SSQ
STRATEGIC STUDIES QUARTERLY
WINTER 2020 VOL. 14, NO. 4

SSQ Strategic Studies Quarterly

https://www.af.mil/

https://www.spaceforce.mil/

https://www.aetc.af.mil/

https://www.airuniversity.af.edu/

Strategic Studies Quarterly

An Air Force–Sponsored Strategic Forum on
National and International Security

WINTER 2020 VOL. 14, NO. 4

POLICY FORUM

FEATURE ARTICLE

PERSPECTIVES

On the Future of Air and Space Power

An Interview with Barbara M. Barrett

Secretary of the Air Force

Conducted 28 September 2020

***SSQ*: Air Force recommendations in the 2021 National Defense Authorization Act (NDAA) focus on divestiture. Was the outcome aggressive or ambitious enough?**

SECAF: Overall, yes. The FY21 president's budget reflects balance. Prior to submission, the Department of the Air Force conducted a comprehensive review of each portfolio and made hard decisions to align with the National Defense Strategy (NDS). Planners wargamed complex scenarios to assess effectiveness and recommended required investments to prevail in a peer fight. In the budget, the Department accepts calculated, modest, short-term risk to achieve the Secretary of Defense's goal of irreversible momentum toward NDS implementation.

***SSQ*: Are you comfortable with the level of R&D funding within the 2021 NDAA?**

SECAF: I think it is reasonable. R&D investment is fundamental to enabling the Department of the Air Force to fulfill the NDS. The FY21 R&D budget includes robust support of the Advanced Battle Management System, the Next-Generation Overhead Persistent Infrared constellation, Vanguard programs, quantum science, advanced communications, directed energy, 5G, microelectronics, and hypersonics. This funding will ensure the Air and Space Forces stay competitive.

***SSQ*: Based on your vast experience in the defense industry, how would you characterize the state of Air Force–industry relations?**

SECAF: The Air Force and Space Force have strong relationships with industry. Still, the COVID-19 pandemic instigated manufacturing challenges that no one anticipated. Concurrently, we jump-started innovation to accomplish missions. Our industry partners often delivered military capabilities despite the challenges encountered this year.

***SSQ*: Do you have any concerns about the defense industrial base and its ability to support our National Defense Strategy?**

SECAF: To meet the demands of the NDS, the Air Force and Space Force depend on reliable, responsive industry that produces and sustains systems. The pandemic has stressed many elements of the defense industry; however, many partners found ingenious ways to mitigate stresses and risks. Importantly, we are accelerating the shift to digital engineering, agile software development, and open systems architectures. Additionally, we are working with the industrial base to reduce dependence on foreign sources of raw materials and microelectronics. Finally, to broaden our collaboration with the defense industry, we are building new partnerships with companies that have never before worked with the Department of the Air Force.

***SSQ*: Are the services organized, trained, and equipped adequately to support the era of great power conflict, particularly the rebalance to Asia?**

SECAF: The NDS calls on the Department of Defense to deter and, if deterrence fails, to defend against adversarial behavior. With the support of Congress, the force has grown over the past three budget years by 7,820 Airmen. Nonetheless, three congressionally mandated reports—one internal and two conducted independently—conclude that the Air Force is too small to meet all the demands of the NDS. The FY21 president's budget will help by adding 1,500 personnel to the F-35, refueling operations, maintenance, and combat support.

***SSQ*: Given the state of artificial intelligence (AI), autonomous weapons, and nanotechnology, what do you see as the greatest opportunities?**

SECAF: Technology presents impactful opportunities for faster, cheaper, and more efficient capabilities. Artificial intelligence, autonomous weapons, and nanotechnology each offer greater flexibility and lethality to Air Force and Space Force operations. Our Vanguard programs use prototyping and experimentation to improve weapons systems and warfighting concepts.

Three Vanguard programs—Skyborg, Golden Horde, and Navigation Technology Satellite-3 (NTS-3)—push boundaries by integrating technology components to deliver new capabilities across multiple domains. The Skyborg initiative integrates AI with autonomous unmanned aerial vehicles to enable manned-unmanned teaming. Golden Horde uses onboard radios to develop networked and collaborative weapons that share

data, interact, and execute coordinated actions with other systems (including manned aircraft). The NTS-3 flight experiment examines field capabilities across the ground, space, and user segments to enhance space-based positioning, navigation, and timing.

Finally, the Air Force and Space Force are leveraging domestic and international partnerships with industry, academia, allies, and partners to advance technological advantages.

***SSQ*: What challenges remain to fully incorporate the U.S. Space Force?**

SECAF: The U.S. Space Force is a good-news story. Logically, space capabilities should be aligned under Space Force. Senior leaders throughout the Department of Defense are analyzing which missions and personnel should transfer into the Space Force to align the organizations with space missions. Simultaneously, we are adapting myriad existing departmental systems to the new service. We are committed to building an agile, innovative, and bold Space Force.

***SSQ*: Where do you see the Space Force 25 years from now?**

SECAF: Twenty-five years from now, systems and technologies employed in space will have advanced unrecognizably. Still, the fundamental mission of Space Force will endure: protecting American interests and preserving free access to, and use of, the space domain for all benevolent actors.

SSQ: Madam Secretary, on behalf of Team *SSQ* and the entire *SSQ* audience, thank you for sharing your views on the bright future of the U.S. Air and Space Forces. SSQ

Space: New Threats, New Service, New Frontier

An Interview with Mir Sadat

Former Policy Director, National Security Council

Conducted 1 August 2020

Dr. Mir Sadat has over 25 years of leadership experience in private industry, the Department of Defense, the intelligence community, and the National Security Council (NSC). At the NSC he was a policy director for interagency collaboration on defense and space policy issues helping to establish the US Space Force and US Space Command. Previously, as a naval officer with intelligence and space expertise, he served on orders as a space policy strategist for the chief of naval operations and as a space operations officer for US Fleet Cyber Command / US Tenth Fleet. Prior to government service, he spent 10 years working for prime defense corporations.

***SSQ*: How would you characterize the great power competition in space?**

MS: The Cold War may be over, but since the early 2010s, a renewed era of great power competition has emerged across the world's land, air, sea, cyber, and space domains. This great power competition is about not only geostrategic positioning but ideological, political, economic, military, and technological dominance. Too often, we have leveraged only a few of our available tools (e.g., the military, diplomacy, or economics). However, the current competition requires the full employment of America's traditional and emerging instruments of national power (i.e., diplomacy, information, the military, economics, finance, intelligence, the law, and science and technology [S&T]). Our global competitors are energized by assessments that the US may be overwhelmed with domestic issues. They suggest that the US is a spent great power in decline. For the last 20 years, America has been laser-focused and resourced on countering terrorism, a real but not existential threat. The US has not resourced with maximum return on investment for great power competition, which does in fact impact our way of life. Worse yet, the US still operates in an Industrial Age mode of operation rather than in an Information or Digital Age. We must change the way we invest in and employ cutting-edge technologies or risk adverse effects to our operations in future conflicts. If we do not take significant measures, we will lose our scientific and technological competitive

advantage in less than a decade. If we lose that competitive advantage, we may be incapable of deterring other great powers or perhaps even regional hegemons. Inaction would also increase our margin of error in assessing our adversaries' intentions and capabilities, resulting in higher risk tradeoffs. Given China's technological investment and our unchanged steady-state planned force structure and budget, such a miscalculation could potentially lead to conflict between two, three, or even more nuclear powers.

Nowhere else is this competition more nebulous and strategic than in space. The US, along with its allies and partners, have recognized space as a war-fighting domain primarily in response to Russian and Chinese counterspace capabilities, military operations, and declarative statements. The stakes are high because there is a race for dominance over cislunar access, operations, and resources.

Since our global competitors and adversaries are dangerously competent and capable of threatening our space equities, a recurring theme in US policy is "maintaining and advancing United States dominance and strategic leadership in space." That is why the bipartisan 2020 National Defense Authorization Act created the US Space Force (USSF), under the Department of the Air Force, to secure our national interests in an increasingly contested domain. The competition is as much about economics, and the other instruments of national power, as it is about military power. Through the US Space Command, the US Space Force will play an integral role in America's competition for leadership in space—whether military, commercial, or civil.

A decade ago, China laid out a 30-year cislunar economic and industrial plan committing vast resources and talent to achieve its "space dream" of becoming a leading global space power. The Chinese government has funded its commercial sector and advanced its customer base via the Belt and Road Initiative at a scale and price point that market-driven firms in the United States cannot match. In fact, China's Belt and Road Initiative Space Information Corridor and Digital Silk Road will supposedly generate $10 trillion by 2050—dwarfing America's estimated space economy of $1.5 trillion by 2040 (pre-COVID-19 estimation) from today's approximately $385 billion.

There is now a gold rush in space because trillions of dollars of economic activity are moving into low Earth orbit and beyond. Our efforts should not focus on preventing China and Russia from participating in this arena if they are engaged in peaceful space activities that follow accepted rules, norms, and behavior. To compete with China, the US cannot

become China, so we must play to our strengths to retain our global competitive advantage.

US strength lies in its position as leader of the world in technological innovation, vibrancy of a true market economy, and, most importantly, democratic norms and values. China attempts to undermine America's traditional leadership role and create schisms between ally and partner spacefaring nations and the US. The US must provide allies and partners—and other nations that view the US as leader of the free and open world—with competitive military, civil, and commercial partnership frameworks. Our example and lead must be so profound that great powers and other nations would have no choice but to follow and replicate our success—although there is no guarantee.

***SSQ*: What is the significance of increased civilian space activities to national security?**

MS: The "NewSpace" sector of private industry has been funded primarily by visionary billionaires with rockets and public R&D. Space entrepreneurs and industrialists are creating new technologies and adapting current innovative technologies for space application. Their efforts are fueled by the decreasing cost of space access and innovative advances in space-enabling technologies. This environment creates the opportunity for an expanded space industrial base beyond the big aerospace companies that have traditionally supported government space missions. These NewSpace entrants are a fast-growing segment of the US space industrial base.

During the last five years, 11 billion dollars of private capital have been invested in NewSpace. However, this model is unsustainable since the COVID-19 pandemic has negatively impacted the entire US space industry. Investments in space-based companies in the second quarter of 2020 were down 23 percent from the record highs hit in 2019, and investments fell 85 percent in the second quarter of 2020 from the first quarter. The US government may also experience near- to medium-term fiscal constraints. Public financing for research and development was already at a historic low even prior to COVID-19.

There is growing recognition by Congress, the White House, the DOD, and NASA that the only long-term path to economic and strategic leadership in space is to catalyze and enable the accelerated growth of a vibrant US space industry. To maintain our lead in space, we must foster a stronger public-private partnership, and our government must resume the sustainable and impactful past levels of support for basic research while also ensuring the empowerment of diverse representation in the space industry.

Without government support, the US would have been unable to maintain its innovation and technological lead over the rest of the world in previous key commercial industries.

Strengthening the US commercial space industry is important to civil space priorities. The civil sector led by NASA is also fundamental to America's national security, as exemplified in the recent NASA Artemis Accords regarding conduct on the Moon and the 6 April 2020 executive order on space resources. NASA is on an ambitious critical path for a return to the Moon by 2024 and the development of the capabilities and infrastructure for a sustained lunar presence as a staging area before the mission to Mars and beyond. While a lunar landing is important, more critical are the readiness and capability to permanently stay on the Moon and to develop the means to get to Mars. NASA and the DOD should provide more precise assessments as to when they expect human settlements on the Moon. Those timelines should become the goals and drive subsequent decisions. This anticipated increase in human visitation and eventual settlement continues both technological and exploration leadership with applications for our military. As such, these efforts bear directly on our national security.

***SSQ*: Recently, it was announced that the 2010 US National Space Policy is being updated. What changes are most needed?**

MS: We need to normalize the space domain just like the other mentioned domains. To do so, we need to think about commerce, civil exploration, and conflict in space with some creativity. Policies need to be addressed within the context of space over the next five to 10 years. The rising economic benefits of space and its increasing importance to national security, along with advances in fundamental technologies, are all intervening factors. These factors will accelerate space activities and improve capabilities of not only traditional great powers such as China and Russia but also other spacefaring nations.

Therefore, our new National Space Policy should include or consider the following (not listed in order of importance):

- Declaring space a zone for economic ventures and civil exploration because emerging commercial ventures and the development of smallsats, cubesats, and satellite constellations are outpacing efforts to develop and implement policies and processes to address these activities;
- Establishing space sustainability, norms of behavior, and codes of conduct;

- Designating space as a critical infrastructure;
- Standardizing space cybersecurity and transmission security;
- Sharing responsibly across the spectrum band;
- Reviewing the overclassification of compartmented and special access programs to allow for greater participation of people with a need to know and not to keep everything black where it serves no deterrent value to foreign adversaries;
- Messaging strategically and publicly to allies, partners, and adversaries;
- Incorporating offensive operations in space in addition to existing defensive operations;
- Advancing solar- and nuclear-powered space propulsion as well as lunar power generation;
- Encouraging US persons to enter and graduate vocational and academic science, technology, engineering, and mathematics (STEM) programs;
- Promoting supply chain hygiene with front-of-the-line contract passes for supply chain illumination;
- Aligning counterintelligence and counterespionage in our laboratories and space industrial base, and also educating participants about potential threats;
- Increasing export-control information sharing across the government for expedient dual-use technological transfers and national security;
- Leveraging US economic offensive and defensive tools to increase American commercial space activities and support the growth of American space companies across the wide spectrum of the domestic space market and international ventures;
- Reforming government procurement and planning to send predictable signals to private space companies;
- Bolstering existing space equities exchanges, creating an eventual separate and unique space commodities exchange along with bond market utilization; and
- Increasing public financing for S&T and research and development (R&D) programs.

We must advance space policy to profoundly benefit life on Earth and for US permanent presence in cislunar and beyond.

***SSQ*: What are your thoughts on the recently released Defense Space Strategy?**

MS: The release of the *2020 Defense Space Strategy* (DSS) is an excellent step forward. The DSS claims to be a strategy for the next 10 years. Within that context, my main concern is how it implicitly perpetuates the notion that space is a domain in which conflict would not occur first. For example, stating that a primary DOD effort is to enable the US to be "capable of winning wars that extend into space" negates the DSS threat section, which affirms that space is a separate warfare domain in which conflict could potentially occur first.

The DSS call for space superiority is reminiscent of space as a sanctuary. Being superior in space vice supreme or dominant does not sufficiently empower us to fully compete with Russia and China. The DSS could have elaborated on the DOD or USSF role in maintaining freedom of space commerce and civil exploration.

The DSS mentions integration of military space power into defense operations. The DSS could have expanded space power beyond only the military and called for the need of a national-level plan emphasizing a whole-of-government space power. Foreign adversaries and US global competitors have integrated their military and national security space entities across their respective governments and even their industry. Now, they are building global partnerships. I would have used the term "integrate" vice "cooperate" in outlining the DSS's fourth line of effort referencing the DOD's relationship to other US government departments and agencies, industry, and US allies and partners.

This DSS is optimal if nothing changes over the next 10 years, and some may think that 10 years is a long time away. However, 2030 will come quickly; much can happen in this span. China sent its first astronaut into orbit in 2003 and by 2018 conducted more space-oriented operations than any other country. Now, it has already declared its intentions for the next 30 years, which will pass in the blink of an eye.

Whether the DSS or another strategy, it should clearly inform our allies and adversaries of our ambitions and intentions. The argument that ambiguity creates flexibility is nonsense when we generalize and make things so nebulous in our policies and strategies that even our closest friends are left baffled. If we do not convey that story explicitly, we are bound to repeat the mistakes of the past and potentially head into conflict.

We should also not classify our general national vision, policies, and overall strategies. We should classify only space operations; tactics, techniques, and procedures; and some of the related S&T/R&D aspects. We

must also bring everyone on the blue team into the same conversation by allowing them into special access programs. How can we prepare for a defense or offense when policy makers, decision-makers, operators, and analysts cannot talk freely to each other?

***SSQ*: What areas or space capability does the US need to be most focused on now?**

MS: Space is more than a war-fighting domain. With each passing second of Planck time, space more and more facilitates our modern way of life: it provides instantaneous global imagery, assures telecommunications, captures humanity's imagination for civil space exploration, and is a burgeoning zone for commercial ventures and investors. American commercial and civil space priorities in space are fundamental to US national security interests. Protecting those activities starting at 100 km from Earth and ranging into deep space fall under the US Space Command's area of operations (AO).

The US needs more than to look down from space to assure support to terrestrial activities. As such, US Space Command must exercise command of its AO by updating the unified command plan for expanded presence to cislunar and to map that operational environment. US Space Command will draw its personnel primarily from the Space Force, which will need to recruit, train, develop doctrine for, and equip that future force and evolving mission.

That future force and evolving mission must have more than just a terrestrial focus. The Space Force may evolve to ensure freedom of US space commerce and civil exploration just as the US Navy stands watch to ensure that the US can freely navigate the world's oceans for sea commerce and exploration. America must have space domain supremacy to ensure unfettered access to, and the freedom to operate in, space. The 2017 *National Security Strategy* (NSS) considers such space access to be a "vital interest," that is, something for which nations have fought over.

To execute this strategy, the US needs to move from the strategic defensive and start planning for the strategic offensive in space. We need to evolve the thinking from defense only to also offense because, in space, first-move advantages have more strategic implications than in the other domains. To align with the 2017 NSS, we should not settle for dominating an adversary at only a specific time and place but strive for domain supremacy, targeting an aggressor whenever we consider "freedom of operation" a vital national interest. For example, the US Navy would never settle

for just a superior naval force. It aspires to sea supremacy and domination of adversaries at any time and location.

We need to evolve our thinking, and both our lexicon and actions must match that thinking. To accomplish this paradigm shift, we need to develop something similar to the infantry assault maxim of "move, shoot, and communicate." In the context of space, moving entails a rapid launch capability to get to space no matter the weather, time, or other impediments. Just as in air operations, this precept would be a game changer because maneuver in, to, and from space is by far the most important element. Offensive action (shooting), if necessary, is next. Finally, communicating effectively is essential to taking advantage of move and shoot. You may not lose if you have a good defense, but to win you need to go on the offensive. And accomplishing any of these objectives requires a space doctrine that sets the strategic context for the Space Force and connects space power to commercial space interests and the cislunar operating environment.

***SSQ*: Do we have too many space-related agencies, such as the Missile Defense Agency (MDA), Space and Missile Systems Center (SMC), Space Development Agency (SDA), and National Reconnaissance Office (NRO)?**

MS: The MDA and SMC have purview beyond the US Space Force because national missile defense and ICBMs were purposely not integrated into the Space Force. The technologies of ICBMs and space launch are operationally different. ICBMs are needed for nuclear deterrence and not necessarily war fighting in space or supporting combatant commands for a conventional conflict. Another argument against merging the ICBM mission into the Space Force is the incompatibility of an ICBM compliance culture with space innovation culture. More evidence is needed to convince opponents that the Space Force could successfully balance ICBM compliance while encouraging space innovation. Some have also argued that if ICBMs are integrated into the Space Force, its focus will always be grounded to the terrestrial theatre. When these concerns are addressed, then separate organizations would perhaps no longer be justified.

The SDA will eventually get incorporated into the Space Force by October 2022. It would be a great outcome if the SDA were first permitted to finalize acquisition of its proliferated low-Earth-orbit architecture. Then the SDA could serve as the ideal model for most or even all Space Force acquisition. The SDA should be afforded the opportunity to succeed before absorption into the Space Force, and if it fails, then absorption allows it to start over with many lessons learned.

As far as the NRO, it may make good sense at some point to incorporate it into the Space Force. Perhaps it would be logical in the form of a dual-hatted Space Force chief of intelligence, surveillance, and reconnaissance (ISR). The topic of NRO more than the other agencies will likely be litigated into the foreseeable future before we see any resolution.

***SSQ*: If you could design a space force, would it look different than today's arrangement? Any advice on this for the new USSF chief?**

MS: General Jay Raymond, Space Force's inaugural chief of space operations, has done a fabulous job considering that he is dual-hatted as a service chief and a combatant commander (US Space Command). I am encouraged by his recent comment that it is important to solicit diverse insights and evaluate their feasibility because America's future in space is a US national interest. The US Space Force should always reflect American societal values, norms, and demography. Everyone wants to be part of a winning team; therefore, the USSF should give all its members something that they can champion.

The active duty component should focus on current operations, space domain awareness, war fighting, space supremacy, and building an international space alliance with nations that share our norms, values, and behavior. In addition to supporting the active duty component, the Reserve component should focus strategically on integrating commercial advances into the Space Force. The Space National Guard should focus on space defense of the homeland, broad-spectrum space integration for states, critical infrastructure, and defense operations from space.

The Space Force should be a cultural blend of all military and space organizations, even embracing some science fiction, to incorporate the best traditions, ranks, and symbols and to create newer ones unique to space. It is very important to consider the future mission of the Space Force between 2060 and 2070, which would resemble an oceanic force. Under no circumstances should the creators of the space culture consist only of, or be dominated by, current or prior Air Force personnel now that we also have Army and Navy personnel detailed to support the Space Force. The next step is to detail, assign, or transfer Army and Navy flag officers with space expertise to ensure diversity of thought and experience as well as to encourage and mentor transfers from their services. Otherwise, we risk creating an Air Force–lite organization that can be folded back as a separate branch of the USAF.

