possible if it is provided with enough weapons and training to irreparably damage invading Chinese forces. Even short of direct Taiwanese victory, inflicting heavy material losses on Beijing undermines its ability to launch further aggression in East Asia and potentially destabilizes the Communist regime. Beyond Taiwan, this article's recommendations focus on either improving capacity of American regional allies to defeat Chinese aggression or enhancing the ability of American forces to subvert China's A2/AD advantages. China presents a formidable but by no means invincible adversary. China can be defeated, and in the process, both the sovereignty of Taiwan and the persistence of American dominance can be secured. In the case of Taiwan, PTT sets the stage for China's revisionist challenge, but strategic specificities in theater provides insight into the likely character of that conflict once it erupts. Far from encouraging war, promoting a disparity in military capability at the operational level is needed to deter Chinese challenges to the security order in East Asia.

Conclusion

Power transition theory has great merit as a systemic theory to guide policy at the grand strategic scale, but it must be synthesized with case-specific strategic studies approaches to predict likely outcomes of conflict. By utilizing military studies methods at the theater-strategic and operational levels, the authors do not believe they are reducing the utility of PTT through theoretical bloat. On the contrary, the authors merely identify the limits of PTT's scope and hand off to more appropriate methods at more finite levels of analysis. From a policy perspective, PTT detects threats, while strategic studies provide insight into confronting those threats. Total measures of power alone such as GDP and CINC overestimate deployable national power and neglect the ability of some "weaker" states to punch above their weight or the inability of "stronger" states to invest in military capabilities.

In each of the cases explored in this article, a stronger power failed to defeat an allegedly *much* weaker adversary. In the Russo-Japanese War, Japan leveraged superior regional power projection and qualitative superiority of its forces to overcome its better equipped but poorly supplied, trained, and deployed Russian adversary. In the Afghan War, America's inability to eliminate the Taliban's Pakistani strongholds due to competing strategic concerns ensured the United States could never translate military superiority into lasting victory. Finally, in the 2022 invasion of Ukraine, Russia's poor military readiness, morale, and ob-

solete doctrines rendered their allegedly overwhelming power mute in the face of more skillful and better-equipped Ukrainian defenders.

Bearing these lessons in mind, the authors apply PTT at the grand strategic scale to explain China's rise as a dissatisfied revisionist and synthesize it with case-specific military and political considerations to inform U.S./Taiwanese countermeasures to Chinese revisionism. This same logic can be applied to virtually any looming power transition conflict. To constructively guide policy, PTT must be used appropriately at the grand strategic level and synthesize itself with traditional strategic studies methods to analyze specific conflicts. When properly used, PTT presents a strong theoretical lens to identify conflict across the globe, but it must divorce itself from the misperception that wars are fought on spreadsheets to reach its fullest potential; total power alone guarantees nothing—only how that power is employed matters.

Endnotes

1. A. F. K. Organski, *World Politics* (New York: Alfred A. Knopf, 1958), 330–32.
2. Ronald Tammen, *Power Transitions: Strategies for the 21st Century* (New York: Chatham House Publishers, 2000), 9.
3. Ronald L. Tammen, Jacek Kugler, and Douglas Lemke, "Foundations of Power Transition Theory," in *Oxford Research Encyclopedia of Politics*, ed. William R. Thompson (New York: Oxford University Press, 2014), 11, https://doi.org/10.1093/acrefore/9780190228637.013.296.
4. Tammen, *Power Transitions*, 21.
5. Douglas Lemke, *Regions of War and Peace* (Cambridge, UK: Cambridge University Press, 2002), 49–50, https://doi.org/10.1017/CBO9780511491511.
6. Jacek Kugler and Douglas Lemke, *Parity and War: Evaluations and Extensions of* The War Ledger (Ann Arbor: University of Michigan Press, 1996).
7. William C. Wohlforth et al., "Testing Balance-of-Power Theory in World History," *European Journal of International Relations* 13, no. 2 (2007): 155–85, https://doi.org/10.1177/1354066107076951.
8. Robert O. Keohane, *Power and Governance in a Partially Globalized World* (New York: Routledge, 1990), 57–59.
9. Tammen, Kugler, and Lemke, "Foundations of Power Transition Theory," 11.
10. Shibley Telhami and Michael Barnett, "Introduction: Identity and Foreign Policy in the Middle East," in *Identity and Foreign Policy in the Middle East*, ed. Shibley Telhami and Michael Barnett (Ithaca, NY: Cornell University Press, 2002), 1–25.
11. Indra de Soysa, John Oneal, Yong-Hee Park, "Testing Power-Transition Theory Using Alternative Measures of National Capabilities," *Journal of Conflict Resolution* 41, no. 4 (1997): 525–26, https://doi.org/10.1177/0022002797041004004.
12. Jonathan DiCicco and Jack Levy, "Power Shifts and Power Problems," *Journal of Conflict Resolution* 43, no.6 (1999): 686–87, https://doi.org/10.1177/0022002799043006001.
13. Bruce Bueno de Mesquita, "Measuring Systemic Polarity," *Journal of Conflict Resolution* 19, no. 2 (1975): 187–216, https://doi.org/10.1177/002200277501900201; and Douglas Lemke and Suzanne Werner, "Power Parity, Commitment to Change, and War," *International Studies Quarterly* 40, no. 2 (June 1996): 235–60, https://doi.org/10.2307/2600958.
14. Tammen, Kugler, and Lemke, "Foundations of Power Transition Theory," 38.
15. Readiness rates for German units in the field hover around 50 percent and bureaucratic inertia cripples procurement and modernization processes across all equipment types. Eva Högl, *Information from the Parliamentary Commissioner of the Armed Forces:*

Annual Report 2021 (Berlin: German Bundestag, 2021), 8, 5. For German GDP in 2021 U.S. dollars, see "GDP (constant 2015 US$)—Germany," World Bank Group, accessed 29 September 2022.

16. Stephen Biddle and Ivan Oelrich, "Future Warfare in the Western Pacific," *International Security* 41, no. 1 (2016): 13, https://doi.org/10.1162/ISEC_a_00249.
17. J. David Singer, "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816–1985," *International Interactions* 14, no. 2 (1987): 115–32, https://doi.org/10.1080/03050628808434695.
18. Kenneth M. Pollack, *Armies of Sand: The Past, Present, and Future of Arab Military Effectiveness* (Oxford, UK: Oxford University Press, 2018), 25, 68, 127.
19. Tammen et al., *Power Transitions*, 82, 91.
20. Ashley J. Tellis et al., *Measuring National Power in the Postindustrial Age* (Santa Monica, CA: Rand, 2000), 6, https://doi.org/10.7249/MR1110.
21. Yegor Gaidar, *Collapse of an Empire: Lessons for Modern Russia* (Washington, DC: Brookings Institution Press, 2007), 75.
22. Michael Beckley, "The Power of Nations: Measuring What Matters," *International Security* 43, no. 2 (2018): 16, 25, 32, https://doi.org/10.1162/isec_a_00328.
23. Mark Gunzinger, *Power Projection: Making the Tough Choices* (Maxwell Air Force Base, AL*:* Air University, 1993), 70–71.
24. Douglas Lemke, "Toward a General Understanding of Parity and War," *Conflict Management and Peace Science* 14, no. 2 (1995): 153, https://doi.org/10.1177/07388942950140020.
25. Consider for instance that during the First Gulf War, American Boeing B-52 Stratofortress bombers based in the United States struck targets inside Iraq and then returned to their bases in the United States with the aid of aerial tankers. Strategic logistics assets enhance power projection in nonlinear ways. See Gunzinger, *Power Projection*, 70–71.
26. Biddle and Oelrich, "Future Warfare in the Western Pacific," 12.
27. Alessandro Scheffler, "Germany: A U-Turn on Defense," in *A Hard Look at Hard Power*, ed. Gary Schmitt, 2d ed. (Carlisle Barracks, PA: Army War College Press, 2020), 107–8.
28. Shawn Davies, Therese Pettersson, and Magnus Öberg, "Organized Violence 1989–2021 and Drone Warfare," *Journal of Peace Research* 59, no. 4 (2022): 593–610, https://doi.org/10.1177/0022343322110842.
29. Carsten Rauch, "Challenging the Power Consensus: GDP, CINC, and Power Transition Theory," *Security Studies* 26, no. 4 (2017): 658, https://doi.org/10.1080/09636412.2017.1336389.
30. Pollack, *Armies of Sand*, 278.
31. Pierre de Dreuzy and Andrea Gilli, "Russia's Military Performance in Ukraine," in *War in Europe: Preliminary Lessons*, ed. Thierry Tardy (Rome: NATO Defense College, 2022).
32. Douglas Lemke, *Regions of War and Peace* (Cambridge, UK: Cambridge University Press, 2002), 50.
33. See "National Material Capabilities v6.0 dataset," Correlates of War, accessed 17 May 2023.
34. Japan's GDP stood at approximately $106 billion USD compared to Russia's GDP of $269 billion in 2015 USD. Russian and Japanese 1904 GDP values calculated from GDP per capita and population from Maddison project data, then adjusted for inflation to 2015 U.S. dollars. See Maddison project dataset, Jutta Bolt and Jan Luiten van Zanden, "Maddison Style Estimates of the Evolution of the World Economy. A New 2020 Update" (Maddison Project Working Paper WP-15, October 2020).
35. Richard Connaughton, *Rising Sun and Tumbling Bear: Russia's War with Japan* (London: Weidenfeld & Nicolson, 2003), 17.
36. Yoji Koda, "The Russo-Japanese War: Primary Causes of Japanese Success," *Naval War College Review* 58, no. 2 (Spring 2005): 21.
37. Prewar Russian planners based their defense plans on the assumption that the Japanese Imperial Army could mobilize 358,809 men; the Japanese Army actually mobi-

lized 1,185,000 men by April 1904. See William Hammac, *The Russo-Japanese War of 1904–1905 and the Evolution of Operational Art* (Fort Leavenworth, KS: U.S. Army Command and General Staff College, 2013), 32.

38. Hammac, *The Russo-Japanese War of 1904–1905 and the Evolution of Operational Art*, 38.
39. Thazha V. Paul, *Asymmetric Conflicts: War Initiation by Weaker Powers* (Cambridge, UK: Cambridge University Press, 1996), 47.
40. Hammac, *The Russo-Japanese War of 1904–1905 and the Evolution of Operational Art*, 45.
41. Koda, "The Russo-Japanese War," 39.
42. Connaughton, *Rising Sun and Tumbling Bear*, 264.
43. Koda, "The Russo-Japanese War," 39.
44. James D. Sisemore, "The Russo-Japanese War: Lessons not Learned" (master's thesis, U.S. Army Command and General Staff College, 2003), 70, 90.
45. Koda, "The Russo-Japanese War," 22.
46. Connaughton, *Rising Sun and Tumbling Bear*, 41.
47. Hammac, *The Russo-Japanese War of 1904–1905 and the Evolution of Operational Art*, 47.
48. Koda, "The Russo-Japanese War," 31.
49. Hammac, *The Russo-Japanese War of 1904–1905 and the Evolution of Operational Art*, 48.
50. Koda, "The Russo-Japanese War," 33.
51. Due to Russia's lack of forward bases, the Baltic Fleet relied on French colonial ports to refuel during its 33,000-kilometer global transit, because most neutral ports refused to service Russian warships. See Hammac, *The Russo-Japanese War of 1904–1905 and the Evolution of Operational Art*, 53–54.
52. Sisemore, "The Russo-Japanese War," 70.
53. Hammac, *The Russo-Japanese War of 1904–1905 and the Evolution of Operational Art*, 42.
54. Sisemore, "The Russo-Japanese War," 37, 58.
55. On paper the total Russian Army outnumbered Japan 5 to 1, with 3.5 million active or trained reserve personnel to Japan's 683,000. However, in Manchuria/Siberia Russia possessed a mere 153,000 at the beginning of the war. See Sisemore, "The Russo-Japanese War," 9–10.
56. Koda, "The Russo-Japanese War," 22.
57. Sisemore, "The Russo-Japanese War," 11.
58. Connaughton, *Rising Sun and Tumbling Bear*, 21, 28.
59. Connaughton, *Rising Sun and Tumbling Bear*, 25, 126.
60. Sisemore, "The Russo-Japanese War,"12.
61. Koda, "The Russo-Japanese War," 25.
62. Hammac, *The Russo-Japanese War of 1904–1905 and the Evolution of Operational Art*, 46.
63. Sisemore, "The Russo-Japanese War," 11, 32.
64. Sisemore, "The Russo-Japanese War," 71.
65. Col M. V. Grulev, quoted in Donald Wright, "Clouds Gathering on the Horizon: The Russian Army and the Preparation of the Imperial Population for War, 1906–1914," *Journal of Military History* 83, no. 4 (2019): 1,137.
66. John Bushnell, *Mutiny Amid Repression: Russian Soldiers in the Revolution of 1905* (Bloomington: Indiana University Press, 1985), 76.
67. Wright, "Clouds Gathering on the Horizon," 1,139.
68. Koda, "The Russo-Japanese War," 29.
69. In 2001, the United States' CINC score stood at .141 to Pakistan's .011 and Afghanistan's .004 in 2000. See "National Material Capabilities v6.0 Dataset."
70. Nan Tian, "20 years of US military Aid to Afghanistan," Stockholm International Peace Research Institute, 22 September 2021.
71. Ryan C. Crocker, quoted in Daniel L. Davis, "Why Is Afghanistan Falling to the Taliban So Fast?," *Guardian*, 14 August 2021.

72. Craig Whitlock, "Unguarded Nation," *Washington Post,* 9 December 2019.
73. Jonathan Schroden, "Afghanistan's Security Forces Versus the Taliban: A Net Assessment," *CTC Sentinel* 14, no. 1 (January 2021): 20–29.
74. Clayton Thomas, *Afghanistan: Background and U.S. Policy: In Brief* (Washington, DC: Congressional Research Service, 2020), 8.
75. Ahmed Rashid, *Taliban: Militant Islam, Oil and Fundamentalism in Central Asia* (New Haven, CT: Yale University Press, 2001), 186–87.
76. Rashid, *Taliban*, 26.
77. Rashid, *Taliban*, 185.
78. Anatol Lieven, *Pakistan: A Hard Country* (New York: Public Affairs, 2011), 409.
79. "Joint Declaration of the United States-Afghanistan Strategic Partnership," press release, White House, 23 May 2005.
80. Central Intelligence Agency, *The CIA World Factbook 2010* (New York: Skyhorse Publishing, 2009), 1.
81. United Nations Development Programme, *United Nations Human Development Report, 2021–2022* (New York: R. R. Donnelley, 2022), 274.
82. "Afghanistan," Transparency International, accessed 23 September 2022.
83. Matt Waldman, "The Sun in the Sky: The Relationship Between Pakistan's ISI and Afghan Insurgents," Crisis States Research Center (discussion paper 18, Harvard University, June 2010), 13–16.
84. Waldman, "The Sun in the Sky," 15.
85. Richard Holbrooke, "Coordinated Support for Afghanistan and Pakistan," *Hampton Roads International Security Quarterly* 9, nos. 2–3 (Spring/Summer 2009): 28–29.
86. Barack H. Obama, "Remarks by the President on a New Strategy for Afghanistan and Pakistan," White House, 27 March 2009.
87. United States Agency for International Development, "U.S. Foreign Assistance by Country: Pakistan," foreignassistance.gov, last updated 10 June 2022.
88. Adrian Ang U-Jin and Olli Pekka Suorsa, "Explaining the PLA's Record-Setting Air Incursions into Taiwan's ADIZ," *Diplomat*, 14 October 2021.
89. Neta C. Crawford, "The U.S. Budgetary Costs of the Post-9/11 Wars," Watson Institute for International and Public Affairs, 1 September 2021, 15.
90. The most recent National Material Capabilities v6.0 calculates Russia's CINC score at .036 versus Ukraine's .007. See "National Material Capabilities v6.0 Dataset." In 2021, Russia's GDP stood at $1.49 trillion compared to Ukraine's $101.45 billion. For Russian and Ukrainian GDP in 2021 U.S. dollars, see "GDP (constant 2015 US$)—Russian Federation, Ukraine," World Bank Group, accessed 30 September 2022.
91. de Dreuzy and Gilli, "Russia's Military Performance in Ukraine," 31–32.
92. Jack Watling and Nick Reynolds, *Operation Z: The Death Throes of an Imperial Delusion* (London: RUSI, 2022).
93. Mason Clark, George Barros, and Kateryna Stepanenko, "Russian Offensive Campaign Assessment, February 27," Institute for the Study of War, 27 February 2022.
94. de Dreuzy and Gilli, "Russia's Military Performance in Ukraine," 30, 32.
95. de Dreuzy and Gilli, "Russia's Military Performance in Ukraine," 26.
96. Mason Clark, George Barros, and Kateryna Stepanenko, "Russian Offensive Campaign Assessment, March 18," Institute for the Study of War, 18 March 2022.
97. Mason Clark et al., "Russian Offensive Campaign Assessment, April 18," Institute for the Study of War, 18 April 2022.
98. Karolina Hird, Mason Clark, and George Barros, "Russian Offensive Campaign Assessment, June 25," Institute for the Study of War, 25 June 2022.
99. Pavel Polityuk and Abedlaziz Boumzar, "Ukraine Says Troops Holding on to Sievierodonetsk, Advance in South," Reuters, 8 June 2022. Throughout the Russian Donbas offensive, Ukraine gradually regained ground in Kharkiv Oblast. See Mason Clark and George Barros, "Russian Offensive Campaign Assessment, May 6," Institute for the Study of War, 6 May 2022.
100. Dan Sabbagh, "Ukraine Continues Kharkiv Offensive Despite Apparent Russian Retaliation," *Guardian*, 12 September 2022.