It is also important to match actions with words. It does not suffice to state only that the Space Force is a high-tech, future-looking service when

there are not going to be programmatic transfers from the Air Force or major investments to keep the service's technology and systems top notch. For example, the X-37B and similar programs need to transition now from the USAF to the USSF.

The Space Force would also not foster a healthy culture if its members are considered elite but others, like those in the intelligence community, play second fiddle to operations people. Every military service has its own separate intelligence center to look after its priorities, mission, and overall domain awareness. The Space Force should be no exception. Arguments against reorganizing space intelligence organizations within the Department of the Air Force should not be about major cost increases or damaging the USAF: it is simply a reassignment of personnel and resources.

Furthermore, the service should create and cultivate a clear war-fighting structure that includes all to, from, in, and through space warfare elements, including terrestrial strike, planetary defense, and space supremacy. It should also craft a unique organizational structure that blends acquisition, engineering, operations, and support at the lowest possible level without favoring a specific career field. The Space Force should have its own maintenance, legal professionals, public affairs, legislative liaison, ISR, labs, recruitment, and other critical service functions.

Every military service also has a career designator of astronaut. Space Force, as the specific military service dedicated to space, does not—even though its first recruitment video says "maybe your purpose on this planet isn't on this planet" and the second features an astronaut. This discrepancy needs to be resolved by permitting other services' astronauts to transfer to Space Force as astronauts and allowing new military recruits to the astronaut career field in the Space Force. Doing so is just one other measure that would permit the Space Force not to be grounded.

These astronauts would also be the connective tissues to build stronger ties with NASA and the private sector because the Space Force will eventually grow to ensure access, operations, and safety of both commercial and civilian space. The earlier that Space Force leadership embraces and supports this momentum, the further ahead we will be in the space competition.

***SSQ*: Looking to the future, what is your sense of our strengths, weaknesses, opportunities, and threats in space 20 years from now?**

MS: In 2019, US Air Force Space Command assessed that by 2060 space will be "a significant engine of national political, economic, and military power" and that the United States "must commit to having a military force structure that can defend this international space order and

defend American space interests, to include American space settlements and commerce."

When we endeavored to put the first two humans on the Moon, we did not do it by cooperating only with the government and industry. We did so by integrating a whole-of-nation approach. The US must create and execute an integrated, comprehensive 2060 American Space Vision and Strategy that fuses national security, civil, and commercial space efforts using to the fullest extent possible all national instruments of power, as mentioned earlier. Integration must not be an end state but a means to assimilate and economize to scale our shared technologies, talents, investments, and innovative discoveries. The US should develop a guiding 2060 American Space Vision to catalyze whole-of-nation efforts and enable the United States to compete and win now and into the future. This vision should be developed to drive a host of actions specific to federal departments and agencies and to update other strategies and policies.

The United States can either prepare and posture to shape a future with American strategic leadership in space or resign itself to follower status—leaving leaders and citizens to ask themselves why we never made the necessary reforms. We can either seize the moment or waste this decade's opportunities for US strategic leadership in space. We cannot achieve this vision by investing only in technology. We must invest in human capital to win in this great power competition.

America's greatest assets are its people's knowledge, innovation, and resolve. Without Americans and their innovative talents, no amount of resources or technological capabilities can ensure that the US will last as a great power or win in great power competition. We must empower Americans to attain the necessary twenty-first-century skill sets for the future economy.

There is no denying that we have a shortage of STEM vocational and educational graduates in the US. The space industrial base and government space organizations compete with each other and with other cutting-edge technology sectors for recruitment of talent. So government and industry need to work together to fix this labor and talent shortage—not just for the space industry but all STEM-dependent sectors.

Space currently provides value because it facilitates the creation, distribution, and selling of data. But in the future, space will become increasingly commercialized and industrialized, which will demand highly skilled human capital. NASA's Artemis program will require an additional 10,000 STEM graduates over the next five years for civil needs alone, and this

does not account for what is needed to support the evolving US Space Force or the enlarged space industry.

Current STEM personnel numbers are insufficient unless we do something to meet the needs of expanded national space capabilities and the industrial base that provides those capabilities. The space industry will also require non-STEM personnel knowledgeable of the space enterprise in a variety of support occupational fields, such as financial engineering, economics, and law. We require a whole-of-government mobilization, especially in light of our STEM statistics as compared to our great power competitors, if we are intent on sourcing those talents.

STEM is a vital innovation multiplier. We must ensure our future generations are afforded access to quality education and training programs—especially in STEM and STEM-related fields. If our future generations don't have this background, then our nation will incur qualitative and quantitative loss in many arenas. We will not have properly trained and educated "women and men of the hour" making sound decisions about our civil, commercial, and national security priorities.

***SSQ*: Dr. Sadat, on behalf of Team SSQ and the entire SSQ audience, thank you for sharing your profound ideas on the future of the US Space Force.**

MS: Thank you for taking an interest in discussing and debating critical space topics facing our nation and allies. I look forward to your readers' reactions and continuing our dialogue. Most importantly, thanks to Mike Guillot for extending an opportunity for me to share my perspective. He deserves our gratitude for his four decades of military and civilian service to our nation. SSQ

FEATURE ARTICLE

Poison, Persistence, and Cascade Effects: AI and Cyber Conflict

Christopher Whyte

Abstract

Few developments seem as poised to alter the characteristics of security in the digital age as the advent of artificial intelligence (AI) technologies. For national defense establishments, the emergence of nefarious AI techniques is particularly worrisome, not least because prototype applications already exist. Cyber attacks augmented by AI portend tailoring and manipulating the human side of important societal systems as well as introducing the risk that comes from moving technical skill from the hacker to an algorithm. The rise of AI-augmented cyber defenses incorporated into national defense postures will likely be vulnerable to "poisoning" attacks that predict, manipulate, and subvert the functionality of defensive algorithms. These AI-enabled cyber campaigns contain great potential for operational obfuscation and strategic misdirection. At the operational level, piggybacking onto routine activities to evade security protocols adds uncertainty, complicating cyber defense particularly where adversarial learning tools are employed offensively. Strategically, AI-enabled cyber operations may be able to pursue conflict outcomes beyond those expected of adversaries. Perhaps more worrisome is that the centrality of the Internet to new AI systems incorporated across all areas of national security—not just to cyber conflict processes—indicates that sophisticated adversaries may be motivated to launch offensive online actions to achieve effects in other domains with some increasing regularity.

*****

In recent decades, few technological developments have captured the attention and sparked the concern of national publics so much as those linked to artificial intelligence (AI). This might seem a remarkable and outlandish statement given that, if prompted, the average consumer would likely be unable to identify that AI sits at the heart of everyday commercial services like Google's search engine or Amazon's marketplace. Nevertheless, the subject of AI has, since at least 2017, been at the heart of prominent conversations about the future of human innovation and the changing shape of societal security.[1] Tech luminaries con-

tinue to expound the revolutionary potential of new machine learning and reasoning techniques that now easily solve endemic issues of overcomplexity that plague the conventional design and operation of digital systems. At the same time, leading voices from Elon Musk to Max Tegmark and Steve Wozniak increasingly refuse to disagree with doomsayers who claim that AI might, if mismanaged, lead to societal disaster.[2] Indeed, some are so concerned that they lean heavily into threat inflation, using extreme examples in an attempt to convince audiences of the stakes involved in getting AI "right."[3]

Around the world, few entities are as focused on the impact AI systems portend for security as are national militaries. In the United States, political and military leaders have variously called for a "Third Offset" that leverages smart machine systems to outpace the capabilities of foreign adversaries in years to come.[4] Indeed, official strategy documents and formal statements maintain something military practitioners and scholars generally take years to realize—that a new technology is changing the character of warfare itself.[5] The resultant expectation, according to some, is that underlying AI processes will lead to an inevitable transformation in the bases of national power and alter security relationships between states in both strategic and operational terms. While there is a small but growing body of work on the potential of AI to affect military and national power writ large, surprisingly few reports attempt to discuss AI developments in the context of state competition online.[6] Moreover, what work does exist tends to involve only descriptive analyses of threat scenarios, without considering how AI's augmentation of cyber capabilities—specifically the application of machine learning techniques to offense and defense—alters the dynamics of strategic engagement in the digital domain.[7]

AI-driven cyber attacks differ dramatically from the more conventional digital threats that have occupied practitioners and researchers for the past three decades. Their effects are also possible—even likely—to be felt outside of cyberspace. However, the centrality of cyberspace to the deployment and operation of soon-to-be ubiquitous AI systems implies new motivations for operations within the cyber domain. The prospect of offensive and defensive cyber operations upgraded by AI challenges several assumptions held by current strategies for cyber conflict prevention and should be a cause of significant concern for policy makers. AI is likely to alter the shape and strategic calculations bound up in interstate cyber conflict and alter the dynamics of interstate cyber conflict processes.[8] However, such transformation will not come simply from the sophistication of attack and defense by AI, but rather from the manner in which AI adds

new complexity and therein intensifies issues of strategic perception and misperception.

Ultimately, AI does not itself imply inevitable advantages for attackers over defenders (or vice versa). But adversarial learning techniques layer complexity on top of already complex operational conditions in cyberspace and may contribute to an uptick in offensive behavior. After all, nested logics of engagement across a heterogeneous global environment make for an even more convoluted battlespace than exists presently. Of greatest concern, however, is the centrality of the Internet to new AI systems that will be incorporated across all areas of national security, not just in cyber conflict processes. This inevitable application of new techniques and technologies across the national defense enterprise suggests that sophisticated adversaries may be motivated to launch offensive online actions to achieve effects in other domains with some increasing regularity. This introduces new challenges for defense at scale and amplifies some risks of AI-enabled engagements, such as the possibility of AI-driven "flash crashes."

This article takes steps to reconcile the task of defining artificial intelligence as it relates to cyber operations by highlighting how the major relevant area of AI development, machine learning, promises to affect many of the assumptions about operating in cyberspace that have been considered standard among security practitioners and researchers for some years.[9] Then, it categorizes the primary advances in AI technologies likely to augment offensive cyber operations, including the shape of cyber activities designed to target AI systems. Finally, the article frames the implications for deterrence in cyberspace by referring to the policy of persistent engagement (PE), agreed competition, and forward defense promulgated in 2018 by the United States.

Before moving forward, one clarification seems worthy of mention. This article is structured around a discussion of the utility of AI learning techniques for cyber offense. It does so as a basis for discussing the totality of strategic cyber considerations pertaining to AI. As implied above, however, it does not fundamentally argue that AI systematically favors the offense as some international relations scholars argue.[10] While new adversarial learning techniques do seem poised to enhance the attacker's toolkit over and above that of the defender, the logic of offense dominance with AI likely mirrors that of cyber operations: offense is dominant and tactical deterrence impossibly hard only where the value of target systems is high. Otherwise, AI stands to favor the defense as much as the offense, at least at the tactical level. This parity of effect may not bear out in the realm of

strategic interaction where new learning capabilities employed at scale in routine operations add complexity to the already murky perspective of operators who must consider interacting operational, institutional, and geopolitical contexts.

Artificial Intelligence and Assumptions on Cyber Operations

The label "artificial intelligence" denotes a basket of technologies whose common attribute is the capability (or a set of capabilities) to simulate human cognition, particularly the ability of the human brain to adaptively reason, learn, and autonomously undertake appropriate actions in response to a given environment.[11] In an even broader sense than is the case with all things cyber, AI encompasses an immensely diverse landscape of technologies and areas of scientific development, from computer science to mathematics and neuroscience. As such, using AI as a descriptor in many studies to describe new capabilities invariably risks, at least on some level, misleading readers by implying that AI is best thought of as a relatively monolithic underlying technology whose design features will define future conflict. The implications of AI are best thought of in terms of unique interactions that will inevitably occur as an incredible array of potential smart machine systems are plugged into extant societal processes. The challenge is to contextualize the diverse forms of what many generically refer to as AI and consider the implications of new techniques on the conduct of cyber conflict.

Machines that Reason, Learn, and Act Autonomously

Machine cognition, which today substantially enables the function of most industrial sectors in advanced economies, has been a topic of significant interest to scientists and philosophers for the better part of two centuries. From Charles Babbage and Ada Lovelace to Alan Turing, many of the greatest minds of the post–Industrial Revolution era have made their names by advancing societal thinking on the possibility of machines that mimic how humans behave, move, and think.[12] More recently, the modern field of *artificial intelligence*—a term that emerged only in the latter half of the twentieth century among cybernetics and computer engineering researchers—has its roots as a discipline in the substantial postwar work of AI pioneers like Marvin Minsky, Norbert Wiener, and John von Neumann.[13] They asked if, given the context of recent advances in computing, a machine might be made that could realistically simulate the higher functions of the human mind.[14] For such

researchers, the challenge of machine intelligence lay in moving beyond the mere programmability of emerging computer constructs to build complex thinking systems capable of concept formation, environment recognition, abstract reasoning, and self-improvement.[15] Such systems are now commonplace in application to narrowly defined societal functions. Moreover, competing schools of thought variously hold—for mathematical, neurological, evolutionary, or computational reasons—that the future will see general learners whose ability to autonomously operate in the world matches and surpasses that of humans.

Today, AI applied broadly across areas of global society is what researchers label "narrow" AI—not the "general" systems that are the focus of science fiction classics like *The Terminator* or *I, Robot*, but limited applications of machine intelligence to discrete tasks.[16] Generally, though there is some crossover and meaningful within-category differentiation, the technologies of AI might be thought of as existing across three main categories—(1) sensing and perception, (2) movement, and (3) machine reasoning and learning.[17] Of these, by far the one most arguably synonymous with AI as it is often portrayed in popular settings is the last. In this category is a range of advances that encompass machine abilities to interpret data, represent knowledge, and understand information imbued with social meaning. By far the most significant area in this category is machine learning, the scientific study and development of approaches to pattern recognition and knowledge construction absent preprogrammed instructions on how to interpret data.[18] Machine learning is relatively simple to understand. We might think of conventional computing as involving the input of data to a (non-learning) algorithm that then outputs some functional result, such as a statistic or perhaps a graphical representation of the data. By contrast, machine learning involves the input of both data and a desired result to an algorithm (often called a "learner") that infers, learns about a given issue represented in the data, and then outputs another algorithm tailored to allow for intelligent engagement.[19] In short, today's sophisticated AI techniques do not overwhelm computational challenges via the application of processing power so much as they more effectively study data to design a better process. In this way, AI promises to solve a traditional challenge in continuing to realize the promise of computers for human society. Specifically, the development of complex software to run on increasingly sophisticated systems means ever-growing demands on computer memory (both in storage and processing terms) and manifestation of human error in programming at scale. Machine learning does not compensate by building a better computer or by just catching those errors more efficiently. Rather,

it does so by allowing computers to sidestep such issues entirely by programming and reprogramming themselves more efficiently.

While machine learning involves those new processes and techniques for the direct mimicry of human cognition, the first two categories above—sensing and perception and movement—include the technologies needed to allow machines to effectively move beyond internal process to survey and operate within an environment. To some degree, of course, better sensing and perception are part and parcel of building better machine reasoning and learning algorithms. After all, effective mimicry of human cognition requires that such algorithms are able to interpret data and make inferences as a human might.[20] This involves an ability to consider language usage as a human might—that is, more effective natural language processing (NLP)—and a capability to construct and represent knowledge via ontological treatment.[21] Thus, learner algorithms can move beyond simplistic statistical treatment of input data to identify concepts and connections that are sociological in nature.

Beyond the syntactic foundations of such advances in perception, however, much AI involves the development of new sensor systems that create data for algorithms to consume. Advances in camera systems and microwave sensors that allow for sophisticated text and imagery recognition via visual feeds, for instance, are critical to the function of new software that helps law enforcement more rapidly assess patterns in criminal behavior or traffic flow. At the same time, AI involves the construction of robotic systems that can more effectively gather data and act as autonomous agents with the help of advanced learning software.[22]

Expected Advances in AI-Enabled Cyber Offense

How might artificial intelligence augment or upgrade offensive cyber operations (OCO)? The conventional answer to such a question is simply that AI (specifically, machine learning) stands to (1) make cyber attacks more insidious, disruptive, and long-lasting; (2) reduce the effectiveness of conventional defensive measures; and (3) make powerful attacks more accessible for the median malicious online actor. Thus, AI portends unprecedented adaptability, rapidity, and opportunity for unexpected malicious behavior than has previously been the case. Four prospective dynamics surrounding AI-enabled cyber offense seem worthy of note.

Attack Surface Analysis at Scale and Speed

AI programming portends a heightened threat to prospective cyber-attack victims insofar as it enables analysis of the attack surface of targeted systems and victim entities at scale.[23] This manifests at two levels. The first is the opportunity for malware to use incoming data obtained via infection of machines to probabilistically judge where and when further infection is likely to lead to some value return. An example of how such future AI-enabled malware might work comes from the financial sector–targeting Trickbot malware encountered in just the past two years.[24] At the point of initial compromise, Trickbot—the target of preemptive cyber operations conducted by Microsoft and US Cyber Command in October 2020 due to its prospective use in election interference activities—functions similarly to other worm-enabled malware seen since the mid-2010s. Once it establishes a foothold, however, within minutes the software targets and compromises additional machines that do not follow a clear pattern of target selection. Not only is the malware able to scale its attack at some speed, it also selects victims based on a "smart" analysis of prospective success in further infection. The word "smart" is placed in quotation marks here because the malware is not truly using the AI techniques that many experts herald as coming soon; rather, it is manually programmed to take more careful action. Nevertheless, the example stands as a case wherein a rapid understanding of the attack surface of a target network has led to an unusual strategy of infection. Not every potential target is hit but only, in the financial services case at least, targets with clear vulnerabilities in the form of outdated Server Message Block (SMB) services. The strategy there proved difficult and costly for defenders set up to handle less persistent threats.

Another manifestation of greater analysis of attack surfaces leading to increased digital insecurity lies in the wealth of data and metadata that either might be obtained via traditional intelligence methods or are already available from criminal sources. The more data available to malicious actors interested in leveraging the advantages of AI for cyber aggression, the more capable the techniques employed might be. The future may very well hold cyber campaigns of either criminal or political natures that are substantially informed by the wealth of data that might be made available to attackers for analysis. The gold standard of AI-enabled OCO, particularly those targeting broad populations or large institutions, is one substantially designed by learning systems that infer lateral approaches to targets—and, in some cases, rapidly and autonomously undertake malicious action informed by such inference—with relatively low risk of detection or mitigation. Indeed,

this threat of attack surfaces under sophisticated machine intelligence analysis is one of the core challenges that promises to impact current thinking on cyber conflict strategy and signaling.

Technique Adaptation

A second dynamic surrounding AI-enabled cyber offense is the inevitable ability of malware to autonomously select from a toolkit of options for further spread. Malware inserted into a machine might undertake environmental analyses and determine that another technique is more suited to attacking new victims than was the exploit involved in the initial compromise. Here, the shape of AI-enabled cyber attack is not much different from the sophisticated software often employed by state security institutions or other advanced persistent threat actors. Rather, it is simply a more accessible, automatable ability to empower hackers of all stripes to use tools smart enough to fit variable elements of an attack toolkit to a diverse attack surface.

Adversarial Tactical Adaptation

The threat of cyber offense upgraded by AI is also one of malware able to adjust its own strategy of approach as operations are underway. Different from a simple ability to assess potential targets and select appropriate methods of approach, AI programming will allow malware to alter its tactics in line with mission parameters as it learns more and more about the operating environment and the defenders and users populating that environment. Faced with diverse defense efforts across a diverse multinetwork attack surface, a sophisticated AI-enabled attack on defense infrastructure could, for instance, determine that the rapid promulgation most advisable for one institution—say, a research laboratory—would be associated with greater risks of detection if executed against another target—say, a military base of operations. In such circumstances, the same piece of malware might be able to select an alternative approach, such as hiding or going "slow and low" in its effort to compromise machines and exfiltrate information. Therefore, AI-enabled malware presents as an adversarial threat that functions even or especially when robust defender efforts are apparent.

Multiple Mindsets

Experts are concerned not only that AI-enabled malware will be able to analyze victim networks at scale and act autonomously to attack in ways that maximize opportunities for further compromise. A sub-element of

the ability of AI-enabled malware to change tactical approach even beyond the point of victim identification and promulgation is the opportunity for multipurpose malware that might change its own task or learn new tasks within the context of an existing operation. AI programming will allow sophisticated malware to learn about the defensive environment and compartmentalize lessons learned such that alternative "mindsets" can drive activity where mission parameters are deemed to have changed (such as upon discovery of a supervisory control system or where information has been retrieved and the task becomes one of exfiltration).

Cyber Artificial Intelligence Attacks: Threat Types

Naturally, if the potential underlying AI for cyber offense can be summed up as greater adaptability, rapidity, and opportunity for unexpected malicious behavior, then something similar can be said for the potential of AI-enabled cyber defenses. And indeed, it would be unfair to broach any discussion of the prospective impact of AI on cyber conflict without considering that the new learning, reasoning, and sensing techniques will also come to—and already have begun to—undergird the efforts of defenders. Just as AI stands to augment and enhance the offense, so too will it become a necessity for those humans in the loop whose conventional perimeter, simulative, and dissimulative defenses become the fodder from which adversarial attack AI builds better offensive routines.[25] Even here, however, it would be disingenuous to suggest that the AI arms race in cyber capabilities can be boiled down to tit-for-tat improvements in the relative capacities of those on the offense or defense. Those on the defense face complex challenges in the form of cyber artificial intelligence attacks (CAIA), which seek to take advantage of approaches to system operations and defender routines in practice to subvert their legitimate functionality.[26] In other words, CAIAs essentially constitute attacks against the AI itself that will increasingly come to underwrite cyber conflict processes. Offense, then, becomes far more attractive to cyber-capable adversaries than it is currently because of the increased potential to achieve second-order effects (i.e., to affect more than just the targeted infrastructure with a single attack by manipulating underlying algorithmic behaviors). Such attacks might fall into two categories: input attacks and poisoning attacks.

Input Attacks

Input attacks are forms of contestation that seek to fundamentally mislead an AI system and skew its efforts to classify patterns of activity.[27] If

the expectations of a model designed by a learning AI program can be subverted, new space opens for unique, hard-to-predict exploits. Notably, input attacks do not involve attacking the code of AI systems or plug-ins themselves. Rather, the point of input attacks is deception that aims to control—or at least partially shape—how an AI system is "thinking" about a given issue or functional challenge. In this way, input attacks are best thought of as counter–command and control (counter-C2) warfare.[28]

Input attacks are highly varied in their form and can functionally be a great many things. This is because input attacks are defined by the function and deployment of those models they target. They might even involve physical activities in aid of cyber outcomes. For instance, a hypothetical rerunning of the Stuxnet attack on Iran's uranium enrichment facility at Natanz—wherein the defenders employed AI in the defense of internal networks—may have necessitated a nascent phase wherein the malware lay dormant vis-à-vis its core purpose. It would then undertake secondary actions to install internal methods of subverting key defender system functions. At the same time, the malware might also benefit from input attacks by human intelligence assets. For instance, a piece of tape placed on computer monitors on-site could conceivably trick security cameras into believing that those monitors are always on. Those cameras would not then flag an anomaly when malware turns a machine on during a period of inactivity.

Poisoning Attacks

In contrast with input attacks, poisoning attacks are activities that fundamentally seek to compromise the AI programming employed in enemy systems.[29] In the Stuxnet redux example above, such an attack on the part of the malware involved might, among other things, entail gradually increasing traffic volume to certain machines during nonpeak hours. Therein lies the primary way AI systems are "poisoned"—the manipulation of data that such systems are trained on so that the model learned by the target system does not accurately reflect reality. In poisoning an AI system, attackers create backdoors through which further offensive action might be taken. This can, naturally, take several formats. An attacker might "train" a defending model to be oblivious to specific forms of anomalous behavior. Likewise, a system might be persuaded to fail or trigger some otherwise unrelated—but useful—process at a particular time when a certain action, such as a diagnostic scan, is taken.

Though the subject of poisoning attacks may be reasonably new in the literature on cyber conflict and national security, design of and defense

against such activities have long been a focus within the machine-learning literature in computer science. It would be disingenuous to suggest here that the threat is insurmountable. While much work has consistently demonstrated the limited access and resources required to engage in poisoning attacks on neural networks, a few strategies seem promising for defense on several fronts.[30] Use of blockchain or watermarking techniques to "sign" data as safe to use, for instance, might prevent compromise even when access by malicious attackers is possible.[31] Statistical optimization techniques using only subsets of data sets also decreases reliance on entire data repositories and allows for self-analysis of data provenance.[32] Others have suggested a strategy of introducing controlled perturbations into data to dramatically reduce the effectiveness of poisoning efforts.[33] Nevertheless, these defensive efforts are vulnerable to many of the conditional vulnerabilities that characterize the best network defense techniques. For instance, the need to apply such defenses at scale clashes with the inevitable complexity of the global information technology landscape and conflicts with commercial interests in product development that emphasize proprietary solutions at speed over best security practices. Thus, poisoning attacks promise to be an increasingly prominent threat to smart systems into the future, particularly as they benefit from the use of self-learning techniques to compensate for defender efforts.

Thinking About Cyber AI Attacks at Scale

While it is tempting to think of the threat of attacks that compromise the function of AI systems that defenders must increasingly come to rely on only at the level of cyber operations themselves, the implications of CAIAs for national security apparatuses go beyond such considerations. Specifically, the problem of poison for modern security institutions exists beyond the implications for cyber conflict; indeed, cyber operations are just one element of the challenge. Given the coming proliferation of AI across military functions, security planners face the threat of skewness from nigh uncountable sources. If adversary militaries wish to skew North Atlantic Treaty Organization (NATO) analytics, they might use conventional military deception methods—such as deploying decoy vehicles during military maneuvers to mislead NATO forces about the normal scale and dispersion of adversary forces—as easily as they might tamper with training data via cyber means. Thus, it would be at least partially disingenuous to argue here that the augmentation of cyber conflict processes by AI constitutes a unique-to-the-domain coming transformation.

Shaping Behavior in an Age of Adversarial Learning

What is particularly unique about the intersection of artificial intelligence and cyber conflict processes, however, is that the centrality of cyberspace to the deployment and operation of soon-to-be ubiquitous AI systems implies expanded motivations—such as an increased interest in using cyberspace to affect extra-domain technological processes—for operations within the domain. The prospect of subverting AI-driven security functions—in particular, the prospect of fundamentally poisoning the deliberative and operational bases of important national security establishment functions—incentivizes operations in cyberspace beyond in-domain effects and outcomes. On the one hand, cybersecurity experts might expect an intensification of cyber conflict and criminal activities around the world based on near-term adoption of advancing AI programming that promises rapid adaptability and sophistication without either major investment or the need for major human presence in the loop. On the other hand, the same experts might expect an intensification of such activities because cyber AI attacks will clearly so often involve effects beyond the domain (e.g., cyber operations not operationally focused on some digital compromise so much as they are intended to affect real-world approaches to risk management, strategic assessment, and resultant military deployments, financial outlays, etc.).

Implications for Deterrence in Cyberspace

What follows is a contextual analysis of the implications of AI-augmented cyber attack for current strategic approaches to mitigating cyber conflict. This includes the strategy of forward defense based around the dynamics of persistent engagement between adversaries in the cyber domain that now constitutes US Title 10 approaches to operations online. It suggests several core problems that either intensify or newly manifest in an era of large-scale proliferation of AI in cyber. The focus on US strategy is intentional; changes to America's force posture in the fifth domain represent the concrete edge of efforts to adapt prevailing approaches to cyber conflict in the context of both intensifying digital interference since 2010 and the failing applicability of legacy security concepts to the challenge. Dynamics of AI-augmented cyber conflict and the ensuing questions that must be addressed vary beyond the scope of such singular focus, of course. But national contextualization allows for more in-depth exploration and produces analytic outcomes generalizable beyond the case.

Defending Forward and Persistent Engagement

In 2018, as it was elevated to the status of unified combatant command in the US military, Cyber Command promulgated a new strategic vision centered around the concept of persistent engagement.[34] To put the concept and strategy that emerge bluntly, PE means that Cyber Command intends be everywhere, constantly maintaining presence and employing necessary tools against US adversaries in networks wherever they might be found. The strategy pushes back against past practices by the US and its allies wherein operations were based on the political desire to mitigate cyber risk principally via norm development and through deterrent efforts that stemmed substantially from the shape of Cold War postures.[35]

In terms of the strategic logic of engagement in the domain, the PE strategy largely emerges from the work of Richard Harknett and Michael Fischerkeller during their time as scholars attached to Cyber Command. The authors argue that the unique character of cyberspace means that traditional deterrent approaches are doomed to failure.[36] Given that deterrence involves strong demonstrations of defense or meaningful statements of punishment, they contend, prospects for developing a sustainable deterrent posture online are limited (or so the architects of the new approach hold).[37] It is extremely difficult to demonstrate defensive capabilities at the scale demanded by a national cyber deterrent strategy, and punishment rarely works in the way it is intended.

Communicating specific meaning in retaliation is difficult, particularly where the diversity of activities that constitute cyber conflict is immensely high. Moreover, response options are often not ready to execute in the time frame required by policy makers that seek to deter. And conceptual agreement on the significance or role of certain elements of the domain is not easy to come by, with poor understanding of what might be meant—if anything—by sovereignty online being a hallmark of the digital world.

The result is an alternative strategy—persistent engagement—that emphasizes "defending forward." This posture involves cyber forces of Western nations operating beyond government and domestic networks to actively contest enemy activities aimed at harming national security or other national interests. Such operations, it is argued, can avoid escalation by embracing the doctrine of selective engagement and can be designed specifically to scale tactical efforts into strategic gains. In doing so, the idea is that the behavior of adversaries can be shaped and the scope of what is deemed to be appropriate competition can be made known.[38] The resultant condition should, it is hoped, be one of "agreed competition" wherein the bounds of cyber conflict deemed to be acceptable can be

consistently made known and where the worst excesses of digital insecurity for states might be avoided by the institution of precise conditions of case-by-case deterrence.[39]

Basic Challenges of AI for Persistent Engagement

Thinking effectively about the problem of poison for cyber conflict processes—particularly as a subset of all national security processes—is difficult in that we fundamentally have to think about learning as it manifests in two different settings, the organizational setting and in the construction of AI systems. It is not simply enough to consider the impact of rapid learning techniques for cyber conflict as we understand it today, though that approach to thinking about the problem of AI in this area does suggest some obvious challenges to be faced by prevailing strategy.

Above almost all other implications, broad-scoped upgrading of "conventional" cyber techniques portend a simple functional challenge for cyber strategy. Specifically, it suggests a narrowing of the space within which adversaries might undertake cost-benefit calculations and come to believe that the benefits of further action are outweighed by the costs that might be imposed in the domain by forward defenders. Simply put, if smart tools exist that can more reliably avoid detection, take lateral routes to targets, or scale effects much more quickly than is the norm today, then adversaries are likely to exhibit increased willingness to continue operating under circumstances they would not have previously. Especially given that the stakes of defection from agreed conditions of competition are not typically very high in political terms, this contraction of that space wherein persuasion is argued to be possible under a doctrine of persistent engagement ostensibly makes meaningful signaling yet more difficult from situation to situation. Likewise, at the most basic level, the proliferation of relatively robust abilities to achieve effects in the digital domain via lateral action—action that takes indirect, harder-to-predict pathways toward targets and outcomes—suggests that we might see recurrent incidents in areas where the threat had previously been thought to have been realized and countered in some form.[40]

It is worth noting on an operational level that AI-enabled cyber conflict adds a new dimension to the traditional perception problem experienced in cyberspace wherein attribution of intent or agency is particularly difficult at the point of threat detection and analysis.[41] Where a probing attack or some other action is detected, it is rare that the investigator is able to discern between run-of-the-mill adversary efforts to conduct espionage or some attacking action. In the near term, another possibility is that cyber

actions may be not linked with either espionage or direct attack but with attempts to interfere with the function of AI programming.[42] The particular danger here is that such attempts may involve activities even less clearly discernable as aggressive than is the case with espionage activities.

AI, Feedback Loops, and the Logic of Persistent Engagement

Beyond functional AI-induced issues of added sophistication and perception, the strategic logic of PE may be made more vulnerable when new learning tools employed at scale also impact second-order conditions relevant to the conduct of cyber conflict in broader international relations. Jason Healey, in his analysis of challenges awaiting the United States as it continues to commit to the strategy of persistent engagement, discusses such logic in the context of feedback loops.[43] Feedback loops describe any system where the outputs of a process either constitute or affect the inputs of that same process as it iterates over time. Positive and negative feedback mean, respectively, outcomes that either amplify the original process or dampen it. With PE, the idea is that forward operation allows the US to see attacks before they occur (informing domestic actors more effectively as a result) and produces "friction" that increases the costs of antagonism for adversaries.[44] Alongside more conventional deterrent operations, this activity should in theory create negative feedback—a dampening, constraining effect on aggressive behavior in cyberspace.[45]

In discussing PE in this fashion, Healey joins others concerned about the risks of such an assertive policy.[46] A main concern, what he refers to as "on-net" challenges, revolves around the issues of misperception and tacit intersubjectivity in direct cyber interactions discussed above.[47] Beyond simple functional difficulties, it is worthwhile reiterating in more detail that AI exacerbates a fundamental problem with PE as a strategy, namely that it includes no concrete method of communication other than conflict actions themselves. This particularly manifests on two fronts.

First, the assumptions of tacit bargaining as a critical pushback against the track record of deterrent efforts in cyberspace now functionally sit at the heart of American cyber conflict policy.[48] This is problematic because strategic assumptions must be based on a range of operational dynamics that are inevitably hard to fully observe from just one side of the screen. Friction designed to produce negative feedback is likely to fail if costs to adversaries are minimal.[49] Certainly, operators can design tactical actions to avoid such an outcome and maximize strategic gains.[50] But to some degree, the impact of forward defense efforts will always be a question of adversary infrastructure and resource commitment, about which the home team will

always have imperfect information. If reconstruction of infrastructure is inexpensive, friction will not work. Today, this is a concerning element of PE because the funding structures and priorities of authoritarian opponents can be relatively opaque. Likewise, a robust defense against PE aggression lies not only in in-domain actions but also in adversary efforts to build operational resilience. This may be the commitment of resources sufficient to regularly make American "friction" ineffectual at cost imposition. Or, somewhat more worrying, this might involve further decentralization of extensive cyber operations infrastructure on the part of adversaries, essentially adding distance and compartmentalization of assets with the use of internal, criminal, and non-state proxies to create redundancy and introduce obstacles to American efforts to map the battlespace.

Second, it is not fully clear what the "acceptable" behavior desired by the strategy of PE might look like.[51] As opposed to a strategy like that of the "fleet in being"—which some scholars have suggested as a more realistic strategic alternative to persistent engagement—that explicitly permits low-intensity antagonism, PE calls for setting norms of behavior to be defined by prevailing military and political stakeholders.[52] This means, as some have noted, that there may easily exist tactical or political reasons over time to attempt to interdict any aggressive behavior. And because the only communication intended under PE is in the method of engagement, mechanisms to quickly clarify expectations promise to be clunky at best.

Artificial intelligence adds to the challenges facing PE on both fronts. Currently, a major concern is that failed friction will lead to "aggression spirals" in which both sides escalate in search of costly digital territory. AI brings new dimensionality to this concern. Simplistically, AI is likely to lower costs of reconstruction of digital assets across the board, making this situation of failed friction more likely. After all, the game-changing fact of the revolution in machine learning amounts to an ability to overcome—via use of self-reprogrammable learner algorithms—the programming bloat that inevitably costs organizations resources as their infrastructure is called upon to provide more diverse specialty functions at scale. Additionally, in-domain escalations might be motivated beyond the link between offensive actions and imposed costs assumed by the strategy. Aggression spirals under controlled conditions—at least, as the adversary judges the risks and intentions involved—provide opportunities to train defensive platforms and to showcase strategies of aggression intended to mislead the peer competitor. Such activity is clearly attractive, as enough evidence of adversary behavioral preferences might create cognitive schema and

operational cultures that dampen tactical adaptability in the face of new patterns in the data.

The question of "acceptable" behavior also looms large given the question of AI in cyber. As laid out above, states are likely to be motivated to directly influence AI systems employed by adversaries, both those pertaining to cyber operations and those functionally at the heart of innumerable national security and societal processes. This dual focus on subversion of process and of process beyond domain-specific capacities makes answering the behavioral question even harder. If subversive attacks that have increasingly real meaning for strategic knowledge capabilities are imperative for competitors heavily invested in use of AI, then what conventional metric can possibly be used to gauge "aggression"? This is particularly salient given the way the PE strategy holds espionage apart as "acceptable" behavior. If low-intensity and lateral engagement begin to threaten core functional capabilities beyond what is currently the case, then strategists will be forced to either by demonstration or explicit declaration attempt to offer tighter definitions of what activity is "unacceptable" that parses apart espionage from poisoning operations. And such a development seems likely. After all, the logic of PE emerges in trusting that an invisible hand of "market" correction will work to produce behavioral equilibrium. The strategy would surely fail, at least in part, if trust in the integrity of that hand faltered. Actors must understand the limits of the game they are playing. The threat of a subverted rule set itself will likely motivate assertive action to stabilize the battlespace, adding yet another layer of complex calculation to daily action and reaction in the domain.

The issue of cyber conflict in an era where cyberspace is the primary highway for the operation of innumerable AI systems spread across important security and societal infrastructure bears additional mention in the context of PE. Forward defense is simply one layer of the US effort to limit aggression experienced via cyberspace.[53] Traditional deterrent operations and efforts to build norms using conventional diplomatic approaches remain as robust pillars of American cyber foreign policy. Persistent engagement is the lynchpin underlying these additional efforts (see fig. 1).

However, the success of PE seems likely only where there will be clear situational alignment with other efforts. In large part, this is because there is so much natural oscillation in the conditions of sophisticated cyber conflict actions and the reactions of complex state military and civilian government infrastructure. The context of much complexity in international affairs—including global and domestic politics, private versus public behavior in cyberspace, intelligence versus military use of the fifth domain, and

more—constructs nested spaces wherein contrasting perspectives about the logic of digital engagement make sense. Simply put, the "AI-ification" of advanced industrial states in the years to come is likely to cause the multiplication of such spaces as cyberspace becomes the central artery through which so much added manipulative traffic flows. This will make it harder for adversaries to be sensitive to each other's signals while at the same time motivating actions targeting non-domain effects as a strategy to degrade state confidence in the value of longitudinal data pertaining to cyber operations.

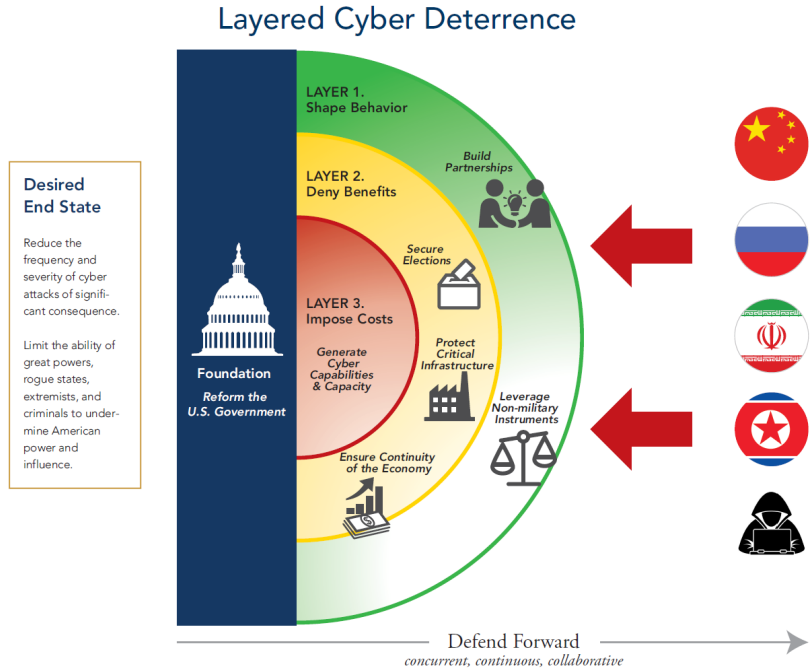


Figure 1. Layered Cyber Deterrence. (Reproduced from Angus King and Mike Gallagher, co-chairs, US Cyberspace Solarium Commission, *Cyberspace Solarium Commission Report* [Arlington, VA: US Cyberspace Solarium Commission, March 2020], 7, https://www.solarium.gov/.)

A final implication of AI for PE and current approaches to cyber conflict is with how efforts to secure cyberspace might degrade, as Healey notes, the reality of "an open, interoperable, reliable, and secure Internet that fosters efficiency, innovation, communication, and economic prosperity."[54] Forward defense naturally relies on a great deal of trust among allies, private sector partners in industry, and other elements of civil society. Yet the actions implied by the strategy are inevitably among the most invasive and assertive imaginable on the part of a national government like that of the United States. This is particularly the case given the way patterns of

engagement are unlikely to ever be the predictable intrusions of terrestrial conflict. This produces a trust challenge to the success of the strategy, with no easy solutions and many fault lines where irritation is not only possible but likely. Global outrage following leaks from Edward Snowden, the Shadow Brokers, and more was not limited to foreign states and persons but was also common in the American private sector, even within the ranks of companies with knowledge of the upstream and provider-sourced data collection efforts of the National Security Agency.

With AI, reliance on distributed smart infrastructure critical to both national security efforts and targets of foreign cyber-enabled manipulation exacerbates the traditional civil-military relations problem already in existence in the digital age. How does the government carry out its security mission and ensure its coercive capability when it is forced to cede ownership of that mission to the de facto governors—including technology companies, Internet service providers, and backbone operators—of the operational domain in question (cyberspace)? Naturally, this problem strikes at the heart of challenges encountered and problematized in recent years regarding attempts to deter foreign digital aggression via cost imposition by denial. The current strategy is, in many ways, a military-oriented solution to challenges that are not—as so many scholars and strategists are wont to suggest—purely driven by domain characteristics but also by legal, normative, and practical government-industry challenges to ensuring national security in democratic states. Persistence underlying more conventional deterrent, norm-building efforts essentially constitutes an effort to define the character of the battlespace, pushing American presence everywhere to shape adversary expectations. With AI, the promise and problem of poisoning the battlespace suggests a (potentially massive) wrinkle for broader American efforts to head a liberal world order, as systematic efforts aimed at subverting algorithmic processes across global society to serve US security objectives spark inevitable outrage. Beyond the obvious broader issues that such outrage might bring about for American foreign policy efforts, the implication is yet another tangled web set to complicate PE as the bedrock of cyber strategy. After all, without additional communications methods baked into the strategy beyond conflict actions themselves, how can democratic states—and particularly the United States—maintain stable deterrent conditions when high political considerations force decision-makers to limit assertive digital activities? Permanent engagement may seem theoretically necessary, but it seems unlikely to be perpetually possible where exogenous changes in political conditions or in the nature of the battlespace threaten. AI stands to produce both.