101. Karolina Hird et al., "Russian Offensive Campaign Assessment, October 5," Institute for the Study of War, 5 October 2022.
102. Russia's largest tank producer, Uralvagonzavod, has publicly announced the creation of dedicated repair and refurbishment facilities for vintage T-62 medium tanks. See Joseph Trevithick, "Russia to 'Modernize' 800 Vintage T-62 Tanks Due to Ukraine Losses: Report," *Drive*, 12 October 2022.
103. Ekaterina Sedlyarova and Ilya Barbanov, "Как мясо—в штурмовую группу. Истории челябинцев, погибших сразу после мобилизации," BBC News, 13 October 2022.
104. Karolina Hird, "Russian Offensive Campaign Assessment, October 14," Institute for the Study of War, 14 October 2022.
105. "Chapter Five: Russia and Eurasia," *Military Balance* 122, no. 1 (2022): 194, 200–1, https://doi.org/10.1080/04597222.2022.2022930.
106. "Chapter Five: Russia and Eurasia," 168.
107. Katarzyna Zysk, "Russian Military Vulnerabilities: Perceptions and Misperceptions," in *Russia Brief Issue 6*, ed. Jeffrey Michaels (Oxford, UK: University of Oxford, 2020), 16.
108. James Byrne et al., "Silicone Lifeline: Western Electronics at the Heart of Russia's War Machine," RUSI, 8 August 2022, 13–14.
109. Frederick W. Kagan and George Barros, "Russian Offensive Campaign Assessment, March 26," Institute for the Study of War, 26 March 2022.
110. Frederick W. Kagan, George Barros, and Kateryna Stepanenko, "Russian Offensive Campaign Assessment, March 29," Institute for the Study of War, 29 March 2022.
111. David Axe, "Russia's Ancient T-62 Tanks Are on the Move in Ukraine," *Forbes*, 6 June 2022.
112. Stijin Mitzer and Jakub Janovsky, "Attack on Europe: Documenting Russian Equipment Losses during the 2022 Russian Invasion of Ukraine," oryxspioenkop.com, 24 February 2022.
113. The British Ministry of Defence credits Russia's failure to deploy explosive reactive armor to poor training and discipline among frontline troops and commanders failing to enforce ERA use. See Ministry of Defence (@DefenceHQ), "The heavy attrition of Russian Main Battle Tanks in Ukraine is highly likely partially due to Russia's failure to fit and properly employ adequate Explosive Reactive Armour (ERA). . . . This suggests that Russian forces have not rectified a culture of poor ERA use, which dates back to the First Chechen War in 1994," Twitter, 18 August 2022.
114. Successful air campaigns target the rival's point(s) of vulnerability most likely to induce the collapse of their forces. Command and logistics present common vulnerabilities, but merely controlling air space and interdicting enemy mobility can be equally dangerous. See John A. Warden III, *The Air Campaign: Planning for Combat* (Washington DC: National Defense University, 1988), 44, 51, 90.
115. Mike Pietrucha, "Amateur Hour Part II: Failing the Air Campaign," *War on the Rocks,* 11 August 2022.
116. Pietrucha, "Amateur Hour Part II."
117. Andrew S. Bowen, *Russia's War in Ukraine: Military and Intelligence Aspects* (Washington, DC: Congressional Research Service, 2022), 3.
118. Justin Bronk, "Is the Russian Air Force Actually Incapable of Complex Air Operations?," RUSI*,* 4 March 2022.
119. Prior to the Russian invasion Ukrainian mobile air defense consisted entirely of Soviet-era systems. See "Chapter Five: Russia and Eurasia," 212.
120. Pietrucha, "Amateur Hour Part II."
121. Justin Bronk, "Getting Serious About SEAD: European Air Forces Must Learn from the Failure of the Russian Air Force over Ukraine," RUSI, 6 April 2022.
122. de Dreuzy and Gilli, "Russia's Military Performance in Ukraine," 37.
123. Stijin Mitzer and Joost Oliemans, "Nascent Capabilities: Russian Armed Drones Over Ukraine," oryxspioenkop.com, 7 April 2022.
124. Watling and Reynolds, *Operation Z*, 11.

125. Pietrucha, "Amateur Hour Part II."
126. Andrew S. Bowen, *Russian Armed Forces: Capabilities* (Washington DC: Congressional Research Service, 2020), 2.
127. "Chapter Four: Europe," *Military Balance* 122, no. 1 (2022): 168.
128. Andrew Radin et al., *The Future of the Russian Military* (Santa Monica, MD: Rand, 2019), 42, https://doi.org/10.7249/RR3099.
129. Bowen, *Russian Armed Forces*, 1.
130. Bowen, "Russia's War in Ukraine," 15.
131. Jim Garamone, "Russian Efforts to Raise Numbers of Troops 'Unlikely to Succeed', U.S. Official Says," DOD News, 29 August 2022.
132. Radin et al., *The Future of the Russian Military*, 66.
133. Kateryna Stepanenko, Grace Mappes, and Frederick W. Kagan, "Russian Offensive Campaign Assessment, September 18," Institute for the Study of War, 18 September 2022.
134. Watling and Reynolds, *Operation Z*, 3.
135. Kateryna Stepanenko, Frederick W. Kagan, and Brian Babcock-Lumish, "Explainer on Russian Conscription, Reserve, and Mobilization," Critical Threats, 5 March 2022.
136. "The Best or Worst of Both Worlds?," Center for Strategic and International Studies, 23 September 2020.
137. Kateryna Stepanenko et al., "Russian Offensive Campaign Assessment, September 23," Institute for the Study of War, 23 September 2022.
138. Stepanenko, Kagan, and Babcock-Lumish, "Explainer on Russian Conscription, Reserve, and Mobilization."
139. de Dreuzy and Gilli, "Russia's Military Performance in Ukraine," 28.
140. Łukasz Kulesa, *Russia's Military after Ukraine: Down but Not Out* (Rome: NATO Defense College, 2022), 3.
141. Ellen Kim and Victor Cha, "Between a Rock and a Hard Place: South Korea's Strategic Dilemmas with China and the United States," *Asia Policy*, no. 21 (January 2016): 112.
142. In 2022, Germany began limited naval and air deployments to the Indo-Pacific. While German deployments remain limited, direct military involvement in the region signals a major transition in German foreign policy toward China. See Sarah Marsh and Sabine Siebold, "Germany Says It Will Expand Military Presence in Indo-Pacific," Reuters, 31 August 2022. Like the United States, Britain views China as a threat to free naval trade in East Asia and still retains a military and logistical presence in the Indo-Pacific. John Hemmings and James Rogers, "Britain and the Quadrilateral," *Journal of Indo-Pacific Affairs*, special issue (2020): 120, 123.
143. Tanner Greer, "Taiwan Can Win a War with China," *Foreign Policy*, 25 September 2018.
144. Greer, "Taiwan Can Win a War with China."
145. Ronald O'Rourke, *China Naval Modernization: Implications for U.S. Navy Capabilities—Background and Issues for Congress* (Washington, DC: Congressional Research Service, 2022), 15.
146. Evan Montgomery, "Contested Primacy in the Western Pacific: China's Rise and the Future of U.S. Power Projection," *International Security* 38, no. 4 (2014): 129, https://doi.org/10.1162/ISEC_a_00160.
147. Biddle and Oelrich, "Future Warfare in the Western Pacific," 13.
148. Biddle and Oelrich, "Future Warfare in the Western Pacific," 8.
149. Montgomery, "Contested Primacy in the Western Pacific," 132.
150. Thomas Shugart, "Mind the Gap, Part 2: The Cross-Strait Potential of China's Civilian Shipping Has Grown," *War on the Rocks,* 12 October 2022.

Intermediate Force Capabilities
Countering Adversaries across the Competition Continuum

Peter Dobias, PhD; and Kyle Christensen

Abstract: This article outlines the relevance of intermediate force capabilities as a key enabler for North Atlantic Treaty Organization (NATO) operations in the gray zone.[1] NATO adversaries, well aware of the NATO thresholds for employment of lethal force, intentionally operate in a way that limits the alliance's options in crisis and conflict situations. At present, these options are often restricted to two extremes of mere presence or the use of lethal force. Summarizing almost two decades of NATO research into nonlethal/intermediate force capabilities, the article examines the applicability of these capabilities across the competition continuum. Finally, the article makes two key observations. First, it identifies future modeling and simulation requirements to represent employment of intermediate force capabilities, and second, it identifies possible promising research and development of subdomains in directed energy nonlethal weapons.

Keywords: hybrid warfare, gray zone, intermediate force capabilities

Introduction

The North Atlantic Treaty Organization's (NATO) 2030 Initiative and Strategic Concept commit the alliance to "prevent crises, manage conflicts and stabilize post-conflict situations" and "ensure that NATO has

Dr. Peter Dobias is the head of Land and Operational Commands Operational Research at Defence Research and Development Canada (DRDC). His earlier roles at DRDC included lead of the Maritime Forces Pacific Operational Research, lead of Metrics Team at Afghanistan-Pakistan Center of U.S. Central Command, and lead of Land Wargaming Group. Kyle D. Christensen is a strategic analyst a the Centre for Operational Research and Analysis, DRDC, Ottawa, Canada. He is currently science advisor at Canadian Joint Operations Command, where he provides strategic analysis advice to the senior leadership. His previous research postings include NATO's Joint Analysis and Lessons Learned Centre, Lisbon, Portugal; the Canadian Joint Warfare Centre, Ottawa; and the Directorate of Maritime Strategy, Ottawa.

Journal of Advanced Military Studies vol. 14, no. 1
Spring 2023
www.usmcu.edu/mcupress
https://doi.org/10.21140/mcuj.20231401010

the full range of capabilities necessary to deter and defend against any threat to the safety and security of our populations."[2] One of the challenges of the current security environment, where our adversaries, well aware of NATO thresholds for employment of lethal force, operate with impunity below the level of armed conflict:

> Adversaries are undertaking acts of aggression that deliberately stay below the lethal force threshold or that ensure a lethal response from NATO would incur costs—undesired escalation, risks of collateral damage including civilian casualties, negative narratives, and other adverse strategic or political outcomes—to the Alliance.[3]

Examples of these activities range from dangerous aerial and maritime approaches, fomenting unrest in third countries, to sponsoring insurgencies and terrorist attacks. In short, NATO's adversaries are undertaking acts of aggression that either deliberately stay below the level of armed conflict threshold or that ensure that any lethal response from NATO would incur costs such as undesired escalation or civilian casualties, all resulting in negative narratives and other adverse strategic outcomes.[4]

Currently, the NATO responses are often limited to two extremes of mere presence or applying lethal force, thus ceding the initiative to the adversaries. As will be discussed in this article, intermediate force capabilities help solve this military problem across the competition continuum, with concept experimentation (wargaming) results highlighting contributions that build through the stages of the *Framework for Future Alliance Operations*.[5] These active means, such as directed energy nonlethal weapons (NLW), cyber, electromagnetic warfare, and information operations help to deliver effects beyond presence but below the threshold of using lethal force.[6]

This article, in summarizing the last 20 years of NATO research in NLW (and more recently intermediate force capabilities), shows how these capabilities could provide NATO with the ability to deter or counter adversaries' activity in the gray zone and facilitate use of lethal force when the latter is justified. The next section discusses the current and future security and operational environment, considering the full spectrum from competition, through to confrontation, to an open conflict. It is then followed by the definition of intermediate force (IF) and intermediate force capabilities as proposed by a recently completed study under NATO Systems Analysis and Studies (SAS) designated SAS-151.[7] After that, the recently developed NATO intermediate force capabilities concept and the associated wargaming campaign are described in greater detail. Finally, it is followed by a discussion of current capability deficiencies and future research and development opportunities as identified by the NATO SAS-151 wargaming campaign.

Security Environment: From Competition to Conflict

Analyses of the international security environment have increasingly drawn at-

tention to what is becoming understood as the "competition continuum," often referred to as the gray zone.[8] Military operations in the space between peace and war where states are currently involved in competition with each other presents unique challenges for military planners. A Rand study exploring these challenges defined the competition continuum (i.e., gray zone) as "an operational space between peace and war, involving coercive actions to change the status quo below a threshold that, in most cases, would prompt a conventional military response, often by blurring the line between military and non-military actions and the attribution for events."[9] *Competition Continuum*, Joint Doctrine Note (JDN) 1-19, defines the competition continuum as "a world of enduring competition conducted through a mixture of cooperation, competition below armed conflict, and armed conflict."[10] These definitions highlight several key aspects/characteristics of the competition continuum as a unique and challenging environment.

First, the competition continuum is truly a continuum and not several distinct zones, points, or steps along an escalation ladder. For instance, academic literature will often break the continuum down for illustrative purposes, such as into cooperation, collaboration, competition, confrontation/conflict, and/or clash/armed warfare.[11] This often gives the false impression that one transitions from one distinct zone to another. However, as noted in *Competition Continuum*, "The competition continuum is not a three-part [or four- or five-part] model substitute for the two-part peace/war model . . . [as] cooperation, competition below armed conflict, and armed conflict can occur simultaneously."[12] As such, one can be involved in simultaneous interactions with the same strategic actor at different points along the competition continuum. Actors can also move up and down the continuum depending on the status of their relationship with another actor and even skip steps along the continuum. Thus, operationalizing activities in the competition continuum involves using all elements of state power (as will be described below) and controlling escalation/deescalation both vertically and horizontally.[13]

Second, adversaries will purposefully and actively try to blur the line(s) between zones in the competition continuum rather than reinforce their differences. One way this is most effective is when an adversary plays on the ambiguity of legal, political, and/or scientific aspects of attribution. This occurs most typically in the cyber domain and has caused challenges for NATO regarding whether a particular malicious act could trigger Article V.[14] In this case, a state knows it has been attacked, but it is not 100 percent sure who is responsible for the attack.[15] The most important of these lines, however, is to blur the line between outright war and the area of competition below armed conflict/open hostilities. This involves generating a situation where it is unclear whether a state of war exists, and if it does, who is a belligerent and who is not.[16] An adversary who can confuse and obfuscate the fact that one is even being attacked holds a distinct strategic advantage in an engagement.

Third, conducting, exploiting, and taking advantage of activities below the

threshold of armed conflict has generally been perceived as a course of action pursued by weaker states against stronger states. Remaining below the threshold of armed conflict enables weaker states to pursue interests in opposition to stronger states and challenge stronger states because they no longer have to engage superior adversaries in head-to-head confrontations.[17] However, exploiting the competition continuum below armed conflict is in fact being used in peer-to-peer or near-peer relationships, and even in situations where stronger states engage in and pursue interests and activities against weaker states.

Consequently, states are actively seeking ways and means to exploit seams and gaps in each other's defense capabilities and security architecture in the current strategic environment. Frank G. Hoffman notes that these activities encompass a "full range of different modes of warfare including conventional capabilities, irregular tactics and formations, terrorist acts including indiscriminate violence and coercion, and criminal disorder. . . . [It] can be conducted by both states and a variety of non-state actors."[18] Thus, activities in the competition continuum involve all elements of state power, actions aimed deliberately below the level of state-on-state use of force, and are typically synchronized and coordinated toward objectives in an organized manner to achieve synergistic effects.[19] This is important because of the immediate relevance of these potential effects on security relationships and agreements. Actions below the level of overt conflict threaten a state's security interests by appearing to call into question its ability to defend its interests.[20]

Finally, activities in the competition continuum are about seizing and maintaining the initiative, causing decision-making challenges, and creating strategic dilemmas for one's adversary. To use a well-known military decision-making framework, it is not about getting inside the observe–orient–decide–act (OODA) loop of an adversary, or about performing the OODA loop faster than an adversary, but about paralyzing and/or breaking the adversary's OODA loop. Ultimately, judicious and controlled application of gray zone tactics, techniques, and capabilities is to create strategic, operational, and/or tactical dilemmas for an opponent. As noted, the aim is to not so much challenge an opponent in a head-to-head confrontation, but rather to constrain the options available to them, thereby maximizing one's operational freedom of movement in the area between peace and war.[21] Because these activities take place below the threshold of armed conflict, they paint opponents into a corner—constraining a state's military, diplomatic, economic, and political decision space by forcing them to either accept the emerging status quo or use force to resolve the dilemma.

Intermediate Force Capabilities and NATO

During the last two decades, NATO pursued a series of nonlethal weapons (NLW) studies. Of these, SAS-078 (NLW Capability-Based Analysis) resulted in the list of NATO NLW requirements approved by NATO Allied Command Operations and Supreme Allied Command Transformation and a list of iden-

tified NATO capability gaps.[22] The study was based on a set of 37 security vignettes that were assessed during two seminar wargames sponsored by the Allied Command Transformation. While reflecting on the realities of the conflicts in Afghanistan and Iraq, the vignette set went beyond counterinsurgency and identified the requirements much more broadly. The identified vignettes encompassed land, maritime, and air operations and included port protection, boarding operations, counterterrorism in maritime and land domain, search and rescue, noncombatant evacuation, convoy operations, megacity failure, and a variety of counterinsurgency vignettes. There were 34 identified counterpersonnel and countermateriel requirements across these 37 vignettes.[23] These included the requirement to warn, tag, stop, move, deny access, suppress, degrade, and disable personnel both individually and in groups in a variety of environments. Similarly, it included requirements to stop, move, degrade, and disable a variety of vessels and vehicles or disable sensor and weapon systems.