The Learning Problem

Cyber conflict driven by the adaptability and rapidity brought on by AI poses several challenges to the strategy of persistent engagement. Policy makers and practitioners must inevitably grapple with increasing uncertainty around the state of common knowledge between actors in the domain. The perception dynamic described above, for instance, is uniquely concerning for current strategic thinking on cyber conflict management insofar as cyberspace is likely to be the domain of political activity most central to efforts to poison or otherwise interfere with AI systems. Moreover, state interest in operations of a poisoning nature via cyberspace is likely to grow over time as opportunities for manipulating processes that underlie strategy development and force posture determination proliferate.[55] Both of these points mean that strategic efforts to constrain adversaries' cyber actions relative to in-domain considerations may fail simply because they are not effectively armed with appropriate assumptions about the motivations of actors to operate online.

More broadly, the advent of narrow AI baked into most functional elements of a state's national security apparatus implies an enduring tension in the conduct of persistent operations intended to shape adversary behavior. All else equal, the existence of robust AI systems on the part of foreign adversaries implies a learning problem: the more security institutions operate to shape behavior, the more adversaries should be empowered to understand and overcome such strategies. Much as in the case of generative adversarial networks (GAN) that study the actions AI models take to continually improve offensive capabilities, AI-enabled cyber forces presented with unique patterns of behavior-shaping attack from abroad will naturally undergo a process of adversarial learning.[56] Foreign action does not so much bound the shape of acceptable behavior as define the criteria under which future aggression is probabilistically less likely to induce some cost. Given the incentive described above toward the use of AI-enabled software agents with dramatically higher track records of success than non-AI-enabled versions, the commonplace existence of such systems seems likely to work against the development of static norms of behavior.

Finally, the result of an emergent era in which AI-driven adversarial learning is the key feature of interstate interactions online is a perpetual challenge of validation. In recent scholarship, there have already been some discussions about the challenges involved in applying relevant metrics to the strategy of PE such that defense practitioners might determine its effectiveness.[57] Such challenges multiply given the AI-ification of cyber conflict processes and the problem of poison as regular features of opera-

tions in the domain. Whereas analysis of broad patterns of activity might otherwise offer some indication as to the effectiveness of forward defensive efforts aimed at dissuading particular adversary behaviors, such metrics may not apply in significant fashion in an era where counteraction from foreign peers is not expected to be tit for tat, but rather an entirely alternative approach. In other words, where the paradigm of operations shifts from in-kind engagement—even if that engagement emerges from an admittedly diverse toolkit—to an imperative of lateral approach and misdirection, attempts to validate current strategic processes seem likely to be ineffective beyond simplistic analysis of major event incidence.

AI and Cyber Conflict Cascades

A final consideration seems particularly worthy of mention at this juncture. As is true in all areas of human interaction, misperception in cyber conflict is naturally not always—or even usually—a one-off occurrence. One action produces an interpretation of that action, which then informs further activity (or is itself that further activity). That reaction is then interpreted in turn, and so on. Misperception can spiral from minor assumption to major failure of interpretation if such a chain of events cannot be stopped. Such failures characterize many of the major conflict episodes in modern history. Of course, in strategic competition between states, one generally assumes that a great many analytic and procedural mechanisms bound up in the complex institutional landscape of international relations serve to backstop spiraling misperception.

Scholars have paid the problem of conflict spirals in cyberspace some sizable amount of attention, not least in the ubiquitous recognition that intention is difficult to ascertain in digital interactions. What may appear to be an attack may simply have been a probe, an effort to understand the battlespace or to engage in a non-warfighting activity. Beyond this level of discussion, however, scholars have given limited attention to the idea of cascading effects. After all, though automated attacks present a particular challenge wherein automated responses may be triggered, cascade effects at some point do tend to cease due to backstops in the algorithm—kill switches or conditional code that end a process without further human interaction.

With the use of AI, there is substantial risk that more interactions might produce a critical mass of activity leading to major unintended effects. One commonly cited example of such a critical mass event is the flash crash of the stock market on 6 May 2010. Though no definitive cause has been agreed upon by researchers, conventional wisdom attributes a Dow Jones

loss of almost 1,000 points in just 36 minutes to automated selling algorithms that reacted to an unusual perturbation of the market—often said to be an accidental sale some orders of magnitude above what was deemed normal. The result was a trillion-dollar loss in the market that then quickly rebounded in the following hours. Looking at the event, it is easy to imagine how dueling AI—or, perhaps more worryingly, a "battle" between AI and dumber automated algorithms—could rapidly and disastrously produce negative effects of strategic consequence. These could range from critical infrastructure shutdowns to counteroffensive cyber volleys of sufficient scale to prompt a state response beyond the domain.

Though this article does not attempt to address the challenge of cascades specifically, it seems clear that planners should avoid formalizing PE-style strategies in procedure and in code. Doing so would invite the opportunity for a diverse prospective set of flash crashes. It also seems reasonable to suggest that national security planners must be mindful of opportunities for such spiraling beyond the practice of cyber conflict. If CAIAs are indeed likely to become the norm of engagement in cyberspace, then we must be consistently mindful of the possibility that unexplained conflict developments not thought to be linked to the fifth domain may yet be affected by it. Thus, the human in the loop must not only be a decision-maker at US Cyber Command, but rather must also represent an assemblage of those stakeholders with jurisdiction over other areas of national defense.

Conclusion

The purpose of this article has been to contribute to the nascent literature on AI and national security activities by outlining how AI is likely to alter the shape and strategic calculations involved in interstate cyber conflict. It is hoped this information will be a resource for those interested in thinking more clearly about how AI stands to alter the dynamics of interstate and cyber conflict processes. Naturally, a substantial part of the effort here has been definitional. Indeed, it is from this effort (i.e., the categorization of different threat forms linked to the augmentation of cyber conflict processes by AI models and systems) that the primary argument of this article emerges.

Broadly, that argument is that the centrality of cyberspace to the deployment and operation of soon-to-be-commonplace AI systems implies new motivations for operations within the domain. More specifically, though AI does not itself imply inevitable advantages for attackers over defenders (or vice versa), adversarial learning techniques add complexity to already complex operational conditions in cyberspace and may contrib-

ute to an uptick in offensive behavior. Perhaps more worryingly, the centrality of the Internet to new AI systems incorporated across all areas of national security—not just to cyber conflict processes—indicates that sophisticated adversaries may be increasingly motivated to launch offensive online actions to achieve effects in other domains. The implications for current cyber conflict strategies—particularly those by Western defense enterprises—are numerous and remain to be assessed in full as literature on the subject is developed in the future. Nevertheless, some immediate takeaways are apparent.

First, strategic planners and policy makers must recognize from the start that there are two levels of challenge when it comes to AI augmentation of cyber conflict processes. At the first level, AI promises to reduce the opportunity to shape competition in cyberspace in favorable terms. At the second, AI intensifies and adds a new dimension to the challenges of validity and attribution already present in cyber operations. Simply put, given opportunities for the poisoning of soon-to-be ubiquitous AI models at work in security apparatuses, how can defenders really know what they think it is they know about the integrity of their systems? At the strategic level, given that broad-scoped attempts to shape competition between AI-enabled adversaries are likely to empower opponents via a process of adversarial learning, how can policy makers and military practitioners really know what to believe about strategic conditions?

Second, success in meeting the challenges of deploying AI for national security purposes will likely hinge on the approach organizations take toward trusting their AI systems and managing the interaction of human and machine operators.[58] To some degree the previous discussion involves the problem of "ghosts in the machine." That is, human assumptions present in the code of machine intelligence systems are the true problem underlying effective AI deployment for national security purposes. While such problems are arguably unavoidable as we move toward more common employment of AI, it seems likely that protocols for keeping humans in the loop at critical junctures are part of the solution to problems of system poisoning (either malicious or self-afflicted).

Finally—and perhaps most significantly—in the forthcoming era of AI-enabled contestation in world affairs, it seems clear that strategy development, assessment, and validation must emerge from the cross-domain understanding of the strategic motivations of adversaries. Cyberspace is not only a domain where unique forms of contestation and signaling can occur but also potentially the most critical terrain over which actions can be taken to affect processes that underlie all areas of modern society. Given

this potential, strategic planners would do well to build from assumptions that move beyond simple logic-of-the-domain characterizations of digital affairs. As some scholars have increasingly argued in both implicit and explicit terms, cyber conflict so often manifests in aid of nondigital contestation that we would do well to couch our analyses in terms of the logic of conflict processes other than cyber.[59] This stands to be especially the case with AI, not least given the fact that its targeting for security purposes is so likely to be tied to the use of computer and Internet systems upon which such programming must inevitably run. SSQ

Christopher Whyte, PhD

Dr. Whyte is an assistant professor at the L. Douglas Wilder School of Government and Public Affairs, Virginia Commonwealth University. His research interests include cyber conflict, information warfare, and emerging technology. He was lead editor for *Information Warfare in the Age of Cyber Conflict* (Routledge, 2020) and is co-author of a forthcoming Georgetown University Press book on military innovation surrounding artificial intelligence.

An earlier version of this article appeared in the *Proceedings of the 12th International Conference on Cyber Conflict: 20/20 Vision: The Next Decade*.

Notes

1. It should be noted that the topic of involving AI in the organization and application of military functions is not new, particularly in popular media. Instances of storytelling and more factual exploration can be found in film and written work stretching back through the early–mid twentieth century.

2. See, among others, Stephen Hawking et al., "Transcendence Looks at the Implications of Artificial Intelligence—but Are We Taking AI Seriously Enough?," *The Independent*, 1 May 2014, https://www.independent.co.uk/; and Max Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence* (New York: Knopf, 2017).

3. For example, the well-publicized threat of autonomous machine "slaughter bots" that, in a fictional future, catalyze societal breakdown as governments and private actors alike are empowered to kill opponents anonymously and at scale—in an attempt to convince audiences of the stakes involved in getting AI "right." For an overview of expert opinion on AI, see Vincent C. Müller and Nick Bostrom, "Future Progress in Artificial Intelligence: A Survey of Expert Opinion, in *Fundamental Issues of Artificial Intelligence,* ed. Vincent C. Müller (Synthese Library; Berlin: Springer, 2016), 555–72, https://www.nickbostrom.com/.

4. The "Third Offset" is a strategy intended to be used by the US Department of Defense to counter and overcome advances being made by key peer competitors, such as China and Russia, in areas of military modernization and technology development. The term "Third Offset" refers to previous efforts to overcome perceived positional, military, or technological advantages held by the Soviet Union during the Cold War—the first of which originated with the famed Project Solarium convened by President Dwight Eisenhower in the 1950s. Robert Work, deputy secretary of defense (speech, "Third Offset Strategy," Brussels, Belgium, 28 April 2016), https://www.defense.gov/; Cheryl Pellerin, "Deputy Secretary: Third Offset Strategy Bolsters America's Military Deterrence," *Defense News*, 31 October 2018, https://www.defense.gov/; and Katie Lange, "3rd Offset Strategy 101: What It Is, What the Tech Focuses Are," *DODLive* (blog), Defense Department, 30 March 2016, https://www.doncio.navy.mil/.

5. This point refers to the oft-cited manifestation of revolutions in military affairs (RMA) that dot human history. On the historical emergence of the RMA, see Dima Adamsky, *The Culture of Military Innovation: The Impact of Cultural Factors on the Revolution in Military Affairs in Russia, the*

US, and Israel (Stanford, CA: Stanford University Press, 2010); and Benjamin Jensen, "The Role of Ideas in Defense Planning: Revisiting the Revolution in Military Affairs," *Defence Studies* 18, no. 3 (2018): 302–17, https://doi.org/10.1080/14702436.2018.1497928. On the distinction between a revolution in military affairs and military revolutions more broadly, see MacGregor Knox and Williamson Murray, eds., *The Dynamics of Military Revolution, 1300–2050* (Cambridge: Cambridge University Press, 2001).

6. For the limited work to date on AI and strategic studies, see, for example, Benjamin M. Jensen, Christopher Whyte, and Scott Cuomo, "Algorithms at War: The Promise, Peril, and Limits of Artificial Intelligence," *International Studies Review*, 2019, viz025, https://doi.org/10.1093/isr/viz025; Joe Burton and Simona R. Soare, "Understanding the Strategic Implications of the Weaponization of Artificial Intelligence," in *2019 11th International Conference on Cyber Conflict (CyCon)* (Tallinn: NATO CCD COE Publications, 2019), 249–65, https://ccdcoe.org/; and Kareem Ayoub and Kenneth Payne, "Strategy in the Age of Artificial Intelligence," *Journal of Strategic Studies* 39, no. 5-6 (2016): 793–819, https://doi.org/10.1080/01402390.2015.1088838; Heather Roff, *Advancing Human Security through Artificial Intelligence* (London: Chatham House, May 2017), https://www.chathamhouse.org/; Michael C. Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1, no. 3 (May 2018): 36–57, https://doi.org/10.15781/T2639KP49; Kenneth Payne, *Strategy, Evolution, and War: From Apes to Artificial Intelligence* (Washington, D.C.: Georgetown University Press, 2018); Heather M. Roff, "COMPASS: A New AI-Driven Situational Awareness Tool for the Pentagon?," *Bulletin of the Atomic Scientists*, 10 May 2018, https://thebulletin.org/; Kenneth Payne, "Artificial Intelligence: A Revolution in Strategic Affairs?," *Survival* 60, no. 5 (2018): 7–32, https://doi.org/10.1080/00396338.2018.1518374; Michael Horowitz et al., "Strategic Competition in an Era of Artificial Intelligence," Center for a New American Security (CNAS), 25 July 2018, https://www.cnas.org/; and Miles Brundage et al., "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation," February 2018, *arXiv:1802.07228,* https://img1.wsimg.com/.

7. See, for instance, Enn Tyugu, "Artificial Intelligence in Cyber Defense," in *Proceedings of the 2011 3rd International Conference on Cyber Conflict*, eds. C. Czosseck, E. Tyugu, and T. Wingfield (Tallinn: NATO Cooperative Cyber Defence Centre of Excellence, 2011), 95–105; and Mariarosaria Taddeo and Luciano Floridi, "Regulate Artificial Intelligence to Avert Cyber Arms Race," *Nature* 556 (April 2018): 296, https://media.nature.com/.

8. For a broad overview of the scope and dynamics of cyber conflict, see, for example, Brandon Valeriano and Ryan C. Maness, *Cyber War versus Cyber Realities: Cyber Conflict in the International System* (Oxford: Oxford University Press, USA, 2015); and Christopher Whyte and Brian Mazanec, *Understanding Cyber Warfare: Politics, Policy and Strategy* (Abingdon: Routledge, 2018).

9. Machine learning is technically a subfield of AI research that, according to many, now virtually demands consideration as its own technology.

10. For instance, Brundage et al., "Malicious Use of Artificial Intelligence."

11. Jensen, Whyte, and Cuomo, "Algorithms at War," 10.

12. For a contemporary description of such efforts, see, for example, Alan Turning, "Computing Machinery and Intelligence," *Mind* 49 (1950): 433–60; John von Neumann, *The Computer and the Brain* (New Haven: Yale University Press, 1958); Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (New York: Cambridge University Press, 2010); and Herbert Simon, "Artificial Intelligence: An Empirical Science," *Artificial Intelligence* 77, no. 2 (1995): 95–127, https://pdfs.semanticscholar.org/.

13. Randolph Kline, "Cybernetics, Automata Studies, and the Dartmouth Conference on Artificial Intelligence," *IEEE Annals of the History of Computing* 33, no. 4 (October–December 2011): 5–16.

14. See Kline, 5–16; J. Moor, "The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years," *AI Magazine* 27, no. 4 (Winter 2006): 87–91, https://www.aaai.org/; and Bruce Buchanan, "A (Very) Brief History of AI," *AI Magazine* 26, no. 4 (Winter 2005): 53–60, https://www.aaai.org/.

15. J. McCarthy et al., "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence," 31 August 1955, http://www-formal.stanford.edu/.

16. Burton and Soare, "Understanding the Strategic Implications," 5.

17. Jensen, Whyte, and Cuomo, "Algorithms at War."

18. For an overview of machine learning, see Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep Learning" *Nature* 521 (May 2015): 436–44. Also see Volodymyr Mnih et al., "Human-Level Control through Deep Reinforcement Learning," *Nature* 5, no. 18 (2015): 529–33, http://dx.doi.org/10.1038/nature14236; and David Silver et al., "Mastering the Game of Go without Human Knowledge," *Nature* 550, no. 7676 (October 2017): 354–59, https://doi.org/10.1038/nature24270.

19. For perhaps the most accessible description of machine learning at the point of operation, see Pedro Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake our World* (New York: Basic Books, 2015).

20. For a seminal description of perception as a component element of broader attempts to build deep learning and reasoning systems, see Nicola Jones, "Computer Science: The Learning Machines," *Nature* 505, no. 7482 (2014): 146–48, https://www.nature.com/.

21. For further information on NLP, see Stephen F. DeAngelis, "The Growing Importance of Nature Language Processing," *Wired*, February 2014, https://www.wired.com/; and Erik Cambria and Bebo White, "Jumping NLP Curves: A Review of Natural Language Processing Research," *IEEE Computational Intelligence Magazine* 9, no. 2 (May 2014): 48–57, https://doi.org/10.1109/MCI.2014.2307227.

22. For further reading on intelligent machine vehicle systems, see Mario Gerla et al., "Internet of Vehicles: From Intelligent Grid to Autonomous Cars and Vehicular Clouds," *2014 IEEE World Forum on Internet of Things* (*WF-IoT)*, Seoul, 2014, 241–46, https://doi.org/10.1109/WF-IoT.2014.6803166; and Alberto Broggi et al., "Intelligent Vehicles," in *Springer Handbook of Robotics*, 2d ed., eds. Bruno Siciliano and Oussama Khatib (Berlin: Springer, 2016), 1627–56, https://link.springer.com/.

23. "Attack surface" is a term of art used to describe the sum of weak points of a given system. According to Tim Stevens, "the attack surface is less a physical boundary to be defended than a logical membrane of potential vulnerability distributed in space and time." More than just a set of functional components, an attack surface typically includes these elements: technical (i.e., infrastructure), social (i.e., the behaviors and psychology of system users/operators), and economic (i.e., the competing interests that characterize a system's usage). Tim Stevens, "Knowledge in the Grey Zone: AI and Cybersecurity," *Digital War*, 2020, https://www.researchgate.net/.

24. For a description of the episode in context, see DarkTrace, *The Next Paradigm Shift: Cyber-Attacks, AI-Driven*, research white paper (San Francisco: DarkTrace, 2018), https://www.oixio.ee/. Also see Lior Keshet, "An Aggressive Launch: TrickBot Trojan Rises with Redirection Attacks in the UK," *Security Intelligence* (2016); and Darrel Rendell, "Understanding the Evolution of Malware," *Computer Fraud & Security* 2019, no. 1 (January 2019): 17–19, https://doi.org/10.1016/S1361-3723(19)30010-7.

25. For discussion of simulation as an element of strategic interactions in cyberspace, see Erik Gartzke and Jon R. Lindsay, "Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace." *Security Studies* 24, no. 2 (2015): 316–48, https://doi.org/10.1080/09636412.2015.1038188.

26. The term "cyber artificial intelligence attacks" is inspired by its recent usage in Marcus Comiter, *Attacking Artificial Intelligence: AI's Security Vulnerability and What Policymakers Can Do about It* (Cambridge, MA: Belfer Center for Science and International Affairs, 2019), https://www.belfercenter.org/, 28.

27. Comiter, 19.

28. See Norman B. Hutcherson, *Command and Control Warfare: Putting Another Tool in the War-Fighter's Data Base*, No. AU-ARI-94-1 (Maxwell AFB, AL: Air University Press, 1994), https://apps.dtic.mil/; and Jeffrey A. Harley, *The Role of Information Warfare: Truth and Myths* (Newport, *RI:* Naval War College, Joint Military Operations Dept. 1996), https://apps.dtic.mil/.

29. See Comiter, *Attacking Artificial Intelligence*, 28.

30. See recent work, for instance, Ali W. Shafahi et al., "Poison Frogs! Targeted Clean-Label Poisoning Attacks on Neural Networks," presented at the 32nd Conference on Neural Information Processing Systems (NIPS), Montréal, Canada, 2018, https://arxiv.org/; Pang Wei Koh, Jacob Steinhardt, and Percy Liang, "Stronger Data Poisoning Attacks Break Data Sanitization Defenses," *arXiv preprint arXiv:1811.00741* (2018), https://arxiv.org/; Saeed Mahloujifar and Mohammad Mahmoody, "Can Adversarially Robust Learning Leverage Computational Hardness?," *arXiv preprint arXiv:1810.01407* (2018), https://arxiv.org/; and Chen Zhu et al., "Transferable Clean--Label Poisoning Attacks on Deep Neural Nets," in *Proceedings of the 36th International Conference on Machine Learning*, Long Beach, California, PMLR 97, 2019, 7614–23, *arXiv preprint arXiv:1905.05897* (2019), http://proceedings.mlr.press/.

31. Shafahi et al., "Poison Frogs!"

32. Matthew Jagielski et al., "Manipulating Machine Learning: Poisoning Attacks and Countermeasures for Regression Learning," in *2018 IEEE Symposium on Security and Privacy (SP)* (New York: IEEE, 2018), 19–35, https://arxiv.org/.

33. Tribhuvanesh Orekondy, Bernt Schiele, and Mario Fritz, "Prediction Poisoning: Utility-Constrained Defenses against Model Stealing Attacks," *arXiv preprint arXiv:1906.10908* (2019), https://arxiv.org/.

34. The White House, *National Cyber Strategy of the United States of America* (Washington, D.C.: The White House, 2018), https://www.whitehouse.gov/.

35. Paul M. Nakasone, "An Interview with Paul M. Nakasone," *Joint Force Quarterly* 92 (1st Quarter 2019): 4–9, https://ndupress.ndu.edu/.

36. Michael P. Fischerkeller and Richard J. Harknett, "Deterrence Is Not a Credible Strategy for Cyberspace," *Orbis* 61, no. 3 (2017): 381–93, https://doi.org/10.1016/j.orbis.2017.05.003.

37. For the broad literature on deterrence in cyberspace, see, for example, Martin C. Libicki, *Cyberdeterrence and Cyberwar* (Santa Monica, CA: RAND Corporation, 2009), https://www.rand.org/; Amir Lupovici, "Cyber Warfare and Deterrence: Trends and Challenges in Research," *Military and Strategic Affairs* 3, no. 3 (2011): 49–62, https://pdfs.semanticscholar.org/; Matthew D. Crosston, "World Gone Cyber MAD: How "Mutually Assured Debilitation" Is the Best Hope for Cyber Deterrence," *Strategic Studies Quarterly* 5, no. 1 (2011): 100–116, https://deepsec.net/; Eric Talbot Jensen, "Cyber Deterrence," *Emory Int'l L. Rev.* 26 (2012): 773, https://law.emory.edu/; Dorothy E. Denning, "Rethinking the Cyber Domain and Deterrence," *Joint Force Quarterly* 77 (2d Quarter 2015): 8–15, https://ndupress.ndu.edu/; Emilio Iasiello, "Is Cyber Deterrence an Illusory Course of Action?." *Journal of Strategic Security* 7, no. 1 (2014): 54–67, https://scholarcommons.usf.edu/; and Uri Tor, " 'Cumulative Deterrence' as a New Paradigm for Cyber Deterrence," *Journal of Strategic Studies* 40, no. 1-2 (2017): 92–117, https://doi.org/10.1080/01402390.2015.1115975.

38. Michael P. Fischerkeller and Richard J. Harknett, "Persistent Engagement, Agreed Competition, Cyberspace Interaction Dynamics and Escalation," *Orbis* 61, no. 3 (Summer 2017): 381–93.

39. See, for example, Department of Defense, Defense Science Board, *Defense Science Board Task Force on Cyber Deterrence* (Washington, D.C.: Defense Science Board, 2017), https://apps.dtic.mil/; and Amb. John Bolton, "Transcript: White House Press Briefing on National Cyber Strategy – Sept. 20, 2018," https://news.grabien.com/.

40. This point references the oft-cited framing of cyber conflict history in the West as emerging via a series of realization episodes that have prompted a series of institutional and doctrinal adaptations over the past three decades. See Jason Healey, ed., *A Fierce Domain: Conflict in Cyberspace, 1986 to 2012* (Arlington, VA: Cyber Conflict Studies Association, 2013).

41. See, for example, Nicholas Tsagourias, "Cyber Attacks, Self-Defence and the Problem of Attribution," *Journal of Conflict and Security Law* 17, no. 2 (2012): 229–44, https://papers.ssrn.com/; Jon R. Lindsay, "Tipping the Scales: the Attribution Problem and the Feasibility of Deterrence against Cyber attack," *Journal of Cybersecurity* 1, no. 1 (2015): 53–67, https://doi.org/10.1093

/cybsec/tyv003; and Thomas Rid and Ben Buchanan, "Attributing Cyber Attacks," *Journal of Strategic Studies* 38, no. 1-2 (2015): 4–37, https://ridt.co/.

42. This issue lies at the heart of what Buchanan labels the "cybersecurity dilemma." See Ben Buchanan, *The Cybersecurity Dilemma: Hacking, Trust, and Fear between Nations* (New York: Oxford University Press, 2016).

43. See Jason Healey, "The Implications of Persistent (and Permanent) Engagement in Cyberspace," *Journal of Cybersecurity* 5, no. 1 (2019): tyz008, https://doi.org/10.1093/cybsec/tyz008.

44. Healey, 5.

45. Healey, 5–6. For the original discussion of the notion of persistence feeding deterrent norm-building, see Fischerkeller and Harknett, "Persistent Engagement," 388–91.

46. Among others, see Max Smeets, "Cyber Command's Strategy Risks Friction with Allies," *Lawfare*, blog, 28 May 2019, https://www.lawfareblog.com/; Brandon Valeriano and Benjamin Jensen, "The Myth of the Cyber Offense: The Case for Restraint," CATO Institute Policy Analysis 862, 15 January 2019, https://www.cato.org/; Herb Lin and Max Smeets, "What Is Absent from the U.S. Cyber Command 'Vision,'" *Lawfare*, blog, 3 May 2018, https://www.lawfareblog.com/; and Max Smeets and H. A. Lin, "A Strategic Assessment of the U.S. Cyber Command Vision," in *Bytes, Bombs and Spies: The Strategic Dimensions of Offensive Cyber Operations*, eds. Herbert Lin and Amy Zegart (Washington, D.C.: Brookings Institution Press, 2019), http://www.jstor.org/.

47. Healey, "Persistent (and Permanent) Engagement," 7–8.

48. Michael P. Fischekeller and Richard J. Harknett, "What Is Agreed Competition in Cyberspace?," *Lawfare*, blog, 19 February 2019, https://www.lawfareblog.com/.

49. Even the head of Cyber Command admits this. See "An Interview with Paul M. Nakasone," *Joint Force Quarterly* 92 (1st Quarter 2019): 4–9, https://ndupress.ndu.edu/.

50. As the wording of the strategy itself suggests, taking reference from Fischerkeller and Harknett's original analysis in "Deterrence Is Not a Credible Strategy for Cyberspace."

51. Max Smeets, "There Are Too Many Red Lines in Cyberspace," *Lawfare*, blog, 20 March 2019, https://www.lawfareblog.com/.

52. Valeriano and Jensen, "Myth of the Cyber Offense," 8. This argument is based on the original notion of the "fleet-in-being" developed by Corbett. See Julian S. Corbett, *Some Principles of Maritime Strategy*, ed. Eric Grove (Annapolis: US Naval Institute, 1988).

53. See the recent US Cyberspace Solarium Commission report for a description on prevailing thought on the relationship between cost imposition via persistent engagement, deterrent operations, and norm building. Angus King and Mike Gallagher, co-chairs, US Cyberspace Solarium Commission, *Cyberspace Solarium Commission Report* (Arlington, VA: US Cyberspace Solarium Commission, March 2020), 7, https://www.solarium.gov/.

54. Healey, "Implications of Persistent (and Permanent) Engagement," 25.

55. This assertion is quite arguably backed by work that demonstrates in both quantitative and qualitative terms in increasing turn toward political warfare as an adjunct of cyber conflict in line with the proliferation of digital services and social platforms that undergird major societal functions. See, for instance, Brandon Valeriano, Benjamin M. Jensen, and Ryan C. Maness, *Cyber Strategy: The Evolving Character of Power and Coercion* (New York: Oxford University Press, 2018).

56. For GANs, see James Vincent, "Deepfake Detection Algorithms Will Never Be Enough," *The Verge*, 27 June 2019, https://www.theverge.com/. The phrase "adversarial learning" is a common one used by computer scientists to describe how machine learning algorithms are capable of adapting to hostile operational environments by crystalizing alternative—rather than combative—approaches to operation. See Daniel Lowd and Christopher Meek, "Adversarial Learning," in *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, 641–47, ACM, 2005; and Pavel Laskov and Richard Lippmann, "Machine Learning in Adversarial Environments," *Machine Learning* 81, no. 2 (2010): 115–19, https://doi.org/10.1007/s10994-010-5207-6.

57. See, for instance, Jason Healey and Neil Jenkins, "Rough-and-Ready: A Policy Framework to Determine if Cyber Deterrence Is Working or Failing," in *2019 11th International Conference on Cyber Conflict (CyCon)*, vol. 900 (IEEE, 2019), 1–20, https://doi.org/10.23919/CYCON.2019.8756890.

58. This is not a thus-far uncommon argument made by scholars of cyber conflict. See, for instance, Rebecca Slayton, "What Is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment," *International Security* 41, no. 3 (2017): 72–109, https://doi.org/10.1162/ISEC_a_00267.

59. See, for instance, Christopher Whyte, "Dissecting the Digital World: A Review of the Construction and Constitution of Cyber Conflict Research," *International Studies Review* 20, no. 3 (2018): 520–32, https://doi.org/10.1093/isr/viw013.

PERSPECTIVE

Nuclear-Armed Hypersonic Weapons and Nuclear Deterrence

Col Stephen Reny, USAF

Abstract

Nuclear-armed hypersonic weapons, with their ballistic missile defense (BMD) penetrating capability, will provide an overall strategically stabilizing effect in the global arena but will further destabilize regional competitions. Development and deployment of BMD is a strategically destabilizing agent since adversaries perceive that they can no longer hold each other at risk of a retaliatory nuclear strike. Nuclear hypersonic weapons, with their promised capability to defeat missile defenses, will bolster expectations of reciprocal nuclear strikes. When this capability to provide retaliation is undermined, strategic instability ensues and manifests as arms races, aggressive posturing, and bellicose rhetoric. Therefore, global nuclear powers, with their robust counterforce capabilities, should develop nuclear-armed hypersonic weapons to return deterrence to an era of assured vulnerability that keeps nuclear weapons holstered. However, introducing hypersonics, with first-strike counterforce and decapitation capabilities, to regional nuclear power competitions will have the opposite effect, further destabilizing an already uneasy peace. In both cases, some period of greater strategic instability will exist as nuclear-armed hypersonic weapons become operational in an unbalanced manner. That is, as one nuclear power attains BMD-defeating capability, opposing powers will perceive that they are at a disadvantage. To mitigate this transition period of instability, global powers should proceed in developing hypersonic weapons but counter regional instability by banning regional development and curtailing hypersonic technology proliferation.

*****

Over two decades ago, Keith Payne wrote in *Deterrence in the Second Nuclear Age* on the challenges of the changing dynamics of nuclear deterrence in the era following the bipolar Cold War. He cautions, with near clairvoyance, that the US needs to balance assured nuclear retaliation against the Russian Federation while hedging protection against rogue states with ballistic missile defense (BMD) development.[1] In other words, the US must consider the second- and third-order

effects of its missile defense policies and capabilities on strategic stability. Taking Payne's argument one step further confirms that US development of ballistic missile defenses has upset great power strategic stability by violating the key nuclear deterrent principle of assured vulnerability. Essentially, there are two nuclear arenas to explore regarding the effects of hypersonic nuclear weapons: global nuclear powers (e.g., US, China, and Russia) and regional nuclear powers (e.g., India, Pakistan, and North Korea). In the context of global deterrence stability, countries seek equilibrium, and in doing so, they pursue nuclear-armed hypersonic missiles with their high-speed, maneuverable, missile defense–defeating capabilities to bolster counterforce options and return stability to strategic deterrence. The consequences of this pursuit are now materializing as China and Russia accelerate programs in hypersonics. China is considering changes in its nuclear alert posture.[2] It has been less direct about confirming research in nuclear-armed hypersonic weapons. However, intelligence indicates Chinese hypersonic capabilities heading in the nuclear-armed direction for similar missile-defense-penetrating reasons.[3] Russia is racing to develop hypersonic nuclear weapons to defeat ballistic missile defenses.[4] It has stated intentions to mate nuclear warheads to these hyper-fast, maneuverable weapons to counter US missile defenses against nuclear attack.[5] A hypersonics competition is also being sought regionally for defense-penetrating capabilities, increasing instability in regional nuclear standoffs as seen in the Pakistan-India conflict. These competitions will have destabilizing effects due to the respective weak counterforce postures and capabilities combined with the first-strike and decapitation potential that nuclear hypersonics may bring.

Today, most hypersonic weapons research globally is focused on conventional arms primarily for the potential value of these weapons in penetrating anti-access environments.[6] Yet some authors fear that the development of hypersonic nuclear missiles will bring us closer to nuclear holocaust. This logic is not universally applicable across global and regional areas and is not grounded in sound deterrence theory.[7] For global nuclear powers, the anticlimactic good news is that if nuclear-armed, non-ballistic hypersonic missiles become a staple of their military arsenals, the long-term deterrent effect will manifest as greater strategic stability. In other words, nuclear-armed hypersonic missiles, with the promised capability to defeat missile defenses, will usher in a return to assured nuclear vulnerability among the global nuclear powers. However, the unsettling news is that as nuclear hypersonics infiltrate regional nuclear power arsenals, strategic instability will increase. In both cases, the path to this new era of stability

is fraught with tension and uncertainty. The world will likely experience times of greater strategic instability as nuclear hypersonic missiles become operational in an unbalanced manner. As one nuclear power attains BMD-defeating capability, opposing powers will perceive they are at a counterforce capability disadvantage against ever-advancing, increasingly affordable, and proliferating missile defenses.

To buttress these arguments, this article first reviews how introducing new technology may create deterrence instability. It then examines hypersonic capabilities and the effects of these weapons on nuclear deterrence. Finally, it uses Albert Wohlstetter's attributes of stable nuclear deterrence to demonstrate the implications of nuclear-armed hypersonic missiles for strategic stability. This article makes the case that dismantling missile defenses and adding hypersonic technology to a nonproliferation ban may be the best approach to avoid global transition instability periods and overall regional instability. These policy proposals will seem counterintuitive, but they appear logical and necessary to stabilize the changing nuclear deterrence environment.

Strategic Stability and Nuclear Deterrence Instability

A stable nuclear deterrence environment, as described by Wohlstetter and Thomas Schelling, is underwritten by each country's credible second-strike capability to levy extraordinary punitive costs against adversaries.[8] In essence, stability contains two parts: the belief that a target adversary has the capability and the political will to deliver a punishing counterstrike, ensuring any first strike would fail to dismantle the opponent's capability to counterstrike. Specifically, Wohlstetter outlines six attributes of a credible (as believed by a nuclear country and its adversary) second-strike deterrent system: It must (1) be reliable, affordable, and sustainable; (2) survive enemy attack; (3) make and communicate the decision to retaliate; (4) reach enemy territory with enough fuel to complete the mission; (5) penetrate the enemy's active defenses; and (6) destroy the target despite passive defenses.[9]

When this retaliatory capability is no longer perceived to be credible (violates one of the six Wohlstetter stability attributes) and is profoundly costly, then instability ripples throughout nuclear and nonnuclear nations, manifesting as arms races, bellicose rhetoric, force posturing, and universal unease.[10] Thomas Schelling similarly wrote, "It is not the 'balance'—the sheer equity or symmetry in the situation—that constitutes mutual deterrence; it is the stability of the balance. The balance is stable only when neither, in striking first, can destroy the other's ability to strike back."[11] The

importance of Schelling's statement cannot be underscored enough: a nuclear stalemate requires that all nuclear parties have an invulnerable second-strike capability to provide optimal stability. A nuclear-armed country that fields technology either mitigating its opponent's capability to impose cost or enhancing its own benefit for initiating a nuclear first strike (an enhanced preemptive strike weapon) leads other nations to question their capability and/or credibility, upsetting the status quo. For deterrence to be effective, nuclear powers must thoroughly evaluate the effect of technology insertion into the nuclear arena on deterrent stability.

The first example highlighting deterrence-destabilizing technological advantages can be found in the USSR launch of *Sputnik*. The orbiting sphere showcased a first-strike nuclear attack capability of the USSR, upset the perception of a nuclear stalemate, fed the US's fear of a missile gap, and spurred the intercontinental ballistic missile arms race.[12] Prior to the 1957 launch of *Sputnik*, the only intercontinental nuclear delivery capability that existed was long-range bombers. Nuclear-tipped intermediate range ballistic missiles (IRBM) did exist at the time and were deployed throughout the European theater, threatening the USSR, but there were no comparable opposing missile deployments that directly threatened the North American continent. Using a fleet of bombers against the US entailed a cascade of warnings and hours of flight time that made a surprise attack unlikely.

Further, the US established Air Defense Command to intercept and mitigate any bomber-borne nuclear threat the USSR could impose. At the time, defenses against airborne ballistic missiles did not exist, let alone defenses against intercontinental ballistic missiles. With IRBMs deployed along the borders of the USSR and threatening Soviet targets, the USSR was spurred to develop an ICBM force to hold the US at similar risk.[13] As one can imagine, a Soviet satellite shot into space and allowed to fly over the US without challenge fueled a sense of naked vulnerability among strategic nuclear thinkers, politicians, and average civilians.

Looking at the *Sputnik* situation through Wohlstetter's stability lens, the USSR's perceived capability to obliterate the US with virtual impunity upended the nuclear deterrent environment. This instability was further exacerbated by the limited survivability of a US retaliatory force of long-range bombers (a nuclear force susceptible to a preemptive ICBM strike). Altogether, the perception of a missile gap—driven home by the Soviet radio beacon flying above—introduced an instability into nuclear deterrence that manifested as an ICBM arms race.

Another example of this instability-inducing technology is the advent of antiballistic missile (ABM) systems or ballistic missile defense efforts by

the US in 1967. Touted to deny an adversary's advantage to impose cost by claiming the capability to intercept inbound nuclear warheads, US BMD programs naturally alarmed the USSR. Looking at Wohlstetter's stability attributes, a BMD capability violates an adversary's "penetrate enemy active defense" characteristic and diminishes the possibility of ensuring a costly retaliation. Against a US BMD, the USSR perceived its missiles to be less likely to provide a credible retaliatory punch, fundamentally undermining the assured vulnerability concept essential to stable nuclear deterrence. This capability-limiting perception spurred the USSR to develop its own BMD program and pushed both superpowers into a counter-BMD arms race. This arms race manifested in the development of multiple independently targetable reentry vehicles (MIRV) capable of defeating BMD systems. [14] The BMD race spurred the MIRV race, which fundamentally was an attempt to return the nuclear deterrence environment back toward strategic stability between the US and USSR.[15] Distressed by this BMD arms race and its promise to intercept nuclear warheads, the Nixon administration and Soviet leadership signed the ABM treaty, halting any real deployment of a BMD umbrella over the US and USSR.[16]

Understanding nuclear deterrence, the elements of nuclear deterrence stability, and the symptoms of instability underpins the key elements for analyzing new capabilities into the nuclear deterrence arena. The desirable stable nuclear deterrence environment, as defined by Wohlstetter's six criteria to credibly guarantee a costly retaliatory response, underwrites the modern US nuclear deterrence strategy. The symptoms of volatility—including nuclear arms races, bellicose rhetoric, and general international unease—are highlighted as nuclear deterrence instability indicators. Altogether, they serve as the foundation to measure modern BMD and the effects of nuclear-armed hypersonic missiles on global and regional nuclear deterrence.

Capabilities and Effects of Hypersonic Weapons

There has been a tremendous amount of concern about introducing nuclear-armed hypersonic missiles over the past decade. Across the internet and in the press, words like "hypersonic arms race," "hypersonic weapons," and "hyper escalation" are making headlines.[17] These manuscripts attribute to hypersonic vehicles the capability to penetrate BMD and air defenses with impunity, reach targets with absolute precision and accomplish all of these with tactical surprise. However, many of these statements are speculative and are not grounded in physics or even the realm of the possible.

In reviewing literature regarding hypersonics, there seems to be an almost mystical admiration and a general misunderstanding of vehicles traveling in this speed regime. The reality is that hypersonic speed has existed since 12 April 1961 when Russian cosmonaut Yuri Gagarin reentered Earth's atmosphere, traveling at speeds fast enough to ionize air into plasma. US hypersonic testing began with the manned X-15 rocket plane that surpassed the hypersonic speed of Mach 5 in June 1961. Today, weapons that travel at hypersonic speeds are already in the inventories (mostly as part of air and BMD systems) of many nations and are on a trajectory to become mainstream in commercial space travel and military applications in the near future. The world is on the brink of a breakout in hypersonic technology use and employment. This makes it essential for readers to have a basic understanding of the hypersonic flight regime as well as hypersonic missiles and their capabilities. Toward that end, this section defines *hypersonic*, discusses the engineering challenges of traveling at this speed, and addresses types of hypersonic assets and their competencies. Such a fundamental appreciation for these super-fast capabilities will complement the deterrence argument.