Between 2014 and 2017, SAS-094 studied and validated the NLW contribution to mission success.[24] This study reviewed lessons learned from Afghanistan and Iraq and executed a series of wargames and, in conjunction with NATO Defence Against Terrorism Programme of Work, two live NATO Non-Lethal Technology Experiments (NNTEX). The study demonstrated that in many scenarios (checkpoint and access control, patrol, maritime vessel boarding, search and seizure), NLWs enhanced time and space for decision making, improved mission success, and decreased both collateral damage and risk to allied troops.[25] For example, the final report from the maritime experiment observed: "Integrating non-lethal capabilities into the VBSS teams' mission improved their operational effectiveness to warn, move, deny access, suppress individuals, and decrease civilian casualties," and in the conclusions it states, "integrating non-lethal capabilities into VBSS missions demonstrated military utility during NNTEX-15M. The non-lethal systems of the Enhanced CAPSET provided the VBSS teams with additional, and more effective, escalation of force options while conducting their missions."[26] Similarly, the NNTEX-15L concluded that "the non-lethal capabilities improved mission effectiveness, provided additional means of warning and communication, and significantly reduced likelihood of collateral damage."[27]

In 2018, the NATO SAS-133 study introduced the term *intermediate force capabilities* to replace the term nonlethal weapons. The latter term became somewhat controversial and insufficient to describe a wide range of capabilities to deliver often reversible effects that attempt to minimize undesirable casualties or material damage. Furthermore, the definition of nonlethal can be rather controversial, especially when including counter-materiel capabilities.[28]

Finally, in 2019, the NATO SAS-151 study designed and implemented two wargames demonstrating the value of intermediate force capabilities in the maritime domain across tactical, operational, and strategic levels.[29] To enable this cross-level assessment, a novel approach to wargaming was developed.[30] This approach combined several distinct tabletop wargame approaches into

one. For the initial two games this approach consisted of a free kriegspiel (for the tactical/operational levels) and a matrix game (for the strategic level). However, this approach was later improved by adding additional structure to the kriegspiel and modifying the matrix game to include the use of a diplomatic-information-military-economic framework.[31] The impact of intermediate force capabilities and the wargaming approach utilized for the SAS-151 wargame series were demonstrated at the 2020 NATO Concept Development and Experimentation Conference. Later, the NATO Military Committee tasked the Allied Command Transformation to develop an intermediate force capabilities concept, and SAS-151 was asked to support the effort.[32]

In response to the tasking, SAS-151 and the Allied Command Transformation subsequently jointly proposed defining intermediate force as "force below lethal intent to temporarily impair, disrupt, delay, or neutralise targets across all domains" and intermediate force capabilities as "active means below lethal intent that temporarily impair, disrupt, delay, or neutralise targets across all domains and all phases of competition and conflict."[33] Using this definition, intermediate force capabilities became a unifying term encompassing not only NLW (including variety of directed energy capabilities), but also electromagnetic warfare, cyber, influence/information operations, and even stability policing and use of special operations forces. Rand has conducted an independent study that leveraged a custom-built logic model, which was then applied to past operational vignettes.[34] The study addressed the questions of how NLWs contributed to the operation, which NLWs were most applicable in certain contexts, and the effects on adversary actions and tactical risk, among other insights.[35] The results of the study were twofold. One, the model showed that when related to the desired operational outcomes, these capabilities have commonalities that broadly address the hybrid warfare/gray zone requirements. And second, the application to past operational vignettes led to conclusions that the "key outcomes include improved gray-zone capabilities, the ability to operate in environments that would otherwise have been too risky, and enhanced perceptions of U.S. forces."[36]

Intermediate Force Capability Concept Development

The NATO intermediate force capabilities concept was developed and validated by SAS-151 through a series of wargames with the strategic environment progressing from competition to conflict. The first game focused on force protection tasks. These included access point control, handling noncompliant crowds, and dealing with small unmanned aerial systems used to harass or attack protected targets. The considered intermediate force capabilities were largely represented by directed energy nonlethal weapons. The wargame demonstrated the value of directed energy nonlethal weapons, but it also brought up the need to conduct a preemptive information operations campaign to stress the nonlethality/no permanent damage of these systems.[37]

The second wargame focused on the use of intermediate force capabilities in

escalation management at the operational level, again in the maritime domain. While the focus remained on the directed energy nonlethal weapons, there was also an increased role for an information operations campaign. In this case, managing escalation at the tactical level (e.g., managing the threat of the use of force by the adversary's paramilitary units without resorting to lethal force) and extended decision-making space proved invaluable for strategic escalation management (i.e., prevention of a large-scale conflict). The intermediate force capabilities were critical in enabling friendly forces to execute the naval task group's air operations while the adversary simultaneously used small unmanned aerial systems to try to block these operations. The tactical use of intermediate force capabilities enabled the friendly forces to retain task group cohesion and consequently operational and strategic initiative and enabled them to manage escalation. However, the maritime domain wargame reinforced the need for preemptive information operations campaign focused on the safety of directed energy nonlethal weapons. The wargame also brought to attention the need for additional capabilities (electromagnetic warfare and cyber).[38]

The third game shifted to the land domain (capacity building scenario) and began at a higher level on the strategic escalation ladder.[39] The scenario for the wargame involved hostile forces using lethal force against host nation and NATO forces. There was also a threat of an imminent escalation (invasion) of the host nation if NATO forces gave any justification to the opposing forces. The scenario contained the added complexity of hostile forces using civilians as human shields and increasing countermobility challenges for NATO forces. The hostile forces intentionally organized crowds to block NATO quick reaction forces, who were deployed to assist friendly forces under attack. From the friendly force perspective, the possibility of using intermediate force capabilities to facilitate lethal engagement was explored. The intermediate force capabilities enabled more targeted use of lethal force when required, while significantly reducing collateral damage (e.g., intermediate force capabilities stopping hostile vehicles or suppressing the adversary's targeting, to enable the more effective use of lethal force at the place and time of the friendly force's choosing). The wargame also showed that intermediate force capabilities could be effectively used to counter the use of civilians as a countermobility tool. The scenario suggested that the mobility of intermediate force capabilities may be more important than their range/power.[40]

A series of two wargames focusing on the operational use of information operations was then conducted. These two games, apart from introducing a novel approach to wargaming information operations by creating an audience that, while removed from the information operations teams, still participated in information exchange and creation. The game assessed the effectiveness of various information operations capabilities and approaches in forming a strategic situation. One of the key observations was that it would be beneficial (and might be a strategic necessity) for NATO countries to provide intermediate force capabilities to partner/host nations in order to manage domestic escala-

tion, particularly when competing international actors are involved (e.g., another country leveraging its ethnic minorities as a destabilizing factor in another country).[41]

The final game involved joint operations (a contested noncombatant evacuation from a port at the beginning of interstate hostilities). The game tested the previously identified ends, means, and ways of intermediate force capabilities. For example, it considered intermediate force capabilities roles such as countermobility and countering an adversary's use of civilians in a countermobility role, crowd management, stopping/slowing vehicles and vessels, and countering small unmanned aerial systems. It validated the advantages of mobility versus range and power, and it also considered hostile intermediate force capabilities employment against NATO forces. One of the key observations from this particular wargame was the cost of inaction. Without intermediate force capabilities, the NATO forces had only two options: doing nothing or resorting to lethal force. From the strategic perspective both options were costly. NATO forces chose inaction, which led to severe strategic consequences and forced NATO countries to submit to the adversary's conditions. The outcomes of the game informed the final Allied Command Transformation intermediate force capabilities concept workshop and led to the final intermediate force capabilities draft submitted to the Allied Command Transformation.[42]

Apart from informing concept development, the wargaming series led to a recommendation to further develop wargaming and modeling and simulation capabilities to enable better (and higher fidelity) representation of the intermediate force capabilities employment to support doctrine development and options analysis for intermediate force capabilities acquisition. An initial proof of concept was executed in August 2022 under the NATO SAS-MSG-ET-EZ study. The event involved integration of the Command Professional Edition™ (Command PE™) constructive simulation application with a strategic wargame scenario. It led to a series of recommendations that were incorporated in a proposal for a NATO SAS-MSG-180 study that has commenced this year. The study will have two objectives: 1) development of a federated intermediate force capabilities representation in constructive simulations, and 2) integration of modeling and simulation and wargaming to enable high-fidelity validation of tactical capabilities across all levels of warfare.[43]

Capability Deficiencies and Research and Development Opportunities

Apart from leading to the development of the intermediate force capabilities concept, the wargame series provided some insights relevant for future research and development, particularly in the domain of directed energy nonlethal weapons. One system that stood out was the Active Denial System (mm wave).[44] This finding was consistent with the concurrent Rand study.[45] The games suggested that the mobility of the system was often more important than range. This was true about other systems such as the radio-frequency vehicle stopping

device as well. Game three and four also showed quite conclusively that the ability to mount these capabilities on armored vehicles and aircraft, particularly helicopters, would be a game changer.[46] Another capability gap, if addressed, that would provide a significant advantage, is a vehicle-/vessel-stopping device. Ideally, stopping devices should be controllable remotely and should be able to stop large systems, including armored vehicles.[47]

However, the games also conclusively demonstrated that legacy less-lethal and NLW systems (e.g., batons, pepper spray/tear gas, rubber bullets/bean-bag rounds, electro-muscular incapacitation devices [e.g., the Taser™], etc.) can be counterproductive as they can create the impression of the use of excessive force. During the wargame, their employment led to unintended escalation and helped fuel the adversary's narrative that NATO was participating in the oppression of ethnic minorities. In contrast, the directed energy nonlethal weapons were effective at minimizing negative effects. Nevertheless, even long-range directed energy systems were effectively countered by staying out of range of the system. Thus, the notional mobile systems performed much better and had greater operational effect.[48]

Conclusions and Recommendations

NATO adversaries are undertaking acts of aggression that deliberately stay below the lethal force threshold or that ensure a lethal response from NATO would incur undesirable cost to the alliance.[49] NATO capabilities to counter gray-zone actions are limited, particularly at the tactical level, and consequently NATO forces could plausibly find themselves in a situation where the only two options are doing nothing or using lethal force. Both of these reactions might have very negative operational or strategic consequences through emboldening adversaries (former) or unwanted escalation and miscalculation (latter). Intermediate force capabilities provide NATO and its members with a range of options between these two extremes and consequently would enable them to:

- Better manage escalation below the threshold of an armed conflict;
- Better manage escalation in the context of irregular warfare;
- Manage situations where an adversary uses civilians as a weapon (e.g., in countercountermobility scenarios);
- Maintain force protection options in situations where the use of lethal force may be undesirable or problematic from a collateral damage perspective (e.g., in the vicinity of critical infrastructure); and
- When appropriate, facilitate the more effective use of lethal force while reducing undesirable effects.

However, the past research also revealed limitations of the use of tabletop wargames for options comparison between specific intermediate force capabilities.[50] Furthermore, tabletop wargames in general have limited ability to model small unit/platform-level performance. Since further research work is required into doctrine and techniques, tactics, and procedures of intermediate force ca-

pabilities employment, it will be necessary to develop improved fidelity of intermediate force capabilities representation through computer-assisted wargames. At the same time, it is necessary to maintain the ability to assess operational and strategic implications of the intermediate force capabilities employment.[51] At present, there is limited representation of intermediate force capabilities in existing computer wargames; adding capabilities that rely on cognitive responses may be nontrivial.[52] To address these challenges, NATO Science and Technology Board approved a bipanel study designated SAS-MSG-180 that will work on: a) development of better representation of intermediate force capabilities in constructive simulations and b) working on modeling and simulation wargaming integration to improve the ability to assess operational and strategic benefits of intermediate force capabilities (and by extension any cross-domain/cross-level capabilities). The expectation is that this study will result in the intermediate force capabilities representation in existing computer-assisted wargames, while preserving the ability to assess the impact of intermediate force capabilities employment at the operational and strategic level developed by SAS-151.

The second observation, consistent between SAS-151 and the Rand logic model, is that not all intermediate force capabilities have equal tactical and operational benefits. Consistently during wargames, legacy NLW systems (batons, CS gas, rubber bullets, etc.) were the least effective, while the directed energy nonlethal weapons showed the most promise.[53] Of particular benefits in the games were the active denial system (to ensure allied forces' mobility in the presence of civilians), high-power microwave counterunmanned aerial systems, and radio frequency vehicle stopping devices (including the ability to slow or stop heavy vehicles up to and including tanks). One of the observations consistent across all the games was that the range of these systems was secondary to mobility. That means that the capability requirements should prioritize the size, shape, and power requirements over the range. SAS-151 study concluded that small and light enough directed energy nonlethal weapons capabilities to be suitable for airborne and small armored vehicle applications would be more beneficial than having long-range systems.[54]

In summary, more than 20 years of NATO intermediate force capabilities/NLW research showed potential tactical and operational benefits of these capabilities in deterring and countering hostile activities in the gray zone. Future research needs to explore intermediate force capabilities representation in computer-assisted wargames and simulations to enable high-fidelity options comparison for acquisition and doctrine development at individual/platform level.

Endnotes

1. When the authors cite NATO documents, they may reference sources that are not currently publicly available and are part of the authors' personal document collections.
2. *NATO 2030: Making a Strong Alliance Even Stronger* (Brussels, Belgium: NATO, 2021); "Brussels Summit Communique," press release, 14 June 2021; and "Active En-

gagement, Modern Defence: Strategic Concept for the Defence and Security of the Members of the North Atlantic Treaty Organization," NATO, November 2010.

3. NATO Intermediate Force Capability Concept, Fourth Draft, Submitted to NATO Supreme Allied Command Transformation in December 2021.
4. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*, NATO STO TR-SAS-151, December 2022.
5. *Framework for Future Alliance Operations* (Rome, Italy: NATO Defence College, 2018).
6. NATO SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*, Annex E: Draft intermediate force capabilities Concept, NATO STO TR-SAS-151, December 2022.
7. Systems Analysis and Studies is a research panel under NATO Science and Technology Organization. The studies executed by this panel bear designation SAS-XYZ, where XYZ is the number assigned to the study.
8. A review of available literature reveals terms such as irregular, asymmetrical, unconventional, unrestricted, nonlinear, nontraditional, new generation, next generation, full spectrum, political warfare, lawfare, multinodal, multivariant, and pan-domain. Frank G. Hoffman, "Examining Complex Forms of Conflict: Gray Zone and Hybrid Challenges," *PRISM* 7, no. 4 (2018): 30–47.
9. Lyle J. Morris et al., *Gaining Competitive Advantage in the Gray Zone: Response Operations for Coercive Aggression Below the Threshold of Major War* (Santa Monica, CA: Rand, 2019), 8, https://doi.org/10.7249/RR2942.
10. *Competition Continuum*, Joint Doctrine Note 1-19 (Washington, DC: Joint Chiefs of Staff, 2019), 2.
11. Susan LeVine, "Beyond Bean Bags and Rubber Bullets: Intermediate Force Capabilities Across the Competition Continuum," *Joint Force Quarterly*, no. 100 (2021): 19–24; and Mikael Weissmann, "Hybrid Warfare and Hybrid Threats Today and Tomorrow: Towards an Analytical Framework," *Journal on Baltic Security* 5, no. 1 (2019): 17–26, https://doi.org/10.2478/jobs-2019-0002.
12. *Competition Continuum*.
13. Erik Reichborn-Kjennerud and Patrick Cullen, "What Is Hybrid Warfare?," policy brief, Norwegian Institute for International Affairs, January 2016.
14. Christopher Porter and Klara Jordan, "Don't Let Cyber Attribution Debates Tear Apart the NATO Alliance," *Lawfare* (blog), 14 February 2019.
15. A corollary to this situation is when a state is sure who undertook the attack but does not have sufficient evidence. Rory Cormac and Richard J. Aldrich, "Grey Is the New Black: Covert Action and Implausible Deniability," *International Affairs* 94, no. 3 (2018): 477–94, https://doi.org/10.1093/ia/iiy067.
16. Cormac and Aldrich, "Grey Is the New Black."
17. *Countering Anti-Access/Area Denial Challenges: Strategies and Capabilities* (Singapore: S. Rajaratnam School of International Studies and Institute of Defence and Strategic Studies, 2017).
18. Frank G. Hoffman, *Conflict in the 21st Century: The Rise of Hybrid Wars* (Arlington, VA: Potomac Institute for Policy Studies, 2007), 8.
19. Frank G. Hoffman, "The Contemporary Spectrum of Conflict: Protracted, Gray Zone, Ambiguous, and Hybrid Modes of War," in *2016 Index of U.S. Military Strength* (Washington, DC: Heritage Foundation, 2016); and Hal Brands, "Paradoxes of the Gray Zone," Foreign Policy Research Institute, 5 February 2016.
20. Bastian Giegerich, "Hybrid Warfare and the Changing Character of Conflict," *Connections* 15, no. 2 (Spring 2016): 65–72.
21. Andrew F. Krepinevich, Barry Watts, and Robert Work, *Meeting the Anti-Access and Area-Denial Challenge* (Washington, DC: Center for Strategic and Budgetary Assessments, 2003).
22. NATO SAS-078 Research Task Group, *NATO Non-Lethal Weapons Capabilities-Based Assessment*, NATO RTO-TR-SAS-078, December 2012; and NATO SAS-078 Re-

search Task Group, *NATO Non-Lethal Weapons Capabilities-Based Assessment*, Annex C: SAS-078 NLW Requirement Descriptions, NATO RTO-TR-SAS-078, December 2012.
23. NATO SAS-078 Research Task Group, *NATO Non-Lethal Weapons Capabilities-Based Assessment*, chap. 2, NATO RTO-TR-SAS-078, December 2012.
24. NATO SAS-094 Research Task Group, *Analytical Support to the Development and Experimentation of NLW Concepts of Operation and Employment*, NATO STO-TR-SAS-094, April 2017.
25. NATO SAS-094 Research Task Group, "Analytical Support to the Development and Experimentation of NLW Concepts of Operation and Employment."
26. K. Sheehy, *NATO Non-Lethal Technology Exercise 2015 Maritime (NNTEX-15M) Military Utility Assessment,* Defence Against Terrorism Programme of Work, October 2015.
27. P. Dobias and C. Eisler, *NATO Non-Lethal Technology Exercise (NNTEX) 16-Land: First Look*, NATO Science and Technology Organization, October 2016.
28. NATO SAS-133 Research Task Group, *Addressing Obstacles to the Acquisition, Deployment, and Employment of Non-Lethal Weapons—Using Intermediate Force to Bridge the Gap between Presence and Lethal Force*, NATO STO-TR-SAS-133, March 2020.
29. Kyle Christensen et al., *Use of Intermediate Force Capability Game Series: Game 1-NATO Naval Task Group in Port*, Scientific Letter, DRDC-RDDC-2020-L180, Ottawa: Defence R&D—CORA, October 2020; and NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*, Annex F: Use of Intermediate Force Capability Game Series: Game 2—NATO Naval Task Group in Confined Waterway, NATO STO TR-SAS-151, December 2022.
30. Kyle Christensen and Peter Dobias, "Wargaming the Use of Intermediate Force Capabilities in the Gray Zone," *Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* (2021): https://doi.org/10.1177/15485129211010227
31. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*, Annex I: Use of Intermediate Force Capability Game Series: Game 4—Contested Non-Combatant Evacuation Operation Scenario, NATO STO TR-SAS-151, December 2022.
32. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*, Annex A, NATO STO TR-SAS-151, December 2022.
33. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*.
34. Krista Romita Grocholski et al., *How to Effectively Assess the Impact of Non-Lethal Weapons as Intermediate Force Capabilities* (Santa Monica, CA: Rand, 2022), https://doi.org/10.7249/RRA654-1.
35. Grocholski et al., *How to Effectively Assess the Impact of Non-Lethal Weapons as Intermediate Force Capabilities*.
36. Grocholski et al., *How to Effectively Assess the Impact of Non-Lethal Weapons as Intermediate Force Capabilities*.
37. Christensen et al., *Use of Intermediate Force Capability Game Series.*
38. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*.
39. By "higher level on the strategic escalation ladder," the authors mean the third game started closer to the open conflict—below the armed conflict line than the previous games.
40. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*, Annex G: Use of Intermediate Force Capability Game Series: Game 3—Land Operations, NATO STO TR-SAS-151, December 2022.

41. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*, Annex H: Use of Intermediate Force Capability Game Series: Information Operations and Information Warfare, NATO STO TR-SAS-151, December 2022.
42. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*, Annex I: Use of Intermediate Force Capability Game Series: Game 4—Contested Non-Combatant Evacuation Operation Scenario.
43. NATO Exploratory Team SAS-MSG-ET-EZ, *Proof-of-Concept for Integrated Simulation and Wargaming Approach to Representing Intermediate Force Capabilities*, NATO STO-TM-SAS-MSG-ET-EZ, December 2022.
44. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*.
45. Grocholski et al., *How to Effectively Assess the Impact of Non-Lethal Weapons as Intermediate Force Capabilities*.
46. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression,* Annex G and Annex I, NATO STO TR-SAS-151, December 2022.
47. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression,* Annex G and Annex I.
48. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression,* Annex G and Annex I.
49. NATO Intermediate Force Capability Concept, Fourth Draft, Submitted to NATO Supreme Allied Command Transformation in December 2021.
50. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*.
51. Christensen and Dobias, "Wargaming the Use of Intermediate Force Capabilities in the Gray Zone."
52. NATO Exploratory Team SAS-MSG-ET-EZ, *Proof-of-Concept for Integrated Simulation and Wargaming Approach to Representing Intermediate Force Capabilities*, NATO STO-TM-SAS-MSG-ET-EZ, December 2022.
53. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*; Grocholski et al., *How to Effectively Assess the Impact of Non-Lethal Weapons as Intermediate Force Capabilities*; and NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*.
54. NATO Research Task Group SAS-151, *Intermediate Force Capabilities (intermediate force capabilities) Concept Development and Experimentation to Counter Adversary Aggression*, Annex I: Use of Intermediate Force Capability Game Series: Game 4—Contested Non-Combatant Evacuation Operation Scenario.

The Human Weapon System in Gray Zone Competition

Master Sergeant Bonnie L. Rushing, USAF;
and Kyleanne Hunter, PhD

Abstract: Russia's experience in Ukraine highlights the importance of the *human weapon system* in next generation warfare. They show that despite technological superiority and investment in sophisticated weapons and equipment, such as hypersonic missiles, *people* are the core of a successful military strategy. While Russia's invasion of Ukraine has resulted in kinetic and largely conventional warfare, the human weapon system is essential across the range of military operations, particularly in gray zone operations. There may be no place where the human weapon system is more important; strategic and meaningful management of the human weapon system for use in countering gray zone activities may prevent escalation into kinetic operations.

Keywords: hybrid warfare, gray zone, Russia, China, diversity, inclusion, human weapon system, strategic competition, military operations, Joint force, innovation, national security, personnel, talent, recruitment, retention, strategy, policy, workforce management, employment, armed forces, stereotypes, operational effectiveness, equity

At the onset of the conflict between Russia and Ukraine, Russia was (and one could argue still is) the technologically superior force with more than 4,000 military aircraft, 12,000 tanks, 605 naval vessels, 850,000

MSgt Bonnie Rushing is course director and academic faculty at the U.S. Air Force Academy's Military & Strategic Studies Department. She joined the Air Force in 2009 as an airborne linguist and earned her master of science in strategic intelligence from the National Intelligence University. Dr. Kyleanne Hunter is a senior political scientist at Rand. She was a professor of military and strategic studies as well as the director of the Strategy and Warfare Center and associate director of the Institute for Future Conflict at the U.S. Air Force Academy. She earned her MA and PhD from the Josef Korbel School of International Studies at the University of Denver. The views and opinions expressed in the article are the authors' own and do not necessarily represent that of their employers, the Air Force, or the Department of Defense.

Journal of Advanced Military Studies vol. 14, no. 1
Spring 2023
www.usmcu.edu/mcupress
https://doi.org/10.21140/mcuj.20231401011

active duty troops, antisatellite weapons, the world's widest inventory of ballistic and cruise missiles, nuclear forces, hypersonic technology, clandestine and proxy forces, and extensive cyber and information warfare operations.[1] It had a sophisticated military apparatus, coupled with the power of nuclear deterrence. On paper, the invasion of Ukraine should have been a quick victory for Russia. Their military campaign, however, has been largely unsuccessful, and as of fall 2022, Ukraine has fought back and stunningly recovered territory throughout their country.[2]

Russia's experience in Ukraine highlights the importance of the *human weapon system* in next generation warfare. By human weapon system, the authors are referring to the role that individuals play in operations. This includes how diversity, past lived experiences, and unique backgrounds are utilized. Russia's experiences show that despite technological superiority and investment in sophisticated weapons and equipment, such as hypersonic missiles, *people* are the core of a successful military strategy. While Russia's invasion of Ukraine has resulted in kinetic and largely conventional warfare, the human weapon system is essential across the range of military operations, particularly in gray zone operations. The gray zone refers to those activities that take place between peacetime and war. They may be conceived of as gradualist campaigns and employ nonmilitary and quasimilitary tactics that fall below the threshold of conflict. They may include both state and nonstate actors. There may be no place where the human weapon system is more important; strategic and meaningful management of the human weapon system for use in countering gray zone activities may prevent escalation into kinetic operations.

Indeed, Russia can serve as a case study for how poor management of the human weapon system can undermine technological advancements. Intercepted calls, former Russian officials, and social media accounts show the extent to which Russia's human weapon system is in poor shape. "I don't want to be here. I'm not a warrior. I wasn't even f——g trained, to run away from the tanks, for f——k's sake," a disgruntled Russian soldier expressed to his mother over the phone.[3] Another soldier described the deficiencies: "There is simply no discipline, and it will only get worse now that they have mobilized 300,000 people who will be barely trained. . . . The army doctrine is based on punishment, so soldiers get penalized if they mess up. . . . Screw-ups will happen until they change the whole philosophy."[4] A former Russian official shares these details about the state of Vladimir Putin's military forces in the Ukraine war.

As the United States considers next generation warfare, particularly how it is going to shape the force to successfully compete in gray zone operations with Russia and the People's Republic of China (PRC), management of the human weapon system must be a primary concern. As put forward in the *National Security Strategy*, outcompeting the PRC and containing Russia are key strategic priorities.[5] The *National Defense Strategy* prioritizes deterring aggression from these countries and being prepared to prevail in armed conflict if necessary.[6] There are certainly material solutions to addressing these challenges, including

modernized combat systems, tactics, and technologies like long-range precision fires, hypersonic missiles, modernized bomber fleets, littoral combat operations, and other agile mission capabilities.[7] Without a clear understanding of how to leverage the human weapon system, all technological, doctrinal, or strategic advancements are ineffective. In *Force Design 2030*, the Commandant of the Marine Corps further recognizes that while there are structural changes that need to be made, and technological investments (and divestments) to meet the pacing threats, investment in Marines is a linchpin to success in any changes that may occur.[8] This people-focused approach is mirrored in all branches of the U.S. military. The *Army 2030* initiative asserts that despite the need for technical advancements, *people* are the key advantage that the U.S. Army has over its adversaries.[9] The Chief of Naval Operations' *Navigation Plan 2022* asserts that empowering our people is the key warfighting advantage that the Navy brings to the future fight.[10] The chief of staff of the Air Force asserted that "we must empower our incredible airmen to solve any problem" as a key pillar of his *Accelerate Change or Lose* strategy.[11] And the Space Force's *Campaign Support Plan* emphasizes the importance that relationships play in the ability of the Service to fulfill its mission and contribute to national security.[12] The Services clearly recognize that people play a key role in military operations. Continuing to highlight how people contribute to current and future operational needs will strengthen investment in our people and continue to give the United States a competitive edge.

In this article, the authors lay out the case for why management of the human weapon system must continue to be prioritized as a focus of competition in the gray zone. If the United States is going to succeed in outcompeting our near-peer adversaries, it begins with leveraging our people. The authors begin with a discussion of the human weapon system and its application in the military context, including the importance of human weapon systems management to understand internal military capabilities and the external environment in which a military is operating. The article then discusses how the human weapon system will give the United States an edge in gray zone competition. The authors conclude with a discussion of current challenges to human weapon system management and what the Department of Defense and military Services can do to mitigate them.

What Is the Human Weapon System?

The *human weapon system* refers to the role that people play in warfighting. While people have always been essential to warfighting, recent decades have seen a deliberate focus on developing the human weapon system and integrating a whole of person approach to optimizing military operational effectiveness. In 2006, the Uniformed Services University of the Health Sciences hosted the first conference on "Human Performance Optimization."[13] The conference grew out of the realities on the battlefields of Iraq and Afghanistan. In the early years of the Global War on Terrorism, Special Operations Forces leadership de-

clared that "humans are more important than hardware," and in 2004 Deputy Secretary of Defense Paul D. Wolfowitz stated that we need to "develop the next generation of . . . programs designed to optimize human performance and maximize fighting strength."[14]

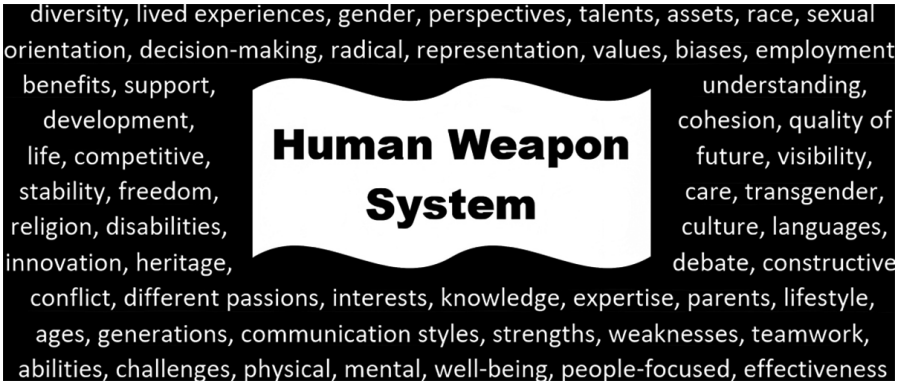
Throughout the early and mid- 2000s, the Department of Defense (DOD) and military Services invested in modernizing human performance.[15] They developed a capabilities-based model of understanding the performance requirements of the operational environment and revamped training, equipment, and standards to meet new needs.[16] Physical performance labs studied the best ways to optimize training to meet the physical demands of military operations. Physical fitness assessments were updated to incorporate both functional movements and provide an overall assessment of individual health and well-being. Nutritionists have been hired to overhaul chow-hall food to ensure that servicemembers are receiving the optimal nutrition to meet physically demanding jobs.[17] As technology rapidly advances, there are additional calls to continue to invest in understanding the human-technology interaction regarding human performance.[18]

While physical performance is one part of the human weapon system, managing the human weapon system is not only about physicality. Mental resilience is just as important as physicality to successful military operations.[19] Developing programs that build mental resilience, investing in mental health care, and incorporating a range of practices to promote self-care, unit cohesion, and build trust are ways that mental resilience has been built into the human weapon system.[20]

Physical performance and mental resilience are aspects of the human weapon system that the military Services can develop within individuals. Fitness and resiliency are largely trainable traits; the Services invest in and customize training for their specific operational needs. Yet, there are untrainable and more intrinsic aspects of the human weapon system that are just as important. Optimizing the human weapon system focuses on ensuring that the United States has the most effective fighting force in the world. It includes both optimizing the physical and mental well-being of U.S. citizens so that they can perform at their best while leveraging the unique backgrounds of individuals to strengthen U.S. national security through employment of diverse skill sets, innovation, and talents.

The nearly two decades of conflict in Iraq and Afghanistan highlighted the importance of how diverse backgrounds were essential for military operations. Lioness teams and Female Engagement Teams were essential for combat operations.[21] All-male infantry units, even with extensive training, could not have the same impact as women, who were able to engage in culturally sensitive and appropriate ways. *Counterinsurgency*, Joint Publication 3-24, codified the need to bring diverse backgrounds to the fight, discussing both the unique role that women play in these types of conflicts and the need for deliberate cultural understanding as part of the way the United States exploits its adversaries.[22]

Figure 1. Examples of the various facets of the human weapon system



Source: Bonnie Rushing, 2022, adapted by MCUP.

Counterinsurgency provides a clear and obvious example of how the unique backgrounds of servicemembers can contribute to military operations. Now, the United States pivots to focus on strategic competition where it is just as critical to understand the importance of the human weapon system and leverage the unique and diverse talents of our servicemembers. Though technology continues to evolve and strategic competition is unfolding, the United States' strength depends on the ability of the DOD to recruit and retain individuals with diverse skills and abilities to take on the country's toughest security challenges.

Leveraging the diversity of our servicemembers is essential for the United States to be competitive across the range of potential military operations—from competition in the gray zone to kinetic combat operations. This idea is reinforced at the executive level. In a recent memorandum, "Memorandum on Revitalizing America's Foreign Policy and National Security Workforce, Institutions, and Partnerships," President Joseph R. Biden notes that diversifying the national security workforce—including the military—is essential for closing mission critical gaps in skills and perspectives.[23] The White House's *National Security Strategy* (NSS) further emphasizes the need to ensure the well-being of our military servicemembers and also to continue to diversify the force as essential components to achieving the United States' strategic goals.[24] As the United States competes against near-peer threats, diversity and innovation are critical. As the NSS states, the primary means by which our national security objectives will be obtained is by "strengthening the national security workforce by recruiting and retaining diverse, high-caliber talent."[25]

The diversity of our workforce is particularly important as our primary adversaries—namely the People's Republic of China and Russia—are engaging in training and recruiting efforts that narrow the opportunities for independent decision making, innovation, morale, and personnel development.[26] Leveraging the human weapon system is thus a strategic asset that is required in today's rapidly changing security environment.

Diversity is critical while considering all potential courses of action. Instead of operating in an echo chamber where a leader surrounds themselves with only like-minded team members that may simply agree on everything or generate similar ideas, diverse teams generate higher levels of innovation, creatively solving problems with higher success rates. Research shows that teams that are diverse consider more facts before deciding and are more likely to accurately interpret facts than homogeneous teams.[27] Diverse teams also create more technologically innovative solutions and are more likely to come up with "radical" solutions that solve the root cause of problems.[28]

Proper management of the human weapon system has an internally and externally reinforcing function. Strategic leaders must understand the needs of their airmen, guardians, soldiers, Marines, and sailors (the internal aspect of the human terrain) while also understanding the ever-changing sociocultural environment (external) in which they operate. The Department of Defense's *Women, Peace, and Security Strategic Framework and Implementation Plan* captures aspects of the reinforcing mechanisms between the internal and external aspects of the human terrain.[29] The ordering of the three defense objectives provides a roadmap for how understanding the human weapon system can help the United States succeed in strategic competition.

> *Defense Objective 1.* The Department of Defense exemplifies a diverse organization that allows for women's meaningful participation across the development, management, and employment of the Joint Force.
> *Defense Objective 2.* Women in partner nations meaningfully participate and serve at all ranks and in all occupations in defense and security sectors.
> *Defense Objective 3.* Partner nation defense and security sectors ensure women and girls are safe and secure and that their human rights are protected, especially during conflict and crisis.

Defense objective 1 is the internal aspect of human terrain. As will be discussed below, it includes understanding how to create policies and pathways that allow for all to meaningfully participate in the institution. From defense objective 1 flows defense objective 2. Success in strategic competition hinges on the United States being a leader in the protection of democracy, human rights, and empowerment. To be an international partner of choice, the United States must model these actions internally. Objective 2 cannot be fully achieved without meaningful investment in objective 1. Finally, defense objective 3 is aimed at creating a more just and secure world. The protection and treatment of women is directly related to the security of states.[30] As the article will show, the ability to have a meaningfully diverse force (objective 1) and build allies and partners around a shared sense of purpose (objective 2) will lead to a more holistic and meaningful understanding of the operational environment, including ensuring that women and girls are protected and empowered during crises.

The Human Weapon System in Gray Zone Competition with the PRC and Russia

The PRC and Russia currently challenge U.S. national security with advanced technology, weapons development, and ceaseless gray zone warfare tactics. Regardless of the ever-changing battlefield and technology environment, the human weapon system remains crucial for operational success, including effectively countering adversarial gray zone threats. Through proactive and positive management of the human weapon system, the United States can succeed in competition and potentially prevent gray zone competition from escalating into kinetic operations.

The *National Defense Strategy* defines these operations as "coercive approaches that may fall below perceived thresholds for U.S. military action and across areas of responsibility of different parts of the U.S. Government."[31] It calls out both the PRC and Russia (as well as other adversaries) for employing gray zone tactics as part of their overall strategies and asserts that campaigning in the gray zone must be a key part of the Joint force's capability in the future threat environment.

The PRC views gray zone activities as a natural extension of how countries exercise power and uses it to build favorable geopolitical conditions without triggering major backlash.[32] The People's Republic of China particularly focused on using gray zone tactics against our allies and partners in the U.S. Indo-Pacific Command area of responsibility, targeting Japan, Vietnam, Taiwan, the Philippines, and India. Russia sees gray zone activities as a way to compete with the United States—and the North Atlantic Treaty Organization (NATO) more broadly—in unconventional ways that go predominantly uncontested because they fall below the threshold of what typically elicits a military response.[33] Russia may also be using gray zone activities to actively shape their near environment to be more favorable for follow-on kinetic military operations (such as the invasion of Ukraine).[34]

Gray zone operations are difficult to counter because they are "gradualist campaigns," combining a mix of traditional military activities with both nonmilitary state and nonstate actors.[35]

As seen in table 1, gray zone operations include a wide range of activities. Military responses to these activities must walk a fine line. Conventional military responses can escalate gray zone activities and draw unwanted international attention, yet the military participates in responding to gray zone activities, and, in many ways, are the key actors responsible for ensuring that activities do not escalate.

The human weapon system is a key component of the military capability to appropriately respond to gray zone activities. Diversity in experiences and backgrounds is essential to countering mis- and disinformation. Diverse understanding of the cues and codes contained in images and language are essential for differentiating real from fake information and for providing important social context as to why certain populations are the targets of falsehoods.[36]

Table 1. Examples of gray zone activities

Tactic	Examples
Information operations	Disinformation campaigns in the media Censorship of dissenting or antigovernment messages
Political coercion/ disruption	Blocking of NATO expansion into Balkans Belt and Road Initiative
Economic coercion/ disruption	Use of military vessels to intimidate or harass commercial shipping Market dominance (i.e., Russia's liquified natural gas energy market dominance)
Disruption of space operations	Jamming and spoofing of satellites Testing offensive space weapons by the PRC and Russia
Proxy forces/ paramilitary	Funding of "little green men" by Russia China's use of commercial fishing vessels to challenge international water access Establishing dual-use bases or ports in contested areas
Military basing in disputed territories	Forward deployed troops or equipment in contested areas Creation of artificial military bases in disputed sea territory Conducting exercises in contested areas
Cyber operations	Breaches of election security systems Hacking into financial systems

Source: courtesy of the authors.

Members of the military are specifically targeted by mis- and disinformation campaigns.[37] A more diverse force will help inoculate servicemembers from falling for this gray zone tactic both through a greater collective understanding of what mis- and disinformation is and through creating more creative solutions to countering fake information.[38] These strategies can also be used to counter mis- and disinformation that appear more broadly in society.