The National Aeronautics and Space Administration (NASA) defines *hypersonic* as the atmospheric speed regime greater than or equal to five times the speed of sound.[18] While hypersonic speed is not a new achievement, it is not a ubiquitously traveled speed realm. Every space reentry vehicle—from Mercury-Redstone space capsules and ICBM reentry vehicles to the space shuttle—traverses the hypersonic regime, sometimes entering the atmosphere in excess of Mach 25.[19] Further, modern military surface-to-air missiles, such as the SA-21 Growler (S-400 as named by the Russian developers), streak to their targets at speeds up to Mach 12.[20] In each hypersonic case mentioned, the technological hurdles were (and still are) quite extreme. Until the recent emergence of commercial space programs, this realm was limited to a few state-sponsored programs using national resources to solve the significant engineering challenges.

All hypersonic literature agrees that heat is the most challenging engineering problem facing hypersonic flight. Figure 1 illustrates the stagnate point temperature calculations at the skin of hypersonic vessels. At hypersonic speeds, the punishing temperatures experienced disassociate and ionize the air, resulting in chemically reactive airflows and plasma.[21] These reactions and plasma-inducing temperatures also produce communication barriers that block reception and transmission of radio signals.[22] Finally—but not least of the problems of hypersonic speed—is shockwave impingement. In certain situations, shockwaves produced by hypersonic

flight can act like a blowtorch wherever they contact the aerospace vehicle frame, burning through the skin and further compounding temperature-related problems.[23]

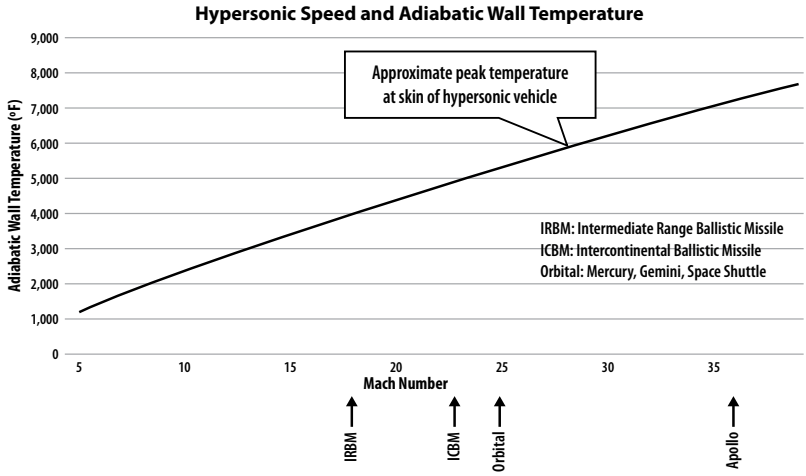


Figure 1. Hypersonic speed versus skin temperature

Note: Mach speeds and temperatures were calculated using a 170,600 (52 km) altitude, a radius of 1 ft, and material emissivity of 0.8. Mach temperature calculator was provided by Mr. Barry Hellman, Air Force Research Laboratory.

(*Source*: For IRBM, ICBM, orbital, and Apollo speeds, see John D. Anderson, Jr., *Modern Compressible Flow:With Historical Perspective*, 3rd ed. [Boston: McGraw-Hill, 2003], 18.)

Despite these extreme difficulties, humans have entered the realm of hypersonic speed with advancements in material, propulsion, and understanding. In vehicles like space capsules, space shuttles, the X-15 and other reentry vehicles, the thermal problems of hypersonic travel have been managed using a combination of ablation heat shields, silicon tiles, carbon composites, zirconia, and high-temperature nickel and titanium alloys.[24] However, these before-mentioned hypersonic vehicles also mitigated heat problems with relatively short durations of exposure to high temperatures. Reduced time exposure lessens the impact of convective heating.

Using these engineering leaps in hypersonics, two basic types of hypersonic vehicles have received most of the attention in military research and application: the boost-glide vehicle and the powered-cruise vehicle (cruise missile). At this time, both vehicle classes require the boost of a large rocket motor to reach hypersonic performance. A boost-glide vehicle is launched using rocket boost systems and glides, much like the space shuttle, to a target. The air-breathing hypersonic cruise vehicle can be launched from ground or airborne assets using a rocket motor boost to achieve hypersonic speeds fast enough to ignite a scramjet[25] (an engine designed to operate at hypersonic speeds) and power the flying vehicle to a target.

Figure 2 compares the flight profiles of a hypersonic cruise missile, hypersonic boost-glide machine, and ballistic missiles. Hypersonic cruise missiles typically travel between ~70,000 to ~120,000 feet above sea level. These altitudes ensure that there are enough oxygen, air volume, and pressure to support combustion for a scramjet engine while mitigating heat-inducing and dynamic pressure properties of lower altitude, higher air densities. Boost-glide vehicles are launched to high altitudes (sometimes leaving the atmosphere), pitch over, and establish a descending glide to a target at hypersonic speeds.[26] As they approach their ground target, both vehicle classes perform a slowing descent to lower Mach numbers for thermal and dynamic pressure management.[27] Lower altitudes also allow for an increase in maneuverability in the terminal phase of flight.[28] Figure 2 depicts the depressed trajectories of glide and cruise missiles as compared to ballistic parabolas. The importance of depressed trajectory and maneuverability attributes are discussed later.

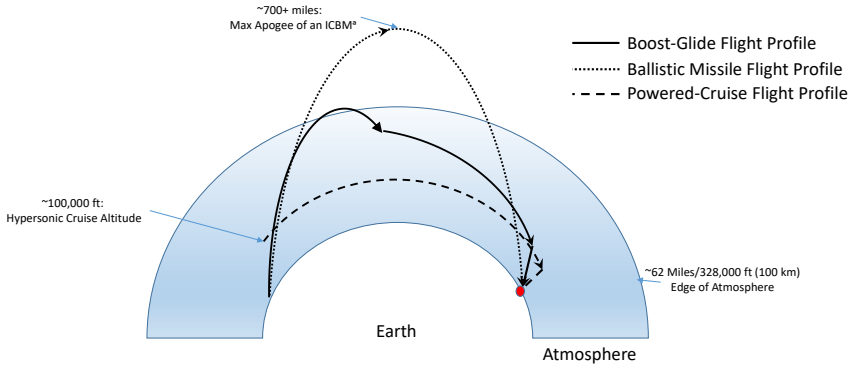


Figure 2. Nominal flight paths of ballistic missiles, boost-glide vehicles, and cruise vehicles

[a]ICBM apogee obtained from Federation of American Scientists, "LGM-30 Minuteman III," accessed 18 January 2017, https://fas.org/.

The different flight profiles are a result of different airframe designs. Hypersonic boost-glide vehicle wedge design generally maximizes glide range and can allow relatively large payload capacities, carrying several thousand pounds of cargo or weapons. The load capacity and size are limited only to the power of the boost vehicle (i.e., more boost = more weight and size available for the glide vehicle).[29] Boost-glide hypersonic vehicles are expected to have a global range comparable to that of intercontinental ballistic missiles (equal to or greater than 5,500 km).[30]

Besides having a boost phase, hypersonic, air-breathing cruise missiles have a different aerodynamic design than boost-glide vehicles. These missiles are engineered to have a sleek, narrow, futuristic bullet shape that manipulates the shockwave for scramjet operation and limits the amount

of drag for maximum speed and range.[31] Although hypersonic cruise missiles can technically be launched from ground-based sites, most US efforts have focused on air-launched hypersonic cruise vehicles, likely in compliance with the former Intermediate Nuclear Force Treaty that banned ground-launched cruise missiles.[32] These cruise missile prototypes are currently launched by bomber-size aircraft due to their relatively large size and weight. Consequently, the air-launched configuration will limit expected ranges (currently 200 to 850+ miles), comparable to short- and intermediate-range ballistic missiles (short missile range: <620 miles; IRBM: 1,800 to 3,400 miles).[33] The reason for these shorter ranges is the payload capacity limits of the combined weight of the booster and a fully fueled hypersonic vehicle. (Recall that the hypersonic vehicle size limit directly relates to the size of the rocket motor that propels it to hyper speed. ICBM-size boosters are generally used for larger boost-glide vehicles while smaller rocket boosters are used for launches from airborne platforms).[34] However, weaponized hypersonic cruise missiles are expected to get smaller as the technology matures to enable higher speeds and launch from smaller, possibly fighter-size, aircraft.[35]

Defense Penetrating Panacea?

The widely touted, missile-defense-defeating capability of hypersonics is triggering speculation and instability in the world. Despite this hype, the reality is that while hypersonic weapons will be better at defeating robust defenses than what is available today, they will not be a panacea against missile and ballistic missile defenses. Physics is the largest limiting factor in the capability of hypersonic flying; understanding these limits is important when judging the true impact hypersonics will have on the nuclear deterrence landscape. When hypersonics are matched against the latest missile defenses of today and tomorrow, the fast-flying projectiles will not be impervious to counter-systems with equal speed and maneuverability capabilities.

Before proceeding, it is helpful to review the hypersonic weapon concepts most of the literature seems to be using. To begin, the promised hypersonic capabilities are a combination of speed, range, accuracy, and maneuverability. With these combined capabilities, hypersonic weapons would conceptually be used for global strike (e.g., a boost-glide vehicle), reaching any target within minutes and penetrating defenses with immunity through a combination of tactical surprise (detected later due to lower altitude flight path when compared to a ballistic missile—see fig. 3) and maneuverability. Additionally, a hypersonic cruise missile could be launched at standoff

ranges and penetrate dense and deadly defensive systems, thereby striking targets with impunity.

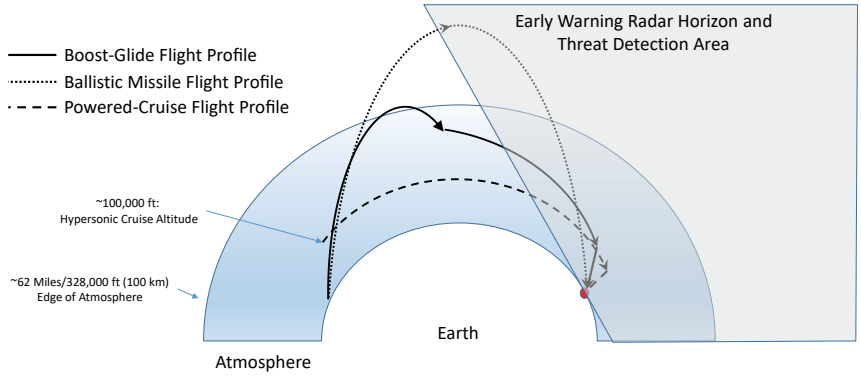


Figure 3. Flight profiles and early-warning radar threat detection

The real characteristics of hypersonic vehicle speed and flight profiles still make these weapons vulnerable to today's modern defense weapons, being only marginally more survivable and effective than ballistic missile–deployed weapons.[36] Reviewing the flight profiles of hypersonic boost-glide and cruise missiles, they must fly at high altitudes (70,000 ft or higher) for aerodynamic load and dynamic pressure limitations (fig. 2). These high-altitude profiles leave them detectable at longer ranges than if they flew at lower altitudes where traditional radar-evading cruise missiles fly.[37] Also, hypersonic boost-glide and cruise vehicles do not move faster than reentering ICBM-launched MIRVs, the very objects some missile defenses are designed to counter.[38] Further, the descending, decelerating end-game profile required of hypersonic vehicles to hit ground targets puts them at greater risk of engagement.

Pitting hypersonics against modern and soon-to-be-fielded advanced anti-missile systems is sobering. James Acton's report *Silver Bullet?* points out that hypersonic cruise and boost-glide weapons can theoretically be detected and engaged by Russian-made S-300 (SA-20 Gargoyle) surface-to-air missile systems' antiballistic missile capability.[39] Furthermore, current antiballistic missile systems, such as the Russian-made S-400 Triumfator (SA-21 Growler), specifically boast the capability to engage hypersonic cruise missiles with an interceptor that can maneuver at 20 times the force of gravity (g) at 100,000 feet. This maneuverability promises to mitigate the advantages of an inbound hypersonic weapon.[40] Also, the soon-to-be-fielded S-500 Triumfator-M advertises advanced air and space defense capabilities and anti-hypersonic warhead ability, further grounding hypersonic speculation.[41] Of course, it is hard to make specific

comparisons against the anti-missile systems and hypersonic flyers without detailed information on the hypersonic vehicles themselves. The point is that anti-missile systems are continually advancing and proliferating.[42] They are already quite lethal to equally fast-flying ballistic missiles and have hypersonic interceptors that are quite maneuverable at high altitudes. Together, these evolving capabilities mean that hypersonic weapons will likely face a formidable challenge around densely defended targets (the very same targets hypersonics are designed against).

The best attribute a hypersonic weapon has is its speed. Using the hypersonic concepts of global strike and defense penetration, speed will likely be used to achieve tactical surprise and compress the timeline required to counter this inbound threat.[43] Therefore, to truly attain tactical surprise against a modern antiballistic and anti-hypersonic missile system, an inbound hypersonic missile would have to fly at an altitude low enough to avoid detection long enough so that by the time it is detected, there is not enough time to defend against it (see table 1). Again, lower altitudes are problematic for hypersonic flight because the lower altitudes overpressure hypersonic engines and prolonged flight creates extreme thermal management issues.[44]

Table 1. Estimated warning times of different hypersonic systems

		Global Boost Glide System (Mach 10–25)	Intermediate-Range Ballistic Missile (~ Mach 15)	Mach 5 Hypersonic Cruise Missile
Strike Range (Miles)		6,800	2,200	930
Warning Time (Minutes)	Early Warning Satellite	33	19	16
	Early Warning Radar	4	14	11
	Air Defense Radar	3	0	8

Source: Table modified from *James M. Acton, Silver Bullet? Asking the Right Questions about Conventional Prompt Global Strike* (Washington, D.C.: Carnegie Endowment for International Peace, 2013), 70, https://carnegieendowment.org/.

Other attributes that put hypersonic vehicles at a further disadvantage are their significant heat signature and relatively limited maneuvering capability. The heat signature produced at hypersonic speeds makes these vehicles very detectable, even visible to the human eye (as hot as 2,000 degrees Celsius [3,600 degrees Fahrenheit] at Mach 10 for the X-43. This is the same temperature as jet engine exhaust identifiable by existing infrared detectors and heat-seeking missiles.[45] Maneuverability, cited as a key survivability attribute of hypersonic weapons, will make them more difficult to engage. However, this maneuvering attribute will likely make hypersonic vehicles only marginally more effective since turning at hypersonic speed is problematic. High-speed turns generate giant turn radii and loss of energy (resulting in slower speed), requiring increased maneuvering

space and decreased range. Taking evasive action at these high-speed ranges has a high potential to throw these swift vehicles miles off course in a fraction of a second. Timed at the right ranges, engaging an inbound hypersonic weapon can force survivability maneuvering, instantly turning the missile far enough off course as to make the weapon miss the intended target. Evasive actions at hypersonic speed will also slow down the missile and/or require more fuel, reducing its range and survivability. Further, antiballistic missile interceptors use a combination of thrusters and aerodynamic devices at high altitudes to achieve high maneuverability—methods that hypersonic vehicles can also use to defeat defenses.[46] As mentioned before, if anti-missile systems are already employing these maneuvering capabilities, then hypersonic weapons' main advantage is speed to delay detection until it is too late for an effective defense. Finally, this hypersonic maneuverability characteristic has been around since the late 1970s with the advent of the Advanced Maneuverable Reentry Vehicle (AMaRV). This warhead was designed to defeat BMDs through maneuvers during reentry and the terminal phase of flight—a stark departure from normal ballistic trajectories.[47] Novel at that time, the AMaRV was declared operational for the Minuteman III in 1982.[48] Over three decades have passed since warhead maneuvering was introduced to the nuclear arena, allowing missile defense system development to mature to the point they can counter trajectory-changing reentry vehicles.

The Nuclear Hypersonic Effect

Despite the reality of hypersonic capabilities, China and Russia have announced they are developing hypersonic boost-glide vehicles and hypersonic cruise missiles to penetrate US antiballistic missile defenses.[49] Furthermore, there is speculation that both countries are developing conventional and nuclear variants of these weapons.[50] India, in a joint venture with Russia, is also developing hypersonic technology as a response to robust air and sea defenses.[51] However, India seems focused on multimission (indications are primarily an anti-ship) hypersonic cruise missiles with ranges around 290 km (180 miles) and has not indicated plans to produce a nuclear variant.[52] Altogether, four nations (US, China, Russia, and India) are developing hypersonic technology in response to sophisticated anti-access and BMD systems.

BMD systems are not new to the realm of nuclear deterrence and have existed in various US and Russian (USSR) forms since the 1950s. Whether hypersonics can penetrate BMD defense or not, the ubiquitous belief that the fast-fling systems can defeat missile defenses is what matters in a stra-

tegic environment. Relying on this defense-penetrating belief and reflecting on Wohlstetter's stability attributes of penetrating enemy active defenses, it is easy to understand the USSR's staunch resistance to US development of nationwide ballistic missile–defeating systems. With the USSR perception that its strategic nuclear arsenal could be made partly or completely impotent, it came to the treaty table to dismantle any US ABM effort. The US Arms Control and Disarmament Agency indicated the ABM Treaty would "decrease the pressures of technological change and its unsettling impact on the strategic balance."[53]

Similar perceptions exist today concerning US missile defense systems. The US has stated that the deployment of BMDs is, according to the 2014 *Quadrennial Defense Review* (QDR) and the 2019 *Missile Defense Review* (MDR), to protect the homeland and its allies from regional actors, namely North Korea and Iran, from limited ballistic missile attack.[54] This position to build strategic missile defenses, according to well-known nuclear deterrent theorist Herman Kahn, is the moral obligation of a country to save lives (saving some is better than saving none), even if the system is not foolproof.[55] Additionally, the QDR and MDR rationale is grounded in the philosophy that BMDs be built to deter "smaller" countries. This argument appears based on the assumption that such a system would be more effective against fewer warheads, rendering a country's small nuclear arsenal impotent.[56] However, what the QDR, MDR, and Kahn fail to adequately address are the second-order effects of developing such defensive systems, specifically, how other global nuclear powers may view and respond to these defenses and how these systems would affect regional nuclear standoffs.

We are presently seeing the second-order effects of such defenses manifest as nuclear deterrence instability and hypersonic arms races.[57] Thomas Schelling predicted this dilemma when he wrote, "ABM systems deployed in both countries would make preemptive war more likely, and the arms race more expensive."[58] Consequently, due to the impression a BMD system can diminish or neuter the effectiveness of a nuclear attack, the MDR, experts, and scholars alike believe Russia and China are developing nuclear-armed hypersonic weapons designed to render these defenses futile.[59]

It is no surprise that several nations are enamored with hypersonic capabilities. After all, the potential of moving military operations at speeds above two miles per second has a tremendous appeal. Militaries that can move weapons or cargo at these speeds will set a tempo of conflict that no adversary can currently match. However, the conclusion regarding hypersonic capability is that rhetoric is proceeding actual capability. The speed, range, and maneuverability of hypersonics are all attributes that will make

them preeminent weapons, but that capability will likely not culminate in the penetrating defense panacea some literature speculates. The engineering problems these speedy vehicles face are titanic and require not only unique material and design solutions but must fly high altitude profiles; both attributes which degrade the promised defense-penetrating capabilities. Understanding these fast-aero vehicle characteristics is fundamental to gaging the effects they will have on nuclear deterrence. Hypersonics will be another arrow in an array of capabilities that, when used, will be part of a holistic force concept to produce desired military effects. In other words, hypersonics are certainly an evolution in weapons technology, not a revolution. It will provide only modest defense-penetrating capability.

Implications for Deterrence: Global and Regional

The development of nuclear-tipped hypersonic missiles is the deterrence "environment" attempting to return the nuclear order to a state of stability. As Wohlstetter implies in "The Delicate Balance of Terror," the international nuclear deterrence system is a balancing act of attributes. As he states, "To deter attack means being able to strike back in spite of it."[60] When the counterstrike option is diminished, as a BMD system has the ability to do—whether actually or perceptually—the deterrent system is shaken and becomes unbalanced and unstable. Russia and China naturally feel disadvantaged by the US development of a credible, albeit limited, ICBM defense capability. However, given the limited number of US BMD defenses, they can easily be overwhelmed.[61] This limited missile defense capability restricts options for a counterstrike, assuming an adversary's doctrine had a spectrum of counterstrike choices versus just massive retaliation (the only way to defeat this limited ballistic missile defense capability is with an overwhelming strike). Also, the deployment of the missile defense system may embolden adversaries: what is to deter them from firing nuclear warning shots if they will be shot down? Again, while the US asserts that the deployment of terminal interceptors in Europe cannot physically challenge Russian missiles and that the deployment of THADD in South Korea cannot surveil all of China, what really matters is the perception of Russian and Chinese leaders that these defenses could mitigate their nuclear missiles.[62]

Conversely, in accordance with Wohlstetter's stability attributes to ensure a costly counterstrike, nuclear-tipped hypersonic missiles will return the nuclear deterrent system between Russia, China, and the US to a condition of higher stability. Even with the additional first-strike and decapitation bolt-from-the-blue capability that hypersonics may be able to

provide, the other Wohlstetter attributes remain in play: a costly counterstrike guaranteed by submarine-launched nuclear weapons, airborne command posts, and possibly air-alerted nuclear-carrying bombers are hardly likely to be simultaneously destroyed provided a counterforce posture is maintained and deployed. See tables 2, 3, and 4 below to compare US, China, and Russian counterforce and stability attributes. (Note that the current state of stability can be uprooted by other technologies outside of BMD and hypersonic nuclear weapons that this article does not consider, which makes it of utmost importance to continue to modernize, conceal, and deploy robust counterforces.[63]) In other words, hypersonics, with their capability to defeat missile defense systems (whether perceived or actual), will return the US-Russia-China nuclear relationship to a state of assured vulnerability and a more stable strategic deterrence environment.

Table 2. US nuclear attributes

Nuclear Weapon System	Reliable, Affordable, Sustainable	Survivable	Credible Perception of Retaliation	Capable of Reaching Adversary	Penetrate Active Defenses	Destroy Target w/ Passive Defenses
ICBM	Positive	Positive	Positive	Positive	Positive	Positive
SLBM	Positive	Positive	Positive	Positive	Positive	Positive
Bombers	Positive	Negative	Positive	Positive	Positive	Positive

Table 3. Russia nuclear attributes

Nuclear Weapon System	Reliable, Affordable, Sustainable	Survivable	Credible Perception of Retaliation	Capable of Reaching Adversary	Penetrate Active Defenses	Destroy Target w/ Passive Defenses
ICBM	Positive	Positive	Positive	Positive	Positive	Positive
SLBM	Positive	Positive	Positive	Positive	Positive	Positive
Bombers	Positive	Negative	Positive	Positive	Negative	Positive

Table 4. China nuclear attributes

Nuclear Weapon System	Reliable, Affordable, Sustainable	Survivable	Credible Perception of Retaliation	Capable of Reaching Adversary	Penetrate Active Defenses	Destroy Target w/ Passive Defenses
ICBM	Positive	Positive	Positive	Positive	Positive	Positive
SLBM	Positive	Positive	Positive	Positive	Positive	Positive
Bombers	Positive	Negative	Positive	Positive	Negative	Positive

Notes: Nuclear attribute table design explanation: deterrence stability tables were developed using Wohlstetter's criteria when compared to each other, with each attribute scored on a basic scale: positive and negative. Positive scores are given for the regional system with attributes that add to deterrent stability. A negative score is given to an attribute based on evidence, logic, or questionable theory that detracts from stability when compared to its adversary. Each attribute is evaluated by itself (i.e., if the system was not found survivable, it may still possess attributes that allow it to penetrate defenses like stealth and be awarded "positive" for the penetrate defenses attribute). See appendix A (online at https://www.airuniversity.af.edu/) for a detailed overview of each county's nuclear capability in relation to Wohlstetter's attributes.

Table 5 matches nuclear countries with and without ballistic missile defense against nuclear countries with and without nuclear (N) hypersonic missiles. This table specifically addresses the Wohlstetter stability attribute of a country's ability to penetrate enemy defenses. It highlights that if a country cannot penetrate defenses or perceives that it cannot), then the rest of the attributes are largely nullified, and the overall deterrent system is unstable. The fundamental calculation used to determine whether a deterrent system was stable or unstable was whether the opposing countries could penetrate each other's defenses. If defenses for both County A and Country B could be penetrated, then the overall system is stable. If defenses could not be penetrated by either country—one country possessed a ballistic missile defense, and the opposing country did not have defense-penetrating nuclear hypersonic missiles in its inventory—then the overall system trends unstable since the guarantee of assured vulnerability is in doubt. Note that a country possessing nuclear hypersonic missiles is alone not a determining factor of whether a system is stable. Countries can possess hypersonic capabilities without upsetting the stability of the deterrent system. If anything, assured vulnerability is bolstered when both nuclear powers have nuclear-capable hypersonic missiles since these weapons have better capability to defeat defenses. The determining factor for stability is whether defenses can be penetrated and opposing countries can hold each other at risk with a robust counterforce capability, underwriting assured vulnerability.

Table 5. Strategic deterrent environment stability scenarios

		Nuclear Country B			
		Ballistic Missile Defense & No N-Hypersonic Missiles	**Ballistic Missile Defense & N-Hypersonic Missiles**	**No Ballistic Missile Defense & N-Hypersonic Missiles**	**No Ballistic Missile Defense & N-Hypersonic Missiles**
Nuclear Country A	No Ballistic Missile Defense & No N-Hypersonic Missiles	Situation G Increases Stability	Situation A Decreases Stability	Situation D Decreases Stability	Situation F Increases Stability
	Ballistic Missile Defense & No N-Hypersonic Missiles	Situation A Decreases Stability	Situation B Decreases Stability	Situation C Decreases Stability	Situation H Increases Stability
	Ballistic Missile Defense & N-Hypersonic Missiles	Situation D Decreases Stability	Situation C Decreases Stability	Situation E Increases Stability	Situation I Increases Stability
	No Ballistic Missile Defense & N-Hypersonic Missiles	Situation F Increases Stability	Situation H Increases Stability	Situation I Increases Stability	Situation J Increases Stability

The assumption in developing table 5 is that hypothetical countries A and B have a robust nuclear capability that satisfies Wohlstetter's other five attributes of reliability/affordability/sustainability and the ability to survive an enemy attack, reach enemy targets with enough fuel, destroy the target, and have effective retaliatory communication. Some may argue that nuclear-tipped hypersonic missiles could be used as a first-strike capability to nullify an opponent's counterstrike force, making situations F and J unstable. However, the assumption used in table 1 for hypersonic capability is much like ballistic missile submarines capability: both these systems could effectively be used in a first strike scenario against a country's nuclear capability, but each country's nuclear strike capability will still maintain an overwhelming counterstrike capability (sea, ground, and/or air) to validate Wohlstetter's stability attributes. This assumption is not valid in regional nuclear stability cases, addressed later in this article, where nuclear forces are relatively small and potentially vulnerable.

Situation A is unstable because one opponent has BMD while the other does not. Fundamentally, this situation violates the "assured vulnerability" criteria for Country B, putting Country A in a precarious position of returning the deterrent system to stability (arms race) or considering a first strike (nuclear or nonnuclear) to nullify the BMD.

Situation B is likely the most unstable of all the scenarios. In this case, both countries have BMDs and no hypersonic missiles to counter such defenses, putting both country's assured vulnerability in question. Both countries are questioning whether their nuclear strike capability is adequate to ensure a powerful counterstrike, with both considering a first strike to nullify each other's defenses and return the deterrent system to a more stable state.

Situations C and D are unstable since one country does not have nuclear hypersonics to nullify the opposing country's BMD. Again, assured vulnerability is not guaranteed in these scenarios.

Situations E, F, G, H, and I all are stable deterrent environments since one or the other country has a counter to BMDs. Further, in situations E, I, and J, both countries possessing nuclear hypersonic missiles keep the deterrent system in an "assured vulnerability" stable state whether BMDs are involved or not.

Nuclear Hypersonics—Regional Deterrence Implications

On the flip side of nuclear deterrence considerations, hypersonic missiles—nuclear or not—will have a destabilizing impact among regional nuclear powers. Fundamentally, this technology will exacerbate the exist-

ing regional nuclear imbalances in Wohlstetter's six attributes of stability. Situation A in table 5 capsulizes this current regional deterrent environment: India has a strong ballistic missile defense capability when compared to Pakistan's nuclear strike capability. However, the underlying assumptions from table 2 are not all applicable since both regional nuclear powers have caveats to their nuclear strike capability when compared to Wohlstetter's stability attributes and require further investigation (see tables 3 and 4). Nuclear deterrence between the regional powers of India and Pakistan relies largely on posture (keeping nuclear warheads disassembled[64] from their launchers and India's no-first-use policy[65]) rather than true Wohlstetter stability in their bilateral relationship. India clearly has a robust and resilient force with solid-fueled missiles (allowing for indefinite alert postures), BMDs, and a nuclear-capable ballistic missile submarine.[66] Further, India's nuclear force attributes, as outlined in table 6, clearly add to deterrent stability.

Table 6. India – Wohlstetter's nuclear attributes

Nuclear Weapon System	Reliable, Affordable, Sustainable	Survivable	Credible Perception of Retaliation	Capable of Reaching Adversary	Penetrate Active Defenses	Destroy Target w/ Passive Defenses
IRBM	Positive	Negative	Negative	Positive	Positive	Positive
SLBM	Positive	Positive	Positive	Positive	Positive	Positive
Bombers	Positive	Negative	Negative	Positive	Negative	Positive

Pakistan, on the other hand, relies primarily on the mobility of its nuclear-capable ballistic missiles for survivability and lacks a submersible, hard-to-locate nuclear capability. Using Wohlstetter's stability attributes, it is evident that this deterrent situation is unstable (table 7). India, with its missile defense system and maturing nuclear triad, can unmistakably weather a first strike from Pakistan and produce a crushing retaliatory nuclear response.[67] Without a credible air and ballistic defense combined with exposed nuclear delivery systems, the same cannot be said for Pakistan following a hypothetical nuclear first strike from India.

Table 7. Pakistan – Wohlstetter's nuclear attributes

Nuclear Weapon System	Reliable, Affordable, Sustainable	Survivable	Credible Perception of Retaliation	Capable of Reaching Adversary	Penetrate Active Defenses	Destroy Target w/ Passive Defenses
MRBM	Positive	Negative	Positive	Negative	Negative	Positive
SLCM	n/a	n/a	n/a	n/a	n/a	n/a
Bombers	Positive	Negative	Positive	Negative	Negative	Positive

India's development of hypersonic cruise or boost-glide missiles is only adding to the instability of the regional deterrent situation, pushing the environment into a situation D (table 5) scenario. Nuclear or not, an arsenal of perceived defense-defeating, first-strike capabilities can theoretically penetrate and eliminate much of Pakistan's nuclear force. However, in addition to Pakistan maturing its nuclear force to include SLBMs and solid-fuel rockets, deterrence stability would improve if Pakistan and India were to develop and procure hypersonic boost-glide or cruise missile capability. This increased stability correlates with Wohlstetter's penetrate-defense attribute. Ideally, it provides both Pakistan and India the capability to defeat antiballistic missile systems, putting both opponents in a stronger assured vulnerability state. Pakistan attaining hypersonic technology is not out of the question; the technology may be available for purchase from a current hypersonic producer, or Pakistan may develop it. In this regional situation, a potential proliferator of hypersonic technology is China because it sees Pakistan as a counterbalance to the India-US strategic relationship.[68]

Another hypothetical regional scenario to consider is the introduction of hypersonics to the Korean peninsula. There is no evidence to indicate that North Korea has a hypersonic missile program. Further, it does not possess a credible anti-BMD system that would require the use of the penetrating attributes of hypersonic missiles (likened to table 4, situation A) where the US is the opponent. North Korea does operate a dense, robust, aging (1960s–1970s era) air defense system, which would complicate fourth-generation warplane access in the event of a conflict.[69] However, this formidable but defeatable air defense system has little to no capability against a hypersonic glide vehicle that the US would likely use to target North Korea's emerging nuclear weapons program. A hypersonic first-strike attack is unlikely since stealth bombers and fighters can easily penetrate such air defenses at less expense than a $10 million hypersonic missile.[70] Regardless, due to the already profound asymmetric match of North Korea's nascent nuclear weapon systems when compared to US mature nuclear capabilities—to include the extended nuclear umbrella over Japan and South Korea—hypothetical nuclear-armed US hypersonic weapons are unlikely to alter this region's current nuclear deterrent dynamic (table 8).

Table 8. North Korea – Wohlstetter's nuclear attributes

Nuclear Weapon System	Reliable, Affordable, Sustainable	Survivable	Credible Perception of Retaliation	Capable of Reaching Adversary	Penetrate Active Defenses	Destroy Target w/ Passive Defenses
ICBM/ IRBM	Negative	Positive	Positive	Positive	Negative	Positive
SLCM	n/a	n/a	n/a	n/a	n/a	n/a
Bomber	n/a	n/a	n/a	n/a	n/a	n/a

Note: North Korea's deterrence stability table was developed using Wohlstetter's criteria when compared to the US, with each attribute scored on a basic scale: positive and negative. Positive scores are given for the regional system with attributes that add to deterrent stability. A negative score is given to an attribute based on evidence, logic, or questionable theory that detracts from stability when compared to its adversary. Each attribute is evaluated by itself (i.e., if the system was not found survivable, it may still possess attributes that allow it to penetrate defenses like stealth and be awarded "positive" for the penetrate defenses attribute). See appendix A (online at https://www.airuniversity.af.edu/) for a detailed overview of North Korea's nuclear capability in relation to Wohlstetter's attributes.

The dynamic changes considerably if North Korea acquired a nuclear hypersonic glide or cruise missile. The 2014 *Quadrennial Defense Review* names North Korea as one of two regional actors that missile defense is designed to deter and defeat.[71] A North Korea hypersonic capability would certainly erode any sense of protection a now operational US missile defense system is providing the West Coast. However, given the theoretical acquisition of nuclear-capable hypersonic weapons, the North Korea-US situation would turn more stable since North Korea could assure the vulnerability of the US and its allies. The assumption underpinning this statement is if North Korea's nuclear arsenal could somehow attain all six Wohlstetter attributes to truly realize this potential deterrent stability. However, North Korea has a long way to go in meeting Wohlstetter's stability attributes.

Another hypothetical case to consider is hypersonic weapons in the Middle East. If the undeclared nuclear power of Israel were to procure hypersonic capability (nuclear or conventional), the instability would remain the same since there are no other opposing nuclear powers in the region (table 4, situations A to D). However, if Iran obtained a nuclear hypersonic missile capability, the deterrent environment would turn more stable (situation H) (assuming Israel only possesses BMD and Iran refrains from proliferating this theoretical acquisition of nuclear hypersonic technology to its proxy forces and terrorists in the region). This newly acquired Iranian capability would increase stability since Iran could theoretically penetrate Israeli missile defenses, putting some of Israel's nuclear capability, conventional forces, and general population at a perceived higher risk, achieving assured vulnerability, fundamentally deterring Israel from striking Iran.[72]

Transition Instability

A period of increased instability will occur during the phase in which nuclear hypersonics become operational. This turbulence will peak as one nuclear country deploys hypersonic weapons while others are still in developmental stages. Once this occurs, nuclear powers without hypersonic capability will perceive a disadvantage and be more vulnerable to a strike from the nation with the defense-penetrating capability. During this time, the disadvantaged power will contemplate and recalculate its options, deciding whether a first strike is warranted because of its perceived vulnerability. As Thomas Schelling stated, "Vulnerable strategic weapons not only invite attack but in a crisis could coerce the . . . government into attacking when it might prefer to wait."[73] Therefore, until opposing powers share the same vulnerabilities and/or comply with Wohlstetter's stability criteria, the mismatch in nuclear attributes will promote instability. Additionally, when competing countries possess ballistic missile defenses and no defense-penetrating capabilities (table 4, situation B), instability will rumble through the nuclear deterrent paradigm: assured vulnerability is completely undermined with neither country convinced it could launch a credible counterstrike. Therefore, as a counter to ballistic missile defenses, hypersonic weapons are a natural evolution in nuclear deterrent systems; they should be anticipated and expected to bring back true assured vulnerability. The danger lies during the transition to assured vulnerability and should be managed in a manner that minimizes risk from the absence of BMD and hypersonics.

Conclusion and Policy Considerations

Two solutions exist to mitigate the impending problems of nuclear hypersonic missiles: dismantle the ground-based missile defense program and add hypersonic specific technology to a nonproliferation ban. As identified here, missile defense is the primary reason for increased nuclear instability and the impetus for the development of nuclear-tipped hypersonic weapons. Specifically, the ground-based midcourse defense system established in Alaska and the West Coast of the US has undermined assured vulnerability by degrading Russia's and China's ability to hold targets in the Western Hemisphere at risk. As Wohlstetter asserts, mutual assured vulnerability is critical to a stable nuclear stalemate. Therefore, the best policy to stabilize the nuclear deterrence environment among the three great nuclear powers (China, Russia, and the US) is to dismantle continental missile defenses and discontinue further development since missile defense is underwriting the emerging hypersonic arms race.

Stabilizing regional nuclear deterrence is a more difficult problem. Many of the regional nuclear powers possess a form of BMD that has at least some capability against each other's nuclear delivery systems. As stated earlier, the natural evolution of these regional nuclear standoffs is to develop nuclear weapon systems that can defeat missile defenses. Consequently, opposing regional nuclear powers will want to develop hypersonic weapons. The dangerous transition periods will emerge during the unbalance of capability—when only one opposing nuclear power attains this hyper-fast, defense-penetrating weapons. During this window, tensions will heighten while adversaries weigh their options, including a preemptive strike to remove this hyper capability from their foe, a heightened nuclear alert posture that sows seeds for a crisis, and/or an increase in bellicose rhetoric. There may well be two options to prevent the regional hypersonics dilemma: quickly proliferate hypersonic technology to all nuclear powers, or add hypersonic technology to the Hague Code of Conduct against Ballistic Missile Proliferation and/or the Missile Technology Control Regime to halt hypersonic proliferation. One option is highly improbable and laced with danger (exporting high-tech weapons to adversaries). The other requires action before the window of technology maturation closes and it is too late to prevent calamity.

After unpackaging hypersonic capabilities and deterrence theory, it should be apparent how maturing hypersonics technology will impact nuclear deterrence. Mature nuclear powers will experience an era of greater strategic stability since each will be more vulnerable to each other's hypercapability, solidifying the desired nuclear stalemate. Regional nuclear powers that develop nuclear hypersonic capability will incite regional instability since one power will have the advantage of the first strike to disable/destroy its opponents' retaliatory nuclear capability. In the long term, regional stability should increase—assuming hypersonic technology proliferates quickly among these nascent nuclear powers and their respective nuclear capability matures to provide a guaranteed costly, retaliatory strike. Overall, the period when this technology is not evenly distributed among nuclear powers will be the highest period of instability. The world is at a crossroads: it can stand by and let the introduction of hypersonic technology sneak in and induce nuclear instability, or it can take action by limiting the export of hypersonic technology and eliminating missile defense. SSQ

Col Stephen Reny, USAF

Colonel Reny is an instructor in the Military and Strategic Studies Department at the US Air Force Academy. He is a former National Defense Fellow at the Woodrow Wilson International Center for Scholars, where he studied nuclear deterrence theory. Colonel Reny is a command pilot with previous assignments at US Central Command, Air Combat Command, and the Air Staff.

Notes

1. Keith B. Payne, *Deterrence in the Second Nuclear Age* (Lexington: The University Press of Kentucky, 1996), 148.

2. Gregory Kulacki, *China's Military Calls for Putting Its Nuclear Forces on Alert* (Cambridge, MA: Union of Concerned Scientists, January 2016), https://www.ucsusa.org/.

3. Congressional Research Service (CRS), *Hypersonic Weapons: Background and Issues for Congress* (Washington, D.C.: CRS, 17 March 2020), updated 27 August 2020, https://crsreports.congress.gov/, 17.

4. Andrew Osborn, "Putin, before Vote, Unveils 'Invincible' Nuclear Weapons to Counter West," 1 March 2018, Reuters, https://www.reuters.com/.

5. Richard Stone, " 'National Pride Is at Stake': Russia, China, United States Race to Build Hypersonic Weapons," *Science Magazine*, 8 January 2020, https://www.sciencemag.org/.

6. Congressional Research Service, *Hypersonic Weapons*, 1.

7. Kyle Mizokami, "Could Hypersonic Weapons Make Nuclear War More Likely?," *Popular Mechanics*, 12 October 2017, https://www.popularmechanics.com/.

8. While the brief citations of Schelling's work in this article may miss some of the nuance of his ideas, the author chose Wohlstetter's six attributes as the best possible criteria to judge the effects and implications of new technology on deterrence. Obviously, many other nuclear deterrence scholars, including Colin Gray and Herman Kahn, have profound ideas on deterrence that are outside the scope of this argument.

9. Albert J. Wohlstetter, "The Delicate Balance of Terror," *Foreign Affairs* 37, no. 2 (1959): 216.

10. Lawrence Freedman, *The Evolution of Nuclear Strategy*, 3rd ed. (New York: Palgrave Macmillan), 245, 255. Freedman describes in his lesson for arms control the cause and effect of improved defenses or offenses, how it spurs instability and an arms race. He states, "The lesson drawn for arms control was that, as every improvement on one side's defense provided no extra security but merely a spurt to the offence of the other, once both sides ceased making defensive moves forces could stabilize at current levels" (245). He also describes how, once the USSR realized the strategic upper hand of the US in nuclear capability, the USSR aggressively pursued a "major military build-up" to catch up with the US (255).

11. Thomas C. Schelling, *The Strategy of Conflict* (New York: Oxford University Press, 1973), 232.

12. Fred M. Kaplan, *The Wizards of Armageddon* (Stanford: Stanford University Press, 1991). In chaps. 9 and 10, Kaplan points out how the 1957 launch of *Sputnik* stoked the already existing fears of a missile gap. Kaplan argues that these fears were largely founded in faulty intelligence reports that supported a general intellectual consensus (fed by the Gaither Report and Wohlstetter's R-290 report) that the USSR was gaining on the US with superior nuclear deterrence technology, numbers of bombers, and estimates they would have hundreds of ICBMs by 1960. Along with these reports, Sputnik's launch reinforced justification for accelerated procurement if US ICBMs.