A diverse force also offers our allies and partners a counter to China and Russia's authoritarian politics. China's Belt and Road Initiative not only brings economic impact to countries, but it also imposes China's narrow political and social norms to trading partners. These include anti-LGBTQ policies, a male-sex preference in children, and single-party rule.[39] Authoritarian regimes—or even authoritarian leaning factions among democracies—center many of their policies around misogyny.[40] As a result, women and girls have been central in countering authoritarian regimes.[41] The political and economic coercion by the PRC and Russia seek to undermine or destabilize democracy. As the United

States engages with allies, partners, and potential partners around the world, it has an opportunity to model an alternative to these authoritarian policies. By embodying a diverse and cohesive force, the United States can counter efforts by China and Russia to deny diversity in society. The DOD's implementation guidance to the Women, Peace, and Security Act of 2017 recognizes the importance of promoting diversity with our allies and partners. Meaningful management of the human weapon system will ensure the United States does.[42]

Additionally, proper management of a diverse human weapon system will help ensure that economic investments are meaningful and less prone to corruption. While the military is not the primary arm of economic investment, it works closely with organizations like the U.S. Agency for International Development (USAID) and other private entities to ensure overall security and stability. When women are involved in economic aid, the outcome results in more security and stability in the place where the aid was distributed.[43] Similarly, when national economic growth is coupled with strategies that help women—such as access to child care, equality in education, and equitable health care benefits, women are able to take better advantage of such opportunities and the whole of society is strengthened.[44] A diverse military will help to see where risks to an equitable distribution of economic opportunity may be, as well as key opportunities.

Finally, a diverse force is essential for countering military posturing—including adversarial basing and exercises. Much of the posturing done by our adversaries' militaries is done to elicit a response from the United States as a means of escalating activity. Yet, rather than countering with direct military action, strategic engagement in military exercises can counter the impact of our adversaries in the region. For example, the U.S.-Australia joint exercise Talisman Saber both had the countries engaging with a near-peer competitor and worked to promote gender equality in vulnerable countries in the region.[45] The Joint exercises Viking 18 and Viking 22 integrated gendered components to planning high-north exercises, and the result was an ability to counter Russia's narrative about hard security outcomes.[46]

While the military alone is not responsible for responding to gray zone activities, it plays a significant role. U.S. forces are forward deployed throughout the world, at permanent bases, temporary assignments, and as part of force projection and quick response packages. As such, they are often the first to respond to a crisis. Additionally, forward deployment serves as a soft-power cultural exchange with allies and partners, which gives them key insights into the risks posed by gray zone activities.

Diversity Directives and Personnel Policies

To effectively manage the human weapon system through the recruitment and retention of a diverse force, the DOD and the Services publish directives and policies related to diversity, equity, and inclusion. Furthermore, Service branch leaders update and implement policies related to diversity initiatives to expand

servicemember lifestyle options, quality of life, more inclusive dress and appearance, increased awareness and combat of biases, and care for victims of harassment and assault.

Directives on diversity at the federal level include: Executive Order 13583, Establishing a Coordinated Government-wide Initiative to Promote Diversity and Inclusion in the Federal Workforce; the Women, Peace, and Security Act of 2017; and the Department of Defense DEI Military Equal Opportunity Program.[47] Collectively, these directives aim to advance equity, inclusion, civil rights, racial and gender rights, and equal opportunity at the highest levels of the country's national security infrastructure. They cultivate diverse and dignified workforces, international security, peace, development and afford equitable opportunities in safe environments, free from prohibited discrimination, retaliation, and harassment.

At the Service levels, there are branch-specific policies and regulations related to diversity that mirror much of the federal-level guidance. Each Service branch of the military has similar regulations and goals: *Diversity and Inclusion*, Air Force Instruction 36-7001; the Army's *Army People Strategy: Diversity, Equity, and Inclusion Annex*; the Navy's "Diversity, Equity & Inclusion" objectives; and the Marine Corps' *Talent Management 2030*.[48] These directives tailor national guidance to Service-specific objectives.

They are intended to guide the Services' recruitment and workforce management to attract, recruit, develop, and retain high quality, diverse personnel, with a culture of inclusion to leverage America's talent pool and power of diversity for strategic advantages in the Joint force. This includes diversity of demographics (personal characteristics, age, race/ethnicity, religion, gender, socioeconomic status, family status, disability, sexuality, gender identity, and geographic origin), cognitive and behavioral diversity (neurodivergent individuals, differences in styles of work, thinking, learning, and personality), organizational and structural diversity (institutional background characteristics and experience), and global diversity (knowledge of and experience with foreign languages and cultures, inclusive of both citizens and noncitizens).[49] Services similarly describe diversity as a critical way to enhance decision making, creativity, and the competitive edge to optimize operational effectiveness.[50]

There are additional policies that enable diversity in the military, including freedom of religion, sexual assault prevention and response (known as SAPR, including mandatory annual training to help shape healthy and safe climate and culture, victim care, and support) the repeal of the "Don't Ask, Don't Tell" policy against nonheterosexual servicemembers, and updated guidance welcoming transgender military personnel to serve openly.

Support policies can also indirectly increase diversity. For example, parental leave policies have expanded to include both caregivers, regardless of gender, lengthened in duration, and the inclusion same-sex couples and adoptions. Support is also provided for miscarriages and other fertility-related concerns. Pregnant personnel may continue to fly actively if they desire. New mothers

also have a longer recovery time available prior to an official fitness test requirement.[51]

Dress and appearance regulations have been updated to include more hairstyles, such as ponytails, increased bun and bangs size, more options for women to wear trousers, jewelry, cell phone use, hands in pockets, and more. These changes consider different hair types, comfort, and quality of life while still upholding good order, discipline, and military effectiveness.[52]

Within different Services, physical fitness standards and programs are being updated to both reflect changes in the demographics of the force and the changing nature of military requirements. While not diversity policies directly, they recognize that outdated physical fitness norms may harm servicemembers.[53] In the Air Force, the physical fitness test is now including alternative event choices, the ability to take a "diagnostic" physical fitness test and choose to save it as official afterward, special considerations for certain career fields where higher standards are required, and updated accounting for gender, age, climate acclimation, injuries, and test location altitude.[54] The Marine Corps is updating its body composition standards, allowing for higher weights and body fat percentages to reflect the strength and body mass requirements of women in newly opened career fields.[55]

There are also efforts to improve diversity and remove potential bias from promotion and special assignment positions. Service directives have worked to eliminate references to race, gender, parental status, or religion from all promotion, award, and special assignment boards.[56] While the intent is to remove conscious and unconscious biases that may be inhibiting diversity, it is a large undertaking to remove all identifying markers. In briefings about the updated process, the Services acknowledge that scrubbing records of all identifying information is not yet complete and identifying information is still a part of some records.[57]

The policy commitment to diversity is essential in setting top-down focus on human weapon system management. Yet, personnel-focused policy alone will not ensure U.S. success in gray zone competition. There are ongoing challenges to fully managing the human weapon system that must be addressed.

Challenges to Human Weapon System Management and Employment: Recruitment and Retention

While the United States is currently more successful than the PRC and Russia in managing the human weapon system, it still faces significant challenges, particularly with recruitment and retention of diverse servicemembers and national security professionals. The United States needs a qualified and diverse talent pool to counter adversary gray zone operations and to harness servicemember expertise and innovation for the next generation of warfare. This diversity is what sets America apart from its adversaries: "Indeed, pluralism, inclusion, and diversity are a source of national strength in a rapidly changing world."[58]

The all-volunteer force presents challenges, in that diverse individuals must

self-select into Service. Propensity to serve in the military for all young people continues to decline, and in fiscal year 2022 the Army missed its recruiting goal by approximately 15,000 soldiers, and while other Services met their goals, they had to rely on unplanned bonuses and financial incentives or changes to recruiting targets.[59]

Some demographics do not join or remain in the armed forces as often, for example, "military service still skews heavily towards men (4 out of 5 active duty enlistees are male)" and there is an overrepresentation of the Black American population in the military—specifically, about twice as many Black men serve in the U.S. military as their White male counterparts, numerically.[60] This "can be seen as a double-edged sword. On the one hand, the military has served as an important means of economic mobility for many Black men. On the other hand, the dominance of Black Americans in military service—and therefore among these most likely to be put in harm's way on behalf of the nation—is striking, especially in light of broader current conversations about race, justice and equity."[61]

It is important for the military to represent the people it serves, and that starts at recruiting stations where there must be visibility on diverse personnel in the office, on the materials, and seen in marketing campaigns. Recruiters must enhance their current approaches by following social trends of America's teenage population.[62] Growing youth propensity to serve also requires engaging with youth in the means they are the most comfortable, including popular social media platforms, other information networks, and in trending applications.[63] Distributing correct facts, debunking military stereotypes and myths, and wholly representing the talent pool is crucial for attracting talent of all demographics. Women and their families, for example, fear possible sexual assault in the military and may not join for this reason.[64] Effective delivery of inclusive practices, accurate narratives, and employment of inclusion-focused recruiting not only builds diversity in our ranks, but it also builds trust with the American people and taxpayers who feel wholly represented by troops of every background.

In addition to recruiting diverse and effective talent, *retaining* talent is a challenge for the U.S. military. The *Department of Defense Diversity, Equity, Inclusion, and Accessibility Strategic Plan: Fiscal Years 2022–2023* prioritizes objectives for career progression and retention. There must be opportunities for all demographics to be promoted and serve in key positions and gain career-enhancing education with selection transparency. Additionally, the roadmap prioritizes mentorship for underrepresented groups and elimination of work environment and policy barriers that inhibit equitable practices.[65] Leaders must ensure all members have fair opportunities to develop and succeed. Furthermore, structural concerns may disproportionately impact certain demographics. Women are almost one-third more likely to leave the Service at any time than their male counterparts. Family concerns, including access to adequate and affordable housing, stability for children, family planning support, reliable

and affordable childcare, and other quality of life factors are cited as top reasons women leave the Services.[66] These issues are "inextricably linked" to military readiness.[67] Addressing and rectifying these problems must be a priority to retain talented personnel.

The National Defense Authorization Act for Fiscal Year 2023 (NDAA) addresses some of these issues. The Services and the Department of Defense are hiring more than 2,000 prevention professionals, aimed at changing the culture around sexual harassment and assault and other adverse behaviors that harm recruitment and retention efforts.[68] The NDAA also called for studies to examine compensation models, barriers to home ownership, and promotion pathways for servicemembers.[69] Other recent DOD actions may also address barriers. A recent memorandum on family planning by the secretary of defense seeks to address both privacy and access to care concerns that arose out of the Dobbs vs. Jackson Women's Health Organization (2022) that overturned the provision of a constitutional right to an abortion.[70] While it is too soon to determine the impact of these changes, they show an understanding of the requirement to meet the needs of servicemembers to recruit and retain a diverse force.

Challenges to Human Weapon System Management and Employment: Ties to Operational Effectiveness

Recruitment and retention are not the only challenges that the United States faces regarding the human weapon system. Most of the directives discussed are focused on personnel systems and policies. However, for the Services to fully embrace an action, there must be direct ties to operational effectiveness. The Independent Review Commission on Sexual Assault in the Military found that while personnel issues are frequently talked about as "readiness issues," they are not measured or tracked as such, allowing them to become afterthoughts in the minds of operationally minded military leaders.[71]

Making the direct connections between personnel actions and operational effectiveness is a missing link for the effective management and employment of the human weapon system. Arguments about the "wokeness" of the military highlight that the operational link between a diverse force and operational necessity is not yet fully understood.[72] To fully engage across gray zone activities, the Services need to incorporate the importance of diversity in their doctrine, planning, and professional military education processes.[73]

These actions have proven tactically, operationally, and strategically effective. Gender advisors and gender focal points at the combatant commands have strengthened the United States' strategic partnerships in key contested regions and improved stability and security during humanitarian and disaster response operations.[74] Additionally, they have proven successful in strengthening ties between the DOD and other government agencies, such as the Department of State and USAID.[75] This whole-of-government approach is necessary for gray zone competition.

Building off the success of combatant commands, the DOD and military

Services would benefit from standardizing aspects of the gender advisor workforce and integrating diverse perspectives throughout the operational planning process. Revising the planning process to consider all perspectives will signal to the force that addressing diversity initiatives is essential and leads to an inclusive culture where the safety and well-being of all members is seen as an essential part of security and military operations.[76]

Conclusion

Success in gray zone operations requires thoughtful management of the human weapon system. To do this, the United States must also leverage the diverse talents of its force and fully integrate diverse perspectives into operational planning and readiness. As our near-peer competitors are becoming increasingly narrow in their view of security, there is an opportunity to leverage diverse perspectives and be successful before competition escalates to conflict. While the United States has enacted various diversity initiatives in the past several years, the importance of personnel cannot be eclipsed by investments in technology. As the United States shifts focus on a new pacing threat, the human weapon system remains the linchpin of our success.

Endnotes

1. "Global Firepower 2023," Global Firepower, accessed 29 December 2022; Deganit Paikowsky, "Why Russia Tested Its Anti-Satellite Weapon," *Foreign Policy*, 26 December 2021; and Ian Williams, "Missiles of Russia," Missile Threat, 10 August 2021.
2. Gian Gentile, Raphael S. Cohen, and Dara Massicot, "Ukraine's 1777 Moment," *Rand* (blog), 19 September 2022.
3. Gerrard Kaonga, "Furious Russian Soldier Complains to Mom after Fleeing Ukrainian Offensive," *Newsweek*, 25 November 2022.
4. Daniel Boffey and Pjotr Sauer, " 'We Were Allowed to be Slaughtered': Calls by Russian Forces Intercepted," *Guardian*, 20 December 2022.
5. *National Security Strategy* (Washington, DC: White House, 2022).
6. *2022 National Defense Strategy of the United States of America* (Washington, DC: Department of Defense, 2022).
7. Andrew Feickert, *U.S. Army Long-Range Precision Fires: Background and Issues for Congress* (Washington, DC: Congressional Research Service, 2021); "B-21 Raider Makes Public Debut: Will Become Backbone of the Air Force's Bomber Fleet," U.S. Air Force, 2 December 2022; and "Littoral Combat Regiment," Marines.mil, 2 August 2021.
8. Gen David H. Berger, *Force Design 2030* (Washington, DC: Headquarters Marine Corps, 2020).
9. "Army of 2030," Army.mil, 5 October 2022.
10. *Chief of Naval Operations Navigation Plan 2022* (Washington, DC: U.S. Navy, 2022).
11. Gen Charles Q. Brown Jr., *Accelerate Change or Lose* (Washington, DC: U.S. Air Force, 2020).
12. "2021 USSF Campaign Support Plan," Spaceforce.mil, 6 December 2021.
13. Patricia A. Deuster et al., "Human Performance Optimization: An Evolving Charge to the Department of Defense," *Military Medicine* 172, no. 11 (November 2007): 1133–37, https://doi.org/10.7205/MILMED.172.11.1133.
14. A. P. Tvaryana et al., "Managing the Human Weapon System: A Vision for an Air Force Human-Performance Doctrine," *Air & Space Power Journal* 23, no. 2 (2009).
15. "Human Performance," AFRL, accessed 8 May 2023.

16. Tvaryana et al., "Managing the Human Weapon System," 34–41.
17. Gina Harkins, "The Military is Overhauling Troops' Chow as Obesity Rates Soar," *Military Times*, 19 August 2018.
18. Daniel C. Billing, "The Implications of Emerging Technology on Military Human Performance Research Priorities," *Journal of Science and Medicine in Sport* 24, no. 10 (2021): 947–53, https://doi.org/10.1016/j.jsams.2020.10.007.
19. Jill R. Scheckel, "Preparing the Human Weapon System: Promoting Warrior Resiliency" (research paper, Air War College, Air University, 2010.
20. "Integrated Resilience," U.S. Air Force, 31 December 2022.
21. Ginger E. Beals, "Women Marines in Counterinsurgency Operations: Lioness and Female Engagement Teams" (master's thesis, Marine Corps Command and Staff College, Marine Corps University, Quantico, VA, 2010).
22. *Counterinsurgency*, Joint Publication 3-24 (Washington, DC: Department of Defense, 25 April 2018).
23. "Memorandum on Revitalizing America's Foreign Policy and National Security Workforce, Institutions, and Partnerships," White House, 4 February 2021.
24. *National Security Strategy* (Washington, DC: White House, 2022).
25. *National Security Strategy*.
26. Eric Danko, "Officer and Enlisted Quality Comparison in the U.S. and the PLA," *WildBlue Yonder*, 13 May 2022.
27. Katherine W. Phillips, Katie A Liljenquist, and Margaret A. Neale, "Is the Pain Worth the Gain?: The Advantages and Liabilities of Agreeing with Socially Distinct Newcomers," *Personality and Social Psychology Bulletin* 35, no. 3, 336–50, https://doi.org/10.1177/0146167208328062; and Sheen S. Levine, "Ethnic Diversity Deflates Price Bubbles," *Proceedings of the National Academy of Sciences* 111, no. 52 (2014): 18,524–29, https://doi.org/10.1073/pnas.1407301111.
28. Cristina Díaz-García, Angela González-Moreno, and Francisco Jose Sáez-Martínez, "Gender Diversity within R&D Teams: Its Impact on Radicalness of Innovation," *Innovation* 15, no. 2 (2013): 149–60.
29. *Women, Peace, and Security Strategic Framework and Implementation Plan* (Washington, DC: Department of Defense, 2020).
30. Valerie M. Hudson et al., *Sex and World Peace* (New York: Columbia University Press, 2012).
31. *National Defense Strategy*, 6.
32. Bonny Lin et al., *Competition in the Gray Zone: Countering China's Coercion Against U.S. Allies and Partners in the Indo-Pacific* (Santa Monica, CA: Rand, 2022), https://doi.org/10.7249/RRA594-1.
33. Anthony H. Cordesman and Grace Hwang, *Chronology of Possible Russian Gray Area and Hybrid Warfare Operations* (Washington, DC: Center for Strategic and International Studies, 2020).
34. "Today's Wars Are Fought in the Gray Zone: Here Is Everything You Need to Know About It," Atlantic Council, updated 23 February 2022.
35. Cordesman and Hwong, *Chronology of Possible Russian Gray Area and Hybrid Warfare Operations*.
36. Peng Qi et al., "Improving Fake News Detection by Using an Entity-Enhanced Framework to Fuse Diverse Multimodal Clues," in *MM '21: Proceedings of the 29th ACM International Conference on Multimedia* (New York: Association for Computing Machinery, 2021).
37. Peter W. Singer and Eric B. Johnson, "The Need to Inoculate Military Servicemembers Against Information Threats: The Case for Digital Literacy Training for the Force," *War on the Rocks*, 1 February 2021; and Kristofer Goldsmith, *An Investigation into Foreign Entities Who Are Targeting Troops and Veterans Online* (Culpeper, VA: Vietnam Veterans of America, 2019).
38. Sonner Kehrt, "Troops, Veterans Are Targets in the Disinformation War Even If They Don't Know It Yet," *Warhorse*, 1 September 2022.
39. Min Ye, "Fragmented Motives and Policies: The Belt and Road Initiative in China,"

Journal of East Asian Studies 21, no. 2 (2021): 193–217, https://doi.org/ 10.1017/jea .2021.15.
40. Nitasha Kaul, "The Misogyny of Authoritarians in Contemporary Democracies," *International Studies Review* 23, no. 4 (December 2021): 1619–45, https://doi.org /10.1093/isr/viab028.
41. Kathleen McInnis, Benjamin Jensen, and Jaron Wharton, "Why Dictators Are Afraid of Girls: Rethinking Gender and National Security," *War on the Rocks*, 7 November 2022; and Erica Chenoweth and Zoe Marks "Revenge of the Patriarchs: Why Autocrats Fear Women," *Foreign Affairs*, March/April 2022.
42. *Women, Peace, and Security Strategic Framework and Implementation Plan*.
43. Andrew Beath, Fotini Christia, and Ruben Enikolopov, "Empowering Women through Development Aid: Evidence from a Field Experiment in Afghanistan," *American Political Science Review* 107, no. 3 (2013): 540–57.
44. Mayra Buvinić and Rebecca Furst-Nichols, "Promoting Women's Economic Empowerment: What Works?," *World Bank Research Observer* 31, no. 1 (2016): 59–101.
45. Stéfanie von Hlatky and Andréanne Lacoursière, *Why Gender Matters in the Military and for Its Operations* (Montreal: Centre for International and Defence Policy, 2019).
46. Lisa Sharland and Genevieve Feely, eds., "Women, Peace and Security: Defending Progress and Responding to Emerging Challenges," *Strategic Insights ASPI*, June 2019.
47. "FACT SHEET: U.S. Government Women Peace and Security Report to Congress," statement, White House, 18 July 2022.
48. *Diversity and Inclusion,* Air Force Instruction 36-7001 (Washington, DC: U.S. Air Force); *Army People Strategy: Diversity, Equity, and Inclusion Annex* (Washington, DC: U.S. Army, 2020); "Diversity, Equity & Inclusion," U.S. Navy, updated 3 April 2023; and *Talent Management 2030* (Washington, DC: Headquarters Marine Corps, 2021).
49. *Diversity and Inclusion*.
50. "Diversity, Equity & Inclusion."
51. "Military Parental Leave Program (MPLP)," U.S. Army, accessed 8 May 2023.
52. "Department of the Air Force Guidance Memorandum to DAFI 36-2903, Dress and Personal Appearance of United States Air Force and United States Space Force," memorandum, U.S. Air Force, 31 March 2023.
53. Jeannette Gaudry Haynie et al., *Impacts of Marine Corps Body Composition and Military Appearance Program (BCMAP) Standards on Individual Outcomes and Talent Management* (Santa Monica, CA: Rand, 2022), https://doi.org/10.7249/RRA1189-1; and "Physical Fitness," in *Defense Advisory Committee on Women in the Services: 2019 Annual Report* (Washington, DC: Department of Defense, 2019).
54. *Department of the Air Force Physical Fitness Program*, Department of the Air Force Manual 36-2905 (Washington, DC: Department of the Air Force, 2022).
55. "Marine Corps Study on Body Composition Leads to Change," Marines.mil, 22 August 2022.
56. Lolita Baldor, "Esper Order Aims to Expand Diversity, Skirts Major Decisions," *Washington Post*, 15 July 2020.
57. DACOWITS December 2022 briefing.
58. *National Security Strategy*.
59. "Fall 2021 Propensity Update," Office of People Analytics, 9 August 2022; and *Hearing to Receive Testimony on the Status of Recruiting and Retention Efforts Across the Department of Defense*, 118th Cong. (21 September 2022).
60. Isabel V. Sawhill and Larry Checco, "Black Americans Are Much More Likely to Serve the Nation, in Military and Civilian Roles," Brookings Institution, 27 August 2020.
61. Sawhill and Checco, "Black Americans Are Much More Likely to Serve the Nation, in Military and Civilian Roles."
62. Douglas Yeung et al., *Recruiting Policies and Practices for Women in the Military: Views from the Field* (Santa Monica, CA: Rand, 2017), https://doi.org/10.7249/RR1538.
63. Yeung et al., *Recruiting Policies and Practices for Women in the Military*.
64. Yeung et al., *Recruiting Policies and Practices for Women in the Military*.
65. *Department of Defense Diversity, Equity, Inclusion, and Accessibility Strategic Plan: Fiscal*

Years 2022–2023 (Washington, DC: Office for Diversity, Equity, and Inclusion, Department of Defense, 2022).
66. *Female Active-Duty Personnel: Guidance and Plans Needed for Recruitment and Retention Efforts* (Washington, DC: Government Accountability Office, 2020).
67. "CMSAF Wright Testifies before Congress on Air Force Quality of Life," news, Joint Base San Antonio, 12 February 2019.
68. "DOD Begins Hiring Prevention Workforce," DOD News, 30 November 2022.
69. Providing for the Concurrence by the House in the Senate Amendment to H. Res. 7776, 117th Cong. (2022).
70. "Ensuring Access to Reproductive Healthcare," Department of Defense, 20 October 2022; and Kyleanne M. Hunter et al., *How the* Dobbs *Decision Could Affect U.S. National Security* (Santa Monica, CA: Rand, 2022), https://doi.org/10.7249/PEA2227-1.
71. *Hard Truths and the Duty to Change* (Washington, DC: Independent Review Commission on Military Sexual Assault, 2021).
72. Thomas Spoer, "The Rise of Wokeness in the Military," Heritage Foundation, 9 September 2022.
73. *Hard Truths and the Duty to Change*. See recommendation 3.4.
74. Jim Garamone, "Women, Security, Peace Initiative Militarily Effective," DOD News, 5 November 2020.
75. "Gender Advisors Key to Effective Policy," Council on Foreign Relations, 14 September 2020.
76. *Hard Truths and the Duty to Change*, recommendation 3.4.

"Trying Not to Lose It"
The Allied Disaster in France and the Low Countries, 1940

Richard J. Shuster, PhD

Abstract: This article argues that the critical point of failure in the Allied catastrophe in France and the Low Countries in 1940 was a military plan that ignored key tenets of operational art and planning. In doing so, it points out that the Allies lacked a strategy oriented toward victory, failed to balance their operational factors of time, space, and force, and planned against a single potential enemy course of action. Together, these components set the conditions for a swift Allied defeat that shocked the world.

Keywords: World War II, strategy, Allies, military planning, France, the Low Countries

Introduction

> *Whatever form the final triumph may take, it will be many years before the stain of 1940 is effaced*.
>
> ~ Marc Bloch, 1940[1]

The Allied debacle in 1940 that resulted in a stunning German victory in the West has been a popular subject for decades. How does France, a major military power considered to have one of the greatest armies in the world, spend 20 years planning for a war and then lose it disastrously alongside British, Belgian, and Dutch forces in a mere six weeks? A number of historians have addressed this question from a variety of perspectives. Nonmilitary studies of the defeat in 1940 have examined political, social, and cultural factors

Dr. Richard J. Shuster is a professor in the Joint Military Operations Department at the U.S. Naval War College. He has a PhD in history from George Washington University and is the author of *German Disarmament After World War I: The Diplomacy of International Arms Inspection, 1920–1931* (2006). From 2004–12, Dr. Shuster worked for the Defense Intelligence Agency as a historian and representative to the Naval War College.

Journal of Advanced Military Studies vol. 14, no. 1
Spring 2023
www.usmcu.edu/mcupress
https://doi.org/10.21140/mcuj.20231401012

that have blamed political instability and weakness, social decay, and cultural malaise. Military explanations have focused primarily on doctrine and tactics. These studies point out that once the Germans executed Case Yellow, the assault on France and the Low Countries, the Allies were crushed, with outdated doctrine and methodical tactics proving unable to combat the revolutionary use of massed armor and supporting air power. This study, however, argues that the critical point of failure in the Allied defeat in 1940 occurred prior to the German assault and must first be laid at the doorstep of planning.

Planning has two rather simple aphorisms that indicate that it is difficult to develop a plan that achieves success without the need for refinement or absolute knowledge of enemy intentions: "No plan survives first contact with the enemy" and "the enemy gets a vote." While it is true that any military force will have to adapt to actual conditions on the ground once a campaign or operation begins, a plan has tremendous influence on one's ability to achieve the objective. Overall, planning can be defined as "the deliberate process of determining how (the ways) to use military capabilities (the means) in time and space to achieve objectives (the ends) while considering the associated risks."[2] A sound plan, or one that will have the best chance to achieve its objective, is one that is steeped in operational art, a collection of theoretical elements that inform the commander's vision for a campaign or operation. If, as Professor Milan Vego of the U.S. Naval War College explains, "operational planning is the synthesis of all aspects of operational art theory and practice," then Allied planning in 1940 illustrates a clear disregard of the principles of operational warfare.[3]

When drafting the campaign plan to defend against a German attack, the Allies failed to develop an effective strategy to defeat Germany or to consider and implement key elements of operational art, especially in balancing a clear military objective with the operational factors of time, space, and force. In addition, their tunnel vision in planning solely against the enemy's most likely course of action and disregarding other more dangerous contingencies had catastrophic results. Together, these three major shortcomings that existed even before the German onslaught began in May 1940—a strategy without victory, an imbalance of operational factors, and the preoccupation with a single course of action—spelled doom for the victors of 1918.

The study of the Allied defeat by Germany in 1940 has garnered considerable attention since the end of the war and has contributed much to the understanding of the dramatic event. Memoirs, books, and articles are plentiful, each with their own unique contribution to the still growing historiography. Memoirs generally have focused on military events and actions, typically either ascribing or denying fault in the process. General Maurice Gamelin, the military mind behind the failed defense of the West in 1940, generally dismissed any personal wrongdoing and placed the onus of defeat on his subordinate commanders. His three-volume collection, *Servir*, contains his postwar analysis of his decisions both prior to and during the fight for France and the Low Countries. It also includes a number of contemporary orders, instructions, and

letters that are critical to any understanding of the events of 1940.[4] Marc Bloch's classic account, *Strange Defeat*, provides the author's insights and anguish as a French staff officer during the events.[5] Other Allied memoirs of the debacle, including those by both French and British senior officers involved in the planning and execution of Allied operations in 1940, add to some of the finer details of the personalities and decisions of the period.[6]

Military studies of 1940, of which there are many, have tended to address the strategic and tactical levels of war. Many of these are excellent and are too numerous to cover here.[7] At the strategic level, the focus of many scholarly works has been on military relations between France and its allies in both Western and Eastern Europe as well as French military policies leading up to the war such as arms production and the construction of the Maginot Line. At the tactical level, they have examined the specific actions and fighting capabilities of the Allies in defending France and the Low Countries. French doctrine and the Allied use of armor have emerged as common reasons for the defeat.[8]

Missing in much of the military side of the discussion is the influence of Allied planning on the outcome of events in the spring of 1940. If, as many studies point out, French doctrine hindered the actual execution of tactical actions on the battlefield, then it was essential that the French devise a watertight plan that would maximize the Allied ability to defend the West. But this was not the case. General Gamelin and his staff, with British acquiescence, set up their forces to fail before the German offensive even began. Inadequate doctrine and faulty tactics merely exacerbated an already hopeless situation.

Strategy without Victory

There can be no doubt that our whole plan of campaign was wrong.

~ Marc Bloch[9]

Allied strategy at the eve of the Second World War was predicated upon a long war. Whereas the French and British at the start of the First World War had stated confidently that their troops would be "home by Christmas," there was no such illusion in 1939. On the contrary, the Allies expected a war that would last years. Instead of projecting a sense of victory, Allied strategy was built around the idea of avoiding defeat. The French and British estimated that they could only muster enough military strength to be able to conduct offensive operations, let alone defeat Germany, after two years. And that was contingent on whether the Allies could defend against a German assault for that long. As it turned out, they could not, at least in regard to the continent of Europe. Instead, the war for France and the Low Countries in 1940 was over quickly, a mere six weeks, a duration that mocked Allied strategy, planning, and operations.

From the signing of the Treaty of Versailles in June 1919 to the declaration of war with Germany in September 1939, France had focused on Germany as the primary threat to its national security. Throughout the 1920s, French diplomats established alliances and relations throughout Europe to thwart any po-

tential German aggression, while French military leadership insisted on physical guarantees of security by stationing troops in strategic locations in the demilitarized Rhineland. Allied weapons inspectors roamed the countryside searching for illegal armaments, fortifications, and military personnel.[10] Military planners focused on how to defeat Germany, revising their plans regularly throughout the interwar period. When Adolf Hitler finally unleashed his forces on the West in May 1940, the French High Command had been preparing for the war for 20 years and was only surprised that it had taken so long.

French military strategy advocated two major phases of a war with Germany. France would first remain on the strategic defensive for upwards of two years and then transition to the strategic offensive once it, and its allied partners, had increased their military power in personnel and equipment.[11] Doctrinal concepts within this strategy included the continuous front and an emphasis on firepower (particularly artillery). The strategy of a long, two-phase war was developed and endorsed by both civilian and military leadership. Although the French general staff produced a campaign plan in support of the defensive half of the strategy, it gave little thought to how to operationalize the offensive phase necessary to defeat Germany. Victory remained something to think about in the future.

Similar to the French, the British national security strategy anticipated a long war that involved a strategic defensive to offensive transition. The British planning staff developed a three-phase military strategy in the spring of 1939: defensive military operations to buy time to increase combat power, strategic bombing of Germany (while defeating Italy in North Africa), and a transition to offensive operations with an alliance with the United States in order to defeat Germany.[12] They anticipated a war that would last three years.[13] Similar to the French, the British conception of future offensive operations to defeat Germany remained vague.

After an interwar period punctuated with differences in how to deal with Germany, the two former alliance partners were drawn together in the face of Hitler's aggression. Always fearful of having to fight Germany alone, the French would not risk war without British support.[14] Gamelin considered a French agreement with the British as most urgent and argued that the French could not defend their borders successfully without British military forces.[15] Intelligence sharing increased and staff talks began in March 1939.[16] Prior to the spring, with each side wishing to avoid war at all costs, no combined planning had been conducted. That all changed once Hitler occupied Czechoslovakia in violation of the Munich Agreement. With war on the horizon, the French and British staffs began formal discussions on a basic Allied military strategy.[17]

When developing their long war strategy in 1939, the Allies had differences about potential operations in Scandinavia and the Balkans but were unanimous in their support of preventing the German occupation of Belgium and the Netherlands. Despite having no formal alliance with Belgium, French prime minister Edouard Daladier and British prime minister Neville Chamber-

lain wanted to provide the smaller nation with "maximum help" in the event of a German invasion. Daladier feared that German occupation of the Low Countries would threaten France's main industrial region in the north while Chamberlain stressed that it would threaten London, southern England, and the maritime approaches with air attacks.[18] They both saw benefits in keeping German forces farther to the east and advocated for the defense of as much Belgian territory as possible.[19] Most importantly, the idea that Germany would focus an attack on the west in central Belgium emerged as a strategic assumption that influenced all subsequent planning.

Ironically, the Allies had planned for years to avoid the bloodletting of the First World War, and now they devised a military strategy aimed at repeating a long war of attrition followed by an ultimate offensive. The initial campaign objective was somewhat amorphous: not the defeat of Germany but the defense of a line that would be defined by the military leadership of France. What or where to defend now lay in the hands of General Maurice Gamelin, the French and Supreme Allied commander, who had been given complete freedom of action by Daladier to draft the plan to defend the West.[20] Although Gamelin developed a French plan, work still needed to be done on an Allied plan. None existed when the war erupted in September 1939.[21]

The philosophy of planning not to lose permeated Allied thinking in 1939/40 and is evident in all of Gamelin's plans in this period. Gamelin developed the Allied campaign plan with a laser-like focus on Belgium, and to a lesser degree the Netherlands. For years, the French High Command had been focused on a German advance through Belgium.[22] Only here, Gamelin thought, could the Germans achieve decisive results.[23] His operational vision, however, suffered from severe myopia. He developed three variations of a campaign plan for the defense of the West, and all three—Escaut (Plan E), Dyle (Plan D), and the Breda variant—had only slight variations of the same concept that required French and British forces to move as rapidly as possible into Belgium to check the expected German advance.