13. Kaplan, 160.

14. Stephan Kieninger, "Diverting the Arms Race into the Permitted Channels: The Nixon Administration, the MIRV-Mistake and the SALT Negotiations," Nuclear Proliferation International History Project (NPIHP) Working Paper No. 9, The Woodrow Wilson International Center for Scholars, November 2016, https://www.wilsoncenter.org/. MIRVs essentially would overwhelm any defensive system with thousands of individual nuclear warheads.

15. Kieninger.

16. Kieninger.

17. See Lt Col Nathan B. Terry, USAF, and Paige Price Cone, "Hypersonic Technology: An Evolution in Nuclear Weapons?," *Strategic Studies Quarterly* 14, no. 2 (Summer 2020): 76, https://www.airuniversity.af.edu/.

18. Tom Benson, "Speed Regimes: Low Hypersonic," Glenn Research Center, US National Aeronautics and Space Administration, accessed 9 January 2017, https://www.grc.nasa.gov/.

19. Tom Benson, "Speed Regimes: Hypersonic Re-Entry," Glenn Research Center, US National Aeronautics and Space Administration, accessed 9 January 2017, https://www.grc.nasa.gov/.

20. John D. Gresham, "S-400 Triumf/SA-21 Growler: New Threat Emerging," Defense Media Network, 8 April 2011, http://www.defensemedianetwork.com/.

21. John D. Anderson, Jr., *Modern Compressible Flow: With Historical Perspective*, 3rd ed. (Boston: McGraw-Hill, 2003), 19.

22. Anderson, 19.

23. William H. Mason, professor, Virginia Tech, "Hypersonic Aerodynamics," chap. 11 in "Configuration Aerodynamics," self-developed textbook for course AOE 4124, accessed 20 January 2017, http://www.dept.aoe.vt.edu/.

24. NASA, "Speed Regimes: Hypersonic Re-Entry." Zirconia was used in the heat shield of the X-43A. T. A. Heppenheimer, *Facing the Heat Barrier: A History of Hypersonic Flight* (Washington, D.C.: National Aeronautics and Space Administration, 2007), 269.

25. Luu Hong Quan et al., "Analysis and Design of a Scramjet Engine Inlet Operating from Mach 5 to Mach 10," *International Journal of Mechanical Engineering and Applications* 4, no. 1 (2016): 11–23. This article describes a *scramjet* as a "supersonic combustion ramjet, a variant of the ramjet engine cycle. Different from other types of air breathing engine like turbojet and turbofan, [the] ramjet engine doesn't rely on turbo machinery but shockwave for compression." For a scramjet, supersonic air is compressed by shockwave "before entering the combustor where fuel is injected and burnt. Air is then accelerated through a nozzle to create thrust." Quan et al., 11–23.

26. The US Army's Advanced Hypersonic Weapon, a hypersonic boost-glide vehicle, has been launched without leaving the atmosphere. The point here is that a hypersonic boost-glide weapon does not have to leave the atmosphere, but longer ranges (intercontinental) will likely require higher exoatmospheric boosts. John Reed, "Army Successfully Tests Hypersonic Weapon Design," *DefenseTech*, 17 November 2011, http://defensetech.org/.

27. James M. Acton, *Silver Bullet? Asking the Right Questions about Conventional Prompt Global Strike* (Washington, D.C.: Carnegie Endowment for International Peace, 2013), 35, https://carnegieendowment.org/. This slowing descent is indicative of the severe aerodynamic loading experienced at lower altitudes primarily due to higher air densities at lower altitude. High-altitude hypersonic flight is desired to limit these loads (pressures) and simplify design efforts to focus on the extreme thermal management issues. Barry Hellman, Air Force Research Laboratory, Aerospace Systems High-Speed Division engineer (interview by the author, 10 February 2017).

28. Acton, *Silver Bullet?*

29. The Minotaur IV Lite boost vehicle (formerly the booster for the Peacekeeper ICBM) was used by Hypersonic Test Vehicle 2, which could carry up to 1,730 kg and produces 500,000 lbs of thrust at launch. "Minotaur IV," GlobalSecurity.org, accessed 23 January 2017, http://www.globalsecurity.org/. The space shuttle (74,800 kg) used two solid-rocket boosters, each producing 3.3 million pounds of thrust at launch. NASA, Kennedy Space Center, "Solid Rocket Boosters," accessed 23 January 2017, https://science.ksc.nasa.gov/. Excerpted from National Space Transportation System (NSTS) News Reference Manual handed out to the press in September 1988. Logic follows that larger booster equals larger hypersonic aircraft, but the limiting factor will likely be cost, which gives the Minotaur (50 or so available in US stock, per "Minotaur IV," GlobalSecurity.org) the edge in being the hypersonic booster of choice due to its ready availability.

30. Acton, *Silver Bullet?*, 80.

31. Mason, "Hypersonic Aerodynamics," 25.

32. Arms Control Association, "The Intermediate-Range Nuclear Forces (INF) Treaty at a Glance," fact sheet, accessed 25 January 2017, https://www.armscontrol.org/. The INF Treaty specifically banned "ground-launched ballistic and cruise missiles with ranges of 500 to 5,500 kilometers." Arms Control Association, fact sheet. Australian researchers, however, have ground-launched hypersonic air-breathing vehicles (with the point being that this mode of launch is possible. Heppenheimer, *Facing the Heat Barrier*, 275.

33. Joseph Cirincione, *The Ballistic Missile Threat*, testimony (Washington, D.C.: Carnegie Endowment for International Peace, 2001), http://carnegieendowment.org/.

34. US hypersonic programs launch air-breathing hypersonic vehicles from a B-52 bomber or L-1011 via boosters are advertised to carry approximately 1,000–1,240 lbs of payload. NASA Armstrong, "Hyper-X Program," fact sheet, 28 February 2014, https://www.nasa.gov/; Orbital ATK, "Pegasus Patented Air Launch System," fact sheet, 2017, https://satsearch.co/; and Florian Ion Petrescu and Relly Victoria Petrescu, *New Aircraft II* (Norderstedt, Germany: Books on Demand GmbH, 2012), 110, http://dx.doi.org/10.13140/RG.2.1.2584.6480. The X-51A Waverider and booster weigh approximately 4,000 lbs (1,800 kg) (the MGM-140 Army Tactical Missile System [ATACMS] booster weighs 2,900–3,600 lbs [1,300 – 1,670 kg] alone), with a scramjet fuel capacity of ~270 lbs (122 kg). When readied for launch, the full stack (booster and vehicle) is 25 feet in length. Andreas Parsch, "MGM-140," in *Directory of U.S. Military Rockets and Missiles*, accessed October 2020, http://www.designation-systems.net/; and US Air Force, "X-51A Waverider," fact sheet, 2 March 2011, https://www.af.mil/. In this configuration, the Waverider advertises a 240-second scramjet burn at Mach 5.1 (~4,000 miles per hour), giving it a range of ~266 miles (428 km). Assuming the same flight time at Mach 10 (~7,610 mph), this would equate to range of 507 miles (815 km). These powered-cuise distances do not include boost and glide time. Assuming expected glide characteristics of a boost-glide vehicle (likely much less due to low lift/drag design of the hypersonic cruise vehicles), I use the 4 to 1 ratio mentioned earlier), assuming 100,000 feet operating altitude, to calculate an X-51 glide range of 75 miles (120 km)—which brings a max range of ~582 miles (936 km). James Acton, "Hypersonic Boost-Glide Weapons," *Science and Global Security* 23, no. 3 (September 2015): 191–219, https://doi.org/10.1080/08929882.2015.1087242. With boost phase included, the X-43 yper-X has an expected range of 850 miles (1,367 km). NASA, "Hyper-X Program."

35. Air Force acquisition official (interview by the author, January 2017).

36. Acton, *Silver Bullet?*, 65.

37. Acton, 156. Review "How a Missile Only Becomes Visible to a Powerful Radar Once It Has Passed through the Radar's Horizon" for a description of missile detection.

38. John D. Anderson, *Hypersonic and High Temperature Gas Dynamics*, 2d ed. (Reston, VA: American Institute of Aeronautics and Astronautics, Inc., 2006), 18.

39. Anderson, 73. Additionally, Dr. Kopp addressed an upgraded S-300 system that boosted tactical antiballistic missile capability comparable to early models of the Patriot system. Carlo Kopp, *Almaz S-300P/PT/PS/PMU/PMU1/PMU2 Almaz-Antey S-400 Triumf SA-10/20/21 Grumble/Gargoyle*, Technical Report APA-TR-2006-1201, Air Power Australia, December 2006, http://www.ausairpower.net/.

40. Carlo Kopp, *Almaz-Antey 40R6 / S-400 Triumf: Self Propelled Air Defence System / SA-21*, Technical Report APA-TR-2009-0503, Air Power Australia, May 2009, http://www.ausairpower.net/.

41. Aerospace Daily & Defense Report, "S-300 Surface-To-Air Missile Systems," *Aviation Week*, 6 August 2015, 6–10; and Carlo Kopp, *Almaz-Antey S-500 Triumfator M Self Propelled Air / Missile Defence System* / SA-X-NN, Technical Report APA-TR-2011-0602, Air Power Australia, June 2011, www.ausairpower.net/.

42. Mark Episkopos, "Introducing Russia's S-500: Can It Take Out an F-22 or F-35?," *The National Interest*, 18 May 2019, https://nationalinterest.org/. Episkopos speculates that the S-500 will proliferate through China and Russia sales.

43. Acton, *Silver Bullet?*, 67.

44. Hellman, interview. Mr. Hellman explained how the air density at lower altitudes creates high dynamic pressures that lead to overpressures in the scramjet, ultimately rendering it inoperative.

45. Charles R. McClinton, "X-43–Scramjet Power Breaks the Hypersonic Barrier: Dryden Lectureship in Research for 2006," 44th American Institute of Aeronautics and Astronautics (AIAA) Aerospace Sciences Meeting and Exhibit, Reno, Nevada, 9–12 January 2006, published online 21 June 2012, https://doi.org/10.2514/6.2006-1; and Acton, *Silver Bullet?*, 77. Acton states

about a video of an HTV-2 test that "the HTV-2 produced so much heat that it is actually visible to the naked eye." Acton, *Silver Bullet?*, 77.

46. Kopp, *Almaz-Antey 40R6 / S-400 Triumf*. Dr. Kopp notes that the S-400 interceptor employs "canards and thrusters to achieve extremely high G and angular rate capability throughout the engagement envelope."

47. William Yengst, *Lightning Bolts: First Maneuvering Reentry Vehicles* (Mustang, OK: Tate Publishing & Enterprises, LLC, 2010), 159–67.

48. Yengst, 167.

49. National Air and Space Intelligence Center (NASIC), *Ballistic & Cruise Missile Threat*, NASIC-1031-0985-13 (Wright-Patterson Air Force Base, OH: NASIC, 2013), 18, https://fas.org/. See also James M. Acton, "Prompt Global Strike: American and Foreign Developments," testimony before the House Armed Services Subcommittee on Strategic Forces, Washington, D.C., 8 December 2015, Carnegie Endowment for International Peace, https://carnegieendowment.org/; and Minnie Chan, "China's latest hypersonic vehicle test seen as 'nuclear deterrent' amid US interference," *South China Morning Post,* 13 June 2015, http://www.scmp.com/.

50. NASIC, *Ballistic & Cruise Missile Threat*, 18.

51. Kelsey Davenport, "India Tests Hypersonic Missile," Arms Control Association, October 2020, https://www.armscontrol.org/

52. Rahul Singh, "India's Tribute to Missile Man: New BrahMos Gets Kalam Name," *Hindustantimes*, 8 August 2015, http://www.hindustantimes.com/.

53. Quoted in Joseph Cirincione, "Brief History of Ballistic Missile Defense and Current Programs in the United States," testimony, 1 February 2000, Carnegie Endowment for International Peace, https://carnegieendowment.org/.

54. Department of Defense, *Quadrennial Defense Review 2014* (Washington, D.C.: Department of Defense, 2014), 14, https://archive.defense.gov/; and Department of Defense, *Missile Defense Review 2019* (Washington, D.C.: Department of Defense, 2019), ii, https://www.defense.gov/.

55. Herman Kahn, "Twelve Nonissues and Twelve Almost Nonissues," Hudson Institute, 1 January 1984, https://hudson.org/.

56. Herman Kahn, *On Thermonucler War* (New Brunswick: Transaction Publishers, 2007), 519.

57. Both Lt Gen Glenn Kent, USAF, an analyst and developer of American defense policy, and Schelling discuss "first-strike stability" where neither force has an advantage in striking first. Ballistic missile defenses put any advesary at a "first strike" disadvantage, sowing instability. Schelling states, "Thus ABM systems deployed in both countries would make preemptive war more likely and the arms race more expensive." See Glenn A. Kent et al., *Thinking About America's Defense: An Analytical Memoir* (Santa Monica, CA: RAND Corporation, 2008), 73–80, https://www.rand.org/; and Thomas C. Schelling, "What Went Wrong with Arms Control?," *Foreign Affairs* 64, no. 2 (Winter 1985–86): 219–33.

58. Schelling, 219–33.

59. For US speculation on Chinese development of a nuclear hypersonic glider, see United States-China Economic and Security Review Commission, "China's Offensive Missile Forces," Hearing before the US - China Economic and Security Review Commission, 114th Cong., 1st sess., Washington, D.C., 1 April 2015, transcript, https://www.uscc.gov/, 1. For further speculation on Russian and Chinese nuclear hypersonic gliders, see Bill Gertz, "Russia and China Are Building Hypersonic Missiles and It's 'Complicating' Things for the US," *Business Insider*, 30 July 2015, https://www.businessinsider.com/; Lora Saalman, *Factoring Russia into The US-Chinese Equation on Hypersonic Glide Vehicles* (Solna: Stockholm International Peace Research Institute, 2017), 6, https://www.sipri.org/; and Department of Defense, *Missile Defense Review 2019*, iii.

60. Wohlstetter, "Delicate Balance of Terror," 213.

61. For specifics on this point, see Missile Defense Agency, "The Ballistic Missile Defense System," fact sheet, January 2018, https://www.mda.mil/; Missile Defense Agency, "Elements: Terminal High Altitude Area Defense (THAAD), accessed 9 February 2017, https://www.mda.mil/; and

Congressional Research Service, *Navy Aegis Ballistic Missile Defense (BMD) Program: Background and Issues for Congress* (Washington, D.C.: Congressional Research Service, 13 October 2020), https://fas.org/.

62. Dean A. Wilkening, "Does Missile Defence in Europe Threaten Russia?," *Survival* 54, no. 1 (2012): 31–52, https://doi.org/10.1080/00396338.2012.657531. For analysis of Chinese rhetoric against THAAD deployment in South Korea, see Jaganath Sankaran and Bryan L. Fearey, "Missile Defense and Strategic Stability: Terminal High Altitude Area Defense (THAAD) in South Korea," *Contemporary Security Policy* 38 no. 3 (2017): 321–44, https://doi.org/10.1080/13523260.2017.1280744; and Eric Gomez, "Why Putin Is Obsessed with America's Missile Defense," CATO Institute, 3 March 2018, https://www.cato.org/.

63. Keir A. Lieber and Daryl G. Press, "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence," *International Security* 41, no. 4 (Spring 2017): 9–49, https://www.belfercenter.org/. In this article, the authors scan sensing technologies that could compromise locations of submarines, possibly jeopardizing the one nuclear triad counterforce capability that is considered the most concealed and invulnerable of the nuclear triad.

64. Until recently, many speculated that Pakistan and India kept their warheads unmated or disassembled from their delivery systems. David J. Karl, "Pakistan's Evolving Nuclear Weapon Posture," *The Nonproliferation Review* 21, no. 3-4 (2014): 317–36, http://dx.doi.org/10.1080/10736700.2014.1072998. However, some literature speculates that some of India's ballistic systems are sealed with a mated warhead. With India's no-first-use policy, this posture change does not alter my conclusion of reliance on posture rather than true stability as outlined by Wohlstetter. Vipin Narang, "Five Myths about India's Nuclear Posture," *Washington Quarterly* 36, no. 3 (2014):143–57, https://doi.org/10.1080/0163660X.2013.825555.

65. Rajesh Rajagopalan, "India: The Logic of Assured Retaliation," in *The Long Shadow: Nuclear Weapons and Security in 21st Century Asia*, ed. Muthiah Alagappa (Stanford: Stanford University Press, 2008), 196.

66. "INS Arihant: Nuclear Submarine Not "Fully Ready" for Patrols Carrying Nukes," *India.com*, 18 October 2016, http://www.india.com/.

67. In April 2016, it was reported that India finalized a deal with Russia to purchase 12 battalions of S-400 integrated air defense systems, which have missile defense capabilities. See Naveed Ahmad, "Analysis: Will India's S-400 missiles checkmate Pakistan?," *Express Tribune,* 29 April 2016, https://tribune.com.pk/. Additionally, India's missiles are all solid-fueled rockets, allowing them to remain on alert indefinitely—a clear advantage over Pakistan's liquid-fueled rockets that can remain on alert for only a day or so once fueled.

68. In the book section of *The Long Shadow*, Kahn states that China is "the only major power that sees the utility of a nuclear Pakistan as a balancer against India." However, there is no indication at this time that China is providing, or plans to provide, hypersonic technology to Pakistan. Feroz Hassan Khan and Peter R. Lavoy, "Pakistan: The Dilemma of Nuclear Deterrence," in Alagappa, *The Long Shadow*, 233.

69. John Reed, "What Do North Korea's Air Defenses Look Like?," *Foreign Policy*, 1 April 2013, http://foreignpolicy.com/.

70. In an interview, an Air Force acquisition official for hypersonic programs stated, "Each [conventional] hypersonic weapon costs 10–15 times less than a surface-to-air missile system" it would destroy. The official noted this was based on a $130M estimate per SAM site and $10M per production copy of a hypersonic missile. Air Force acquisition official, interview.

71. Department of Defense, *Quadrennial Defense Review 2014*, 14–32.

72. For more information on Israeli missile defenses, see Ruth Eglash and William Booth, "Israel to Launch One of the Most Advanced Missile Defense Systems in the World, with US Help," *Washington Post*, 3 March 2016, https://www.washingtonpost.com/.

73. Schelling, *Arms and Influence*, 233.

PERSPECTIVE

Space Traffic Management in the New Space Age

Brian G. Chow

Abstract

Since 2018, at least 11 US space officials and intelligence agencies at the highest level have expressed serious concerns about the threat from dual-use rendezvous and proximity operations. Yet the United States and the world are still not prepared for this rapidly approaching threat. However, its destabilizing nature is prone to turn a crisis into a war. This article analyzes the characteristics of the proximity threat and identifies opportunities—whether technical, economic, or political—to resolve the problem. The United States should declare that it will enact self-defense or warning zones and enforce them with bodyguard spacecraft and urge other countries to do the same. It should lead the way in pursuing a Western space traffic management (STM) system and an international STM version, both of which will have zones and bodyguards. Additionally, the West should offer China and Russia access to Western space markets and technical know-how if they abide by the zone/bodyguard rules under Western STM. The natural consequence would be for all countries to join the international STM system, as both regimes have virtually identical rules.

*****

In November 2015, the U.S.-China Economic and Security Review Commission released its annual report to Congress stating that "since 2008, China has . . . conducted increasingly complex tests involving spacecraft in close proximity to one another."[1] It added, "China is setting a strong foundation for future co-orbital anti-satellite systems that could include jammers, robotic arms, kinetic kill vehicles, and lasers."[2] Two and half years later, in a surge of government statements between June 2018 and February 2020, at least 11 space officials and intelligence agencies at the highest level expressed serious concerns about this proximity threat. In August 2018, Vice President Pence did not mince words: "Both China and Russia have been conducting highly sophisticated on-orbit activities that could enable them to maneuver their satellites into close proximity of ours, posing unprecedented new dangers to our space systems."[3] The other 10 government sources sounded similar alarms.[4] On 27 July 2020, the

United States and Russia held their first space security talks since 2013. Trump administration officials hoped that these talks would lead to a set of voluntary norms for operating in space.[5] The next day, US Space Command nominee Army lieutenant general James Dickson echoed during a confirmation hearing that "norms of behavior" should be established for the space domain.[6]

The proximity threat has unique characteristics that can make traditional remedies ineffective. China, Russia, the United States, the European Union, and others have planned to deploy spacecraft[7] capable of rendezvous and proximity operations (RPO) during the first half of the 2020s, if they have not already done so. Many of these robotic spacecraft will be used to refuel, repair, and upgrade satellites already in orbit and to remove or reposition space debris.[8] However, these spacecraft are dual use. If a robotic spacecraft can grapple space debris, it can also grapple another country's satellite. Russia and China have been proposing to keep peace in space by prohibiting the placement of any weapons there.[9] As these dual-use spacecraft can readily turn into antisatellite weapons (ASAT), weapons will soon be present in space. In the new age of proximity operations, banning dual-use robotic spacecraft is not desirable. The United States, as well as other countries, should learn how to use space traffic management (STM) as a key instrument to deter and defend against these potentially threatening spacecraft.

The 18 June 2018 Space Policy Directive-3, National Space Traffic Management Policy, states that "we must develop a new approach to space traffic management" that must "incorporate national security considerations" and "encourage growth of the US commercial space sector."[10] It also emphasizes that "the contested nature of space is increasing the demand for DOD focus on protecting and defending US space assets and interests."[11] The more recent *Defense Space Strategy Summary