The Allies had complete confidence in the outcome of the upcoming defensive fight, particularly in the French Army's capabilities. Although they would later complain of a clear superiority in German capabilities, in reality the relative combat power was roughly equal, with the exception of a superiority in the size of German air forces. Chamberlain claimed that the Germans had "missed the bus" when they did not begin their offensive in 1939. When the Germans finally attacked the Low Countries on 10 May 1940, Gamelin responded with almost a smug confidence that the Allies would repel the hated enemy. His counterpart in Britain, General William Edmund Ironside, chief of the Imperial General Staff, had no doubt of Allied success.[24]

Allied confidence was misplaced, to say the least. Gamelin's plan to defend the West pleased Allied civilian leadership but his solution to avoid defeat by focusing his efforts on a defense of central Belgium would create a cascading series of disasters that Allied tactics and doctrine could not overcome. Daladier,

like Gamelin, believed that the primary mission of the French Army was to prevent defeat.[25] General Ironside admitted that the Allied outlook was hardly inspirational with the comment that "a war cannot, however, be won merely by trying not to lose it."[26] Yet, he threw his wholehearted support behind the plan. The commander of the British Field Force, General John Vereker Gort, stated at an Allied meeting in September 1939 that "the war can be lost in France or Belgium, even though it perhaps cannot be won there."[27] Arguing that their national interests were at stake, the British Chiefs of Staff Committee went so far as to advocate an Allied advance into Belgium if the Germans invaded the Netherlands, even if the Allies encountered *Belgian* resistance to their movement.[28]

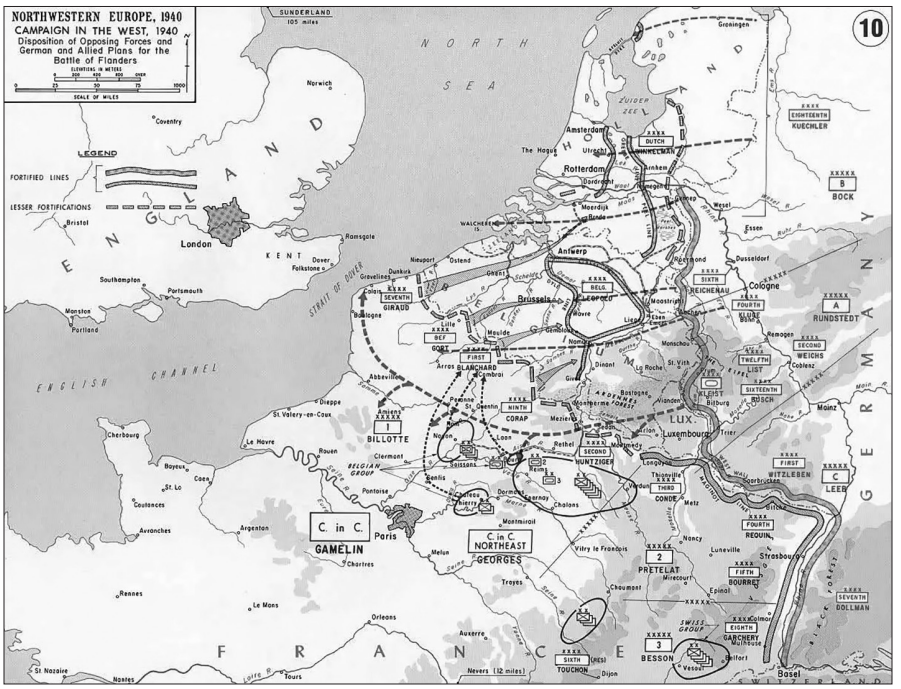
By the time the Germans launched their assault on the West, French and British national and military leadership were in full agreement that they would advance their forces into Belgium to defend a line that was kilometers from the French border. The French would keep the fight away from their northern industrial region and add British, Belgian, and Dutch forces in doing so. The British would protect the coastal areas that could be used to threaten Britain with air and submarine attacks. Any ideas of defeating Germany were put on hold. For now, all they had to do was to fight the Germans to a standstill, and all would not be lost. When the Germans attacked the Low Countries on 10 May, the Allies reacted methodically with their advance and confidently in their expectations to prevent the German occupation of an area that both considered critical to their national interests. When the German main effort, however, appeared in the vicinity of Sedan, far to the south of the expected enemy line of operation, the Allied strategy of a long war and its supporting defensive campaign plan were laid to waste.

Imbalance of Operational Factors

> *Not only did we meet the enemy too often in unexpected places, but for the most part, especially, and with increasing frequency, in a way which neither the High Command nor, as a result, the rank and file had anticipated.*
>
> ~ Marc Bloch[29]

Years after the end of the war, the mastermind behind the German plan to defeat France and the Low Countries, General Erich von Manstein, summarized the intent of Case Yellow simply by explaining that the Germans "just did the obvious thing; we attacked the enemy's weakest point."[30] Those simple words illustrate the essence of operational art, more specifically the balance of operational factors of time, space, and force in order to achieve the objective. Initial German plans in 1940 placed the main effort in the north to swing through Belgium and the Netherlands, but Hitler rejected the idea as too predictable and became enamored with Manstein's idea.[31] Although the German Chief of Staff, General Franz Halder, modified Manstein's plan, he remained true to the critical importance of placing the German main effort against the

Map 1. Plan Yellow vs. Plan D (Breda variant)



Source: Department of History, United States Military Academy, adapted by MCUP.

weak French center, between the onrushing Allied forces in the north and the formidable defenses of the Maginot Line to the south. For the Germans, it was the perfect setup of forces in time and space to defeat the Belgian, Dutch, and the best French and British forces in the north before executing the follow-up operation (Case Red) that would knock France out of the war. A concentration of German armor in highly mobile panzer divisions under the bold leadership of General Heinz Guderian would lead the main assault through the Ardennes to the Channel coast. For the French and British, who would send the bulk of their mobile forces racing into the Low Countries while failing to understand the importance of concentrated armor, it was a recipe for disaster (map 1).

French plans regarding war with Germany had long focused on the need to avoid fighting on the French frontier, choosing instead to base their operational idea on an advance into Belgium. Gamelin put all of his energy into this concept, developing three variations of a campaign plan distinguishable only at the point in which Allied forces would meet their German counterparts on the field of battle in Belgium. In the fall of 1939, Gamelin had drawn up two versions of his plan—Plan E and Plan D—and in the spring of 1940, he added his third version—the Breda variant—that became the plan that the Allies executed in May. All three versions, but particularly the Breda variant, were conceptually flawed. In essence, the plans were not much more than deployment orders, with Gamelin concerned most about the need to "complete his deployment" before contact was made with German forces.[32]

The differences between the three versions of the plan were mainly in the locations of the Allied defensive line. Force dispositions in each plan were similar: the French 16th Corps (and then all of Seventh Army), the British Expeditionary Force (BEF), and the French First Army would move into Belgium and incorporate the Belgians into their defensive line while the French Ninth and Second Armies remained essentially static along the French frontier up until the Maginot Line. Plan E, developed by Gamelin in September, required the Allies to secure and hold the line of the Escaut River, forming a junction with Belgian forces at Ghent, as well as holding the French frontier. Forward elements would push east of the Escaut and fight a delaying action along the Dendre River while much of the Seventh Army was held in reserve near Reims. In November, Gamelin drafted Plan D, which pushed Allied forces roughly another 65–95 kilometers east to the Dyle River, along the Namur-Wavre-Louvain-Antwerp line, and incorporated the entire Seventh Army into the northern part of the line.[33] Allied forces would link up with Belgian forces in the vicinity of Antwerp and attempt to occupy the Dutch islands of Walcheren and Beveland. Cavalry would advance forward to act as a screen to delay any potential German forces.[34] Gamelin then added the Breda Variant in March 1940, which pushed Allied forces, specifically the French Seventh Army that had once been held in reserve, even farther to the east to Breda/Turnhout in the Netherlands. The farther east the Allies moved, however, the higher risk they incurred as they fell into the German trap.

Gamelin's plans lacked the key ingredient for success, or in this case, the ability to employ superior (or even sufficient) force at the right time and place in order to achieve the objective. Instead, they created conditions that ceded any potential advantage to the enemy. When assuming that the Germans would concentrate on central Belgium, he employed his best equipped and most mobile forces in Belgium in the north, relied on the static Maginot Line in the south, and considered the center of his defensive lines in the rugged Ardennes area an economy of effort. The Germans, of course, attacked the weakest part of the line with overwhelming force and maneuvered with alacrity to the English Channel, cutting off Gamelin's most precious forces in the north and then crushing the entrapped Allied forces in a classic hammer and anvil approach.

When developing his plan to defend against a German attack, Gamelin's fatal flaw was to focus on force in his desire to achieve parity regarding overall numbers of divisions vis-à-vis Germany. He expected his forces to defend Belgian and even Dutch territory until sufficient Allied offensive forces could be built up, forcing the Germans into a long war of attrition to offset their advantages in manpower and mobile warfare. Gamelin believed that the Allies would not have a superiority in force at any time before 1941 and would take no decisive action without it.[35] With a long war strategy in mind, Gamelin envisioned a grueling repetition of the fighting in the First Word War, with Allied numbers once again eventually turning the tide in their favor.

Gamelin's focus on the importance of force, particularly regarding numbers

of divisions, was a major factor in how and why he devised his campaign plan to move into Belgium to counter a German assault. The preoccupation with numerical equality, designing a campaign that would tally up Allied divisions and ensure continued manpower over time, was his answer to "trying not to lose it." French preoccupation with Germany's superior combat potential in terms of numbers of personnel can first be seen in the interwar years, when the French worked tirelessly to prevent the militarization of all sources of manpower such as the regular army, police, and paramilitary organizations.[36] As the chances for war increased in the late 1930s, Gamelin had concluded that France did not have enough manpower to defend against a German assault.[37] Throughout the period of the "Phony War," Gamelin reiterated that France had a clear disadvantage in the numbers of divisions vis-à-vis Germany.[38] His answer to this age old problem was to devise a campaign that would have the best chance of adding precious British, Belgian, and eventually Dutch divisions to the Allied cause. In this manner, he could create numerical equality with the Germans and then fight a largely static form of warfare until he was able to build up a numerical superiority to shift to the offensive. When Germany attacked in May, Gamelin had successfully evened the score as far as force. In fact, 135 German divisions faced 151 Allied divisions.[39] In reaching parity, however, Gamelin had actually sacrificed advantages in space and time and increased risk to the mission and to his forces.

Gamelin's intent to increase Allied forces was centered first and foremost on the need to keep Britain in the fight. Although the size of the BEF was small in this period, a mere 10 divisions by May 1940, Gamelin envisaged a long war, and over time Britain would be able to produce a large number of quality divisions to help tip the scales against any German force advantage. According to Colonel Jacques Minart, who served on Gamelin's staff in 1939–40, Gamelin's impetus to move his forces into Belgium was his fear that the German occupation of Belgium and the Netherlands would knock the British out of the war or at least force them to withdraw from the continent.[40] Ironically, of course, this decision helped lead to France's defeat while Britain was able to survive the German capture of the Low Countries and the subsequent air attack on Britain.

Fear of losing British support weighed on Gamelin at the start of the war. In one of his first meetings with the British in September, Gamelin claimed that he needed as much British help as possible to defend against the expected German attack in the Low Countries.[41] French intelligence produced reports that influenced Gamelin in the fall, warning that Germany had the potential to double its current military strength with the reconstitution and training of military age personnel.[42] The French military representatives who had been having staff conversations with the British warned in September that German occupation of the Flemish coast would create serious danger for Britain.[43] In October, Gamelin urged General Ironside to increase the number of British divisions to the continent, emphasizing the "necessity of the common effort which we must undertake in regard to effective strengths so that we may not find ourselves this

coming Spring in a dangerous state of inferiority in face of the German forces."[44] Ironside was well aware of Gamelin's force sensitivity and understood that the French would "continue to pressure us to send the maximum number of divisions to France."[45] The British questioned Gamelin's plans at times, particularly force dispositions and defensive preparations, but never the idea to move into Belgium.[46]

When focusing on force, Gamelin also planned to add Belgium's 22 divisions to the Allied defense against Germany. His desire to incorporate Belgian divisions into the Allied defense was a consistent theme in his planning, another number that he could add to the force balance sheet to offset German force advantages. In drafting Plan E in September, Gamelin argued that his force dispositions in Belgium along the Escaut would allow the Belgian Army to reconstitute its forces (expected to be in combat with the Germans) and "to take its place on the Allied front."[47] He believed that employing his forces along the south bank of the Escaut River had defensive advantages, but more importantly would rally the Belgians, adding their divisions to the Allied defense against the German attack.[48] He repeated his desire to rally the Belgians and incorporate their forces into the Allied defensive line in another meeting with the British on 19 November.[49]

Finally, Gamelin envisioned the Allied move into Belgium could be a way of adding Dutch forces to his overall plan. First considered in a September instruction to Georges, he pointed out that his Plan E to move into Belgium would be a prelude to any land support given to the Netherlands.[50] During Allied meetings in November 1939, when Gamelin presented his Dyle version of his campaign plan to the British, he also began to examine the question of how to add the Dutch to the Allied force mix. At this point Gamelin had long since settled on a plan based on Allied movement into Belgium but now laid out his plan for an Allied move to the Dyle that included sending forces into Dutch territory as well. Elements of the French Seventh Division on the far left of the Allied line would occupy the mouth of the Escaut and the two Dutch islands of Walcheren and Beveland to link up with Dutch forces. More importantly, he emphasized the disadvantage in French force numbers and that the additional 22 Belgian and 10 Dutch divisions were necessary to even out the numbers against Germany.[51]

Gamelin eventually relented completely to his force preoccupation in his Breda variant of the Dyle Plan. After warning Daladier in January 1940 that he needed to address the lack of Allied "numerical equality" with the Germans, he modified his plan further to ensure the addition of Dutch divisions and to help protect Belgian forces.[52] In an instruction to George on 12 March 1940, Gamelin first pointed out that the Dyle Plan placed the Seventh Army north of Antwerp in order to ensure the security of the lower Escaut and to forge a connection with the Belgians and Dutch. To maintain communication with the Dutch and add their forces to the Allied defensive line, Gamelin now pushed the Seventh Army even farther to the east, toward the Breda-Saint-Leonard or

even the Tilburg-Turnhout line. He argued that this extension of the Allied front to the east would actually reduce risk to his forces in helping to reinforce the Belgian forces against the German assault.[53] Now, elements of the Seventh Army would extend another 48 kilometers to the east, farther away from the French frontier, and to the extreme north of the Allied line, isolated from the main fighting that would soon take place 240 kilometers to the south.

Gamelin's subordinate commanders' concerns with the Breda variant fell on deaf ears. Georges was one of the only voices though that brought up the uncomfortable notion that the Germans may not attack in strength in Belgium but rather make their main effort possibly in the center of the French defensive line. He complained that the new modifications to the Dyle Plan stripped away his reserve forces and placed them far to the north. Gamelin, however, rebuffed Georges's critique, arguing that it was out of the question to abandon the Netherlands and that it was necessary "to make an effort to at least give a hand to the Dutch and try to have a land communication with them."[54] In exchange for 10 Dutch divisions that were overwhelmed quickly by the Germans in May, Gamelin had further entrapped some of his best forces far to the north, with little hope of either holding the secondary German effort in the north or supporting the Allied defense against the German primary effort to the south.

The focus on increasing force by advancing into Belgium led to an imbalance with time and space that the Allies could simply not overcome once the fighting erupted. When Allied intelligence reported that the long anticipated German attack had begun, French and British forces followed Gamelin's tragic script. They reached the Dyle line with little resistance, as the Seventh Army moved steadily toward Breda, and along with the BEF, engaged what they thought was the German main effort in central Belgium. Large engagements with German Army Group B occurred in Hannut and Gembloux to prevent the Germans from crossing Gamelin's "open plains" of Belgium, while the bulk of German armored divisions in Army Group A overran the much smaller Belgian forces in the Ardennes and the French forces in Sedan. The Allied line crumbled.

With all their planning focused on moving into Belgium, neither France nor Britain gave much, if any, thought to the time it may take to employ their forces anywhere else. Once locked into combat with German forces, the Allies faced a difficult fighting withdrawal, and any notion of repositioning their best forces to meet the German main effort along the Meuse in the Sedan area was overcome by the tyranny of distance and time. The French had also failed to assess with any accuracy the area facing the center of their defensive line. Overestimating the defensive value of both the Ardennes and the Meuse, these natural defenses were rendered impotent when faced with overwhelming local superiority of force. Most importantly, the Germans had a far superior force-to-space ratio in the sector of main effort and the point of main attack—the area between Sedan and Dinant.[55] With the Allies locked in a ferocious battle with German armored and infantry forces in Belgium, they were unable to dis-

engage their best forces to meet the concentrated German armored forces 240 kilometers to the south that were breaching the Meuse River in the Sedan area and would soon reach the English Channel coast.

Scripting Disaster: Tunnel Vision and Mirror Imaging

> *Only the most elastic of minds can make sufficient allowance for the unexpected—which means, in most cases, what the enemy will do*.
>
> ~ Marc Bloch[56]

When the Germans attacked on the morning of 10 May 1940, Gamelin, in reference to his Dyle/Breda Plan, asked his subordinate commander: "Since the Belgians have appealed to us, can you see how we can do anything else?"[57] Georges affirmed the expected response; there simply was no other plan.[58] The rigid adherence to what amounted to a single course of action reveals a stunning lack of creativity and sound operational thinking. The Allies had written a script on how to fight Germany with a singular focus on moving into Belgium as quickly as possible, memorized it in full, and then performed it with aplomb. The French Seventh Army arrived in the vicinity of Breda, British forces reached the Dyle, and the French First Army arrived on the Wavre-Namur line, all with no significant issues. Now the Germans just had to follow the same script and the Allies would be the saviors of Europe once again. The Germans, however, had other ideas. Manstein's belief that the best solution was not necessarily the most logical solution—because the enemy could be planning along identical lines—is both simple and instructive.[59]

Campaign and operational planning are most effective when multiple courses of action are generated to achieve an objective, and then each course of action is evaluated against potential enemy courses of action. This is an art, not a science, and therefore relies on the application of sound military theory, with a dose of creativity. Yet, the Allies had developed only slight variations of one course of action that matched up perfectly with a single, most likely German course of action. In most cases a commander does not have a perfect awareness of enemy intentions so it is imperative to consider the impact that various potential enemy actions could have upon one's forces in order to improve the effectiveness of the plan. In theory and practice, the Germans understood that it was wise to adopt the enemy's most dangerous enemy course of action as a basis for one's planning in order to reduce risk.[60] Current U.S. joint doctrine, for instance, stipulates that each friendly course of action should be analyzed (or wargamed) against the enemy's most likely *and* most dangerous courses of action. In his postwar memoirs, Gamelin even admits that "one must always plan for the worst"[61] In 1939–40, however, the French and British ignored theory and logic and instead based their plan on wishful thinking.

Gamelin personally devised the Dyle/Breda Plan based on an enemy response that illustrated what he would have done—classic mirror imaging. In this case, his lack of creativity and application of sound theory led him to be-

lieve that the main German effort would be on the "open plains" of Belgium.[62] This mirror imaging satiated his desire to secure British support and add Belgian and Dutch divisions to the Allied defense but left his forces unprepared to deal with any contingencies. Ironically, the original German plan was to advance exactly as Gamelin had anticipated, but Adolf Hitler dismissed it as another Schlieffen Plan that lacked any original thinking. In the Allied case, however, strategic leadership never questioned its creativity or potential to deceive the enemy. They all assumed, like Gamelin, that the Germans would focus their main effort in that area.

Allied tunnel vision on a single course of action planned against a single German course of action was apparent early in the planning process. In September 1939, as Gamelin pondered his Escaut Plan, he had already assumed the German weight of main effort would be across the Belgian plains. In an instruction to Georges, Gamelin was only concerned with the amount of time it would take Allied forces to reach the proper defensive line in Belgium before meeting the German main effort (*gros de l'effort*).[63] General Howard Vyse, the British director of military operations, reported that Gamelin was preoccupied with a German attack on the Low Countries, thinking that it represented an "audacious" move.[64] Ironically, Gamelin referred to a German move into Belgium as "the most dangerous" because it could have the fastest results.[65] He told Ironside in mid-September that the Germans would attack through the neutral countries, and he never wavered from this belief.[66] Preoccupation turned to negligence, as the Allies were completely unprepared to deal with the ultimate German plan to breach the Meuse in the center of the Allied line and race to the English Channel.

The only exceptions to the Allied exclusive focus on Belgium were some fleeting thoughts that the Germans could attempt to attack through Switzerland or to outflank the Maginot Line. During a meeting with his Allied counterparts on 6 October 1939, Gamelin raised the idea of a German attack through Switzerland but quickly dismissed it as unlikely.[67] Gamelin also revealed that he had considered the possibility of a German attack through Luxembourg and the Ardennes, moving southward behind the Maginot Line.[68] This potential German course of action, however, never emerged in the critical Allied discussions in November, or frankly at all. If such a contingency had been planned, and then executed as a branch plan in May 1940 once the Germans revealed their true intentions, the outcome of 1940 could have been a far cry from what occurred. At the start of the war, Gamelin had pointed out to Georges the need to maintain large strategic and tactical level reserves behind the lines.[69] In the end, it is curious that Gamelin designed his campaign plan to meet the defensive Allied strategy of a long war by stripping his reserves away from the center, where they could most easily reach any part of the defensive line and employing them in the far reaches of his left wing.

Historians have examined the role of intelligence in how it supported the Allied response to the German assault, but a study on the link between op-

erational intelligence and Allied planning in 1940 has yet to be written. Was Gamelin's fixation on a German attack in Belgium supported by intelligence? It is difficult to say. Gamelin later claimed to be ill-informed of the direction of the main German attack through the Ardennes.[70] Allied intelligence, at least at the tactical level, did report German armored columns snaking through the Ardennes in the early hours of the German assault. At the operational level of war, however, Allied intelligence focused on German capabilities, particularly the number of divisions available for combat in the West but ignored potential German intentions that could have influenced Allied planning. Ernest May, in *Strange Victory*, argues that French intelligence had uncovered many clues that pointed to an attack through the Ardennes but that nobody was able to synthesize these snippets of information in an accurate estimate of German intentions.[71]

It is also difficult to ascertain whether Gamelin would have listened to such heresy and changed his plan accordingly. Making significant modifications to all the detailed planning of meticulous timetables that focused on getting Allied forces to the proper defensive line in Belgium would have been quite challenging after months of beating the same drum over and over. The British were also in complete agreement concerning what they considered to be the German intent.[72] A sound plan, however, should always reflect any changes in the situation. As it turned out, when the situation did reveal that the Allies had erred in their assumption of the German main effort in Belgium, it was left to subordinate commanders to conduct ad hoc/crisis action planning. The German tempo, however, disrupted any potential Allied decision-making cycle to produce a coordinated response.

Stealing a phrase from Neville Chamberlain in reference to Hitler not launching an assault on the West in 1939, the Allies "missed the bus" on German intentions. In planning, an assumption is made to continue planning when something is unknown. In other words, a likely conclusion or judgment is made in the absence of facts. Much like Allied leaders assumed that the Germans would turn west after the completion of their campaign in Poland, they assumed that Germany would concentrate their forces in Belgium because the terrain was suited for the offense and that was where the Germans had been successful in the previous war. Of course, going back to the war prior to that one could have shed some light on a more dangerous possibility—that the Germans would employ superior force in the area of Sedan. Instead, Allied plans began to treat their assumption of the main German assault in Belgium as a fact instead of the grave risk that such an assumption represented.

In his postwar memoirs, Gamelin incredulously defended his plan to advance into Belgium by claiming that "staying on our border was the easy way out . . . it was, indeed, tempting."[73] In his mind, a German takeover of Belgium would physically and diplomatically sever France and Britain and put an end to a united front against the common enemy.[74] He only had to keep the Germans locked in a stalemate and eventually the British would send enough troops to

give him the force superiority that he believed was a prerequisite of success. This stubborn adherence to a preconceived idea, with little attention paid to any other contingency, played right into German hands. Avoiding defeat instead became a self-fulfilling prophecy.

Conclusion

> *In other words, the German triumph was, essentially, a triumph of intellect*.
>
> ~ Marc Bloch[75]

During the six weeks of the campaign in the west in 1940, the Allies suffered approximately 100,000 killed in action, roughly equal to the amount the United States lost in all four years of bitter fighting in the Pacific. This staggering number of dead in a relatively brief period underscores the point that Allied soldiers paid a steep price for the mistakes of their strategic civilian and military leadership. Armed with a strategy that had no clear vision of victory, Gamelin had devised a campaign plan that ignored key facets of operational art and sound planning. Many more soldiers and civilians would soon pay the price for that failure.

Lessons learned from the Allied debacle in 1940 are numerous. As far as the execution of the Allied plan, previous studies have drawn conclusions on the faulty employment of Allied armor, the exposure of the linear front concept to maneuver warfare, the methodical nature of the French employment of forces, and antiquated command and control. This study has examined the topic from a strategic and operational viewpoint prior to the actual campaign, examining the rationale and implications of a strategy that lacked a clear vision of victory, a campaign plan that did not balance operational factors, and planning that never accounted for a potential most dangerous enemy course of action. These were the critical ingredients that set up the French and British forces to fail.

The Allied long war strategy, developed at the start of the war, had no clear vision of victory. Instead, the French and British planned to defend their national interests with a war of attrition in Belgium for at least two years to build up superior combat power for offensive operations in the future. The War to End All Wars was now a blueprint for success. The Allied focus on defending Belgium operationalized the long war strategy, as it appealed to both nations' strategic concerns and addressed Gamelin's quest to even the force score with Germany. It was the *only* contingency, and with few exceptions everyone accepted it without question. More than just an example of group think, it was a plan that appeased each Allied nation's fears. The French could move the dreaded front away from French territory with the added prestige of protecting their neighbors. For the British, it meant that they could concentrate their land forces on the continent as close to Britain as possible and use their naval and air assets in support without compromising national security. But a second phase of the Allied campaign to conduct offensive operations to defeat Germany nev-

er even reached the planning stages. Gamelin had planned a half measure; it was a defensive campaign with no concept to link it to subsequent operations.

Gamelin's fixation on force, illustrated by his campaign plan to ensure the addition of British, Belgian, and Dutch forces through a concentrated move into Belgium, tipped the balance of time, space, and force and prevented the Allies from achieving their objective of defending France and the Low Countries. In support of the Allied long war strategy, Gamelin had ensured the continued support of British forces and anticipated adding Belgian and Dutch divisions to his force ledger. In doing so, he satiated his own preoccupation with force that he believed was instrumental in preventing defeat. Gamelin had concluded that this could only be done by defending Belgium and Dutch territory. He therefore positioned his best forces at the northernmost point in the defensive line, ready to prevent the Germans from exploiting the flat terrain in Belgium and outflanking them at the coast. As a result, he was unable to meet the actual German main effort in the area of Sedan, as the superior enemy concentration of force at the right time and place, coupled with a high operational tempo, shredded the Allied long war strategy and revealed Gamelin's campaign plan as a paper tiger.

The Allies had put all their effort into one plan against the most likely enemy course of action—a German advance across Belgium—accepting enormous risk in doing so. They ultimately fought the campaign that they had envisioned for months, not the campaign that the actual situation demanded. When the Germans simply focused on the weakest part of Gamelin's long-planned Allied defensive line, there was no contingency plan to meet it and no chance to reposition their best forces in time. Months of planning around a single option had led to a predictable plan, and for the Allies, a predictable result. With no serious consideration of contingency plans, Gamelin had gambled everything on a German most likely course of action that pleased his political masters but proved to be nothing more than his own wishful illusion. In the end, the Allied plan did not survive first contact with the enemy. The enemy had gotten a vote too.

Endnotes

1. Marc Bloch, *Strange Defeat: A Statement of Evidence Written in 1940* (New York: Octagon Books, 1968), 175.
2. *Joint Planning*, Joint Publication 5-0 (Washington DC: Office of the Chairman of the Joint Chiefs of Staff, 2017).
3. Milan Vego, *Joint Operational Warfare: Theory and Practice* (Newport, RI: U.S. Naval War College, 2007), ix–63.
4. Maurice Gustave Gamelin, *Servir*, 3 vols. (Paris: Plon, 1946).
5. Bloch, *Strange Defeat*.
6. See, for instance, Jacques Minart, *P.C. Vincennes, Secteur 4* (Paris: Editions Berger-Levrault, 1945); Adolphe Goutard, *The Battle of France, 1940* (New York: Ives Washburn, 1959); and Roderick Macleod and Denis Kelly, *Time Unguarded: The Ironside Diaries, 1937–1940* (New York: David McKay, 1962).
7. Classic studies of the events include Alistair Horne, *To Lose a Battle: France 1940*, rev. ed. (London: Penguin Books, 1990); and William L. Shirer, *The Collapse of the Third*

Republic: An Inquiry into the Fall of France in 1940 (New York: Simon and Schuster, 1969); and Martin S. Alexander, *The Republic in Danger: General Maurice Gamelin and the Politics of French Defence, 1933–1940* (Cambridge, UK: Cambridge University Press, 1992). More recent studies include Ernest R. May, *Strange Victory: Hitler's Conquest of France* (New York: Hill and Wang, 2000); Julian Jackson, *The Fall of France: The Nazi Invasion of 1940* (Oxford, UK: Oxford University Press, 2003); Joel Blatt, ed., *The French Defeat of 1940: Reassessments* (New York: Beghahn Books, 1998); Lloyd Clark, *Blitzkrieg: Myth, Reality, and Hitler's Lightning War, France 1940* (New York: Atlantic Monthly Press, 2016); and Philip Nord, *France 1940: Defending the Republic 1940* (New Haven, CT: Yale University Press, 2015).

8. Robert A. Doughty, *The Seeds of Disaster: The Development of French Army Doctrine, 1919–1939* (Hamden, CT: Archon Books, 1985); and Robert A. Doughty, *The Breaking Point: Sedan and the Fall of France, 1940* (Hamden, CT: Archon Books, 1990). The author argues that French emphasis on the methodical battle and centralized decision making were critical to the overall French failure in 1940.
9. Bloch, *Strange Defeat*, 39.
10. French and British enforcement of the military clauses of the Treaty of Versailles, despite German obstruction efforts, were largely successful in diminishing German military strength from 1920–31. See Richard J. Shuster, *German Disarmament of Germany after World War I: The Diplomacy of International Arms Inspection, 1920–31* (London: Routledge, 2006).
11. Alistair Horne, *To Lose a Battle: France 1940*, rev. ed. (London: Penguin Books, 1990), 169; and Robert Young, "La Guerre de Longue Dureé: Some Reflections of French Strategy and Diplomacy in the 1930s," in *General Staffs and Diplomacy Before the Second World War,* ed. Adrian Preston (London: Croom Helm, 1978), 46–49.
12. David Dilks, "The Unnecessary War?: Military Advice and Foreign Policy in Great Britain, 1931–1939," in *General Staffs and Diplomacy Before the Second World War*, 129.
13. CAB[inet papers] 80/9, "Certain Aspects of the Present Situation," 26 March 1940, Kew, United Kingdom, National Archives; and Nick Smart, *British Strategy and Politics During the Phony War: Before the Balloon Went Up* (Westport, CT: Praeger, 2003), 73.
14. Young, "La Guerre de Longue Dureé," 58.
15. [Vincennes, France, Service historique de la Défense, Centre historique des archives], GR 5N 579, Gamelin to Minister of National Defense, 4 April 1938; and GR 5N 579, Gamelin, "Note sur la Collaboration Militaire Franco-Britannique," 24 April 1938.
16. Peter Jackson and Joseph Maiolo, "Strategic Intelligence, Counter-Intelligence and Alliance Diplomacy in Anglo-French Relations Before the Second World War," *Militärgeschichtliche Zeitschrift* 65, no. 2 (2006): 453–54.
17. Up to this point, there had been only low-level contacts between service attachés. See Martin S. Alexander and William J. Philpott, ed., *Anglo-French Defence Relations Between the Wars* (New York: Palgrave Macmillan, 2002), 106, 211; Eleanor M. Gates, *The End of the Affair: The Collapse of the Anglo-French Alliance, 1939–40* (Berkeley: University of California Press, 1981), 16; L. F. Ellis, *The War in France and Flanders, 1939–1940* (London: Her Majesty's Stationary Office, 1953), 4; and Allan R. Millett and Williamson Murray, *Military Effectiveness*, vol. 2, *The Interwar Period* (Boston: Allen & Unwin, 1988), 109–10.
18. Gamelin, *Servir*, vol. 1, 145–46. His quest for military manpower played a significant role in his decision to move into Belgium and will be explored in the next section.
19. GR 5N 580, Decisions of the Supreme Council, 17 November 1939; Smart, *British Strategy and Politics During the Phony War*, 86; and Gamelin, *Servir*, vol. 3, 146.
20. W[ar]O[ffice papers] 106/1684, Director of Military Operations and Plans, Notes of a Meeting Held at the Headquarters of General Gamelin, 9 November 1939, Kew, United Kingdom, National Archives.
21. WO 106/1684, War Office to British Military Attaché, 4 September 1939.
22. Doughty, *The Seeds of Disaster*, 65–67.
23. Smart, *British Strategy and Politics During the Phony War*, 81; and Millett and Murray, *Military Effectiveness*, vol. 2, 46.

24. Roderick Macleod and Denis Kelly, *Time Unguarded: The Ironside Diaries, 1937–1940* (New York: D. McKay, 1962), 297.
25. Smart, *British Strategy and Politics During the Phony War*, 77.
26. CAB 80/104, Ironside Memo, 24 January 1940.
27. WO 106/1684, Hore-Belisha Note, 21 September 1939.
28. CAB 80/105, Chiefs of Staff Committee, German Invasion of Holland: Allied Policy Draft Memorandum, 5 April 1940; CAB 80/105, Chiefs of Staff Committee, German Invasion of Holland: Air Policy Report, 8 April 1940; and CAB 80/10, Chiefs of Staff Committee, Allied Action in the Event of a German Attack on the Netherland Islands, 1 May 1940.
29. Bloch, *Strange Defeat*, 48.
30. Alistair Horne, *To Lose a Battle: France 1940* (London: Penguin, 2007), 677 notes.
31. Karl-Heinz Frieser, *The Blitzkrieg Legend: The 1940 Campaign in the West* (Annapolis, MD: Naval Institute Press, 2005), 60–67.
32. CAB 66/1/44, Conversation with M. Daladier and General Gamelin, Note By the Secretary of State for War and Lord Hankey, 21 September 1939.
33. WO 106/1684, Notes of a Meeting Held at the Headquarters of General Gamelin, 9 November 1939; WO 106/1684, General Staff to C.I.G.S., 20 April 1940.
34. Gamelin, Note (Georges), 5 November 1939, *Servir*, vol. 3, 140–41.
35. GR 5N 580, Gamelin to Minister of National Defense, "Plan de guerre pour le printemps 1940," 26 February 1940.
36. Shuster, *German Disarmament of Germany After World War I*, 72–73.
37. Gamelin, *Servir*, vol. 1, 36.
38. See, for instance, CAB 80/3, General Gamelin's Observations, 27 September 1939; CAB 80/3, Chiefs of Staff Committee, Comments on General Gamelin's Observation on our Appreciation on the Possible Course of the War, 2 October 1939; and Gamelin to Daladier, 7 January 1940, *Servir*, vol. 3, 153–54.
39. Karl-Heinz Frieser, *The Blitzkrieg Legend: The 1940 Campaign in the West* (Annapolis, MD: Naval Institute Press, 2005), 36; and Clark, *Blitzkrieg*, 94. There is some variation among sources in the overall and individual nation numbers but there certainly was not a German force advantage in terms of number of divisions.
40. Minart, *P.C. Vincennes, Secteur 4*, Tome I, Titre II, 151–52.
41. CAB 80/104, Notes of a Meeting at the Headquarters of Gamelin, 21 September 1939.
42. CAB 66/2/50, Gamelin to Secretary of State for War, 20 October 1939.
43. CAB 80/3, French Permanent Military Representatives, Note Relative to Operations in Holland, Belgium, and Luxembourg, 24 September 1939.
44. CAB 66/2/50, 2nd Bureau, Note on the Available Personnel of Germany, 14 October 1939.
45. CAB 80/104, Ironside Memo, 24 January 1940.
46. WO 106/1684, Dewing to Pownall, 23 September 1939; WO 106/1684, British Chiefs of Staff Draft Letter to Gamelin, September 1939; and WO 106/1684, General Staff Criticism of General Gamelin's Plan, September 1939.
47. WO 106/1684, Extract of a Letter from Gamelin to CIGS, 23 September 1939.
48. Gamelin, Instruction No. 6, 29 September 1939, *Servir*, vol. 3, 83.
49. WO 106/1684, Suggestions, 19 November 1939.
50. Gamelin, Instruction No. 7 (Georges), 30 September 1939, *Servir*, vol. 3, 85.
51. Gamelin, Instruction Personnelle et Secrète No 8, 15 November 1939, *Servir*, vol. 1, 82–93.
52. Gamelin to Daladier, 7 January 1940, *Servir*, vol. 3, 153–54.
53. Gamelin, Instruction Personnelle et Secrète No 11 (for Georges), 12 March 1940, *Servir*, vol. 3, 177.
54. Gamelin to Georges, 15 April 1940, *Servir*, vol. 3, 343.
55. Vego, *Joint Operational Warfare*, iii–55.
56. Bloch, *Strange Defeat*, 115–16.
57. May, *Strange Victory*, 385.

58. L. F. Ellis, *The War in France and Flanders, 1939–1940*, 22.
59. Frieser, *The Blitzkrieg Legend*, 69.
60. Vego, *Joint Operational Warfare*, ix–45.
61. Gamelin, *Servir*, vol. 1, 346.
62. Gamelin, *Servir*, vol. 1, 336.
63. Gamelin, *Servir*, vol. 3, 83–84.
64. WO 106/1684, Notes Sent by General Howard Vyse, 12 September 1939.
65. CAB 80/104, Notes of a Meeting Between Gamelin and CIGS, 6 October 1939.
66. WO 106/1684, Gamelin to Ironside, 18 September 1939. Gamelin dated his letter 1839.
67. CAB 80/104, Notes of a Meeting Between Gamelin and CIGS, 6 October 1939.
68. WO 106/1684, CIGS to Wavell, 10 October 1939.
69. Gamelin, Note (Georges), 21 September 1939, *Servir*, vol. 1, 245–47.
70. Gamelin, *Servir*, vol. 1, 336.
71. May, *Strange Victory*, 5.
72. CAB 80/104, Ironside Memo, 24 January 1940.
73. Gamelin, *Servir*, vol. 1, 337.
74. Gamelin, *Servir*, vol. 3, 135.
75. Bloch, *Strange Defeat*, 36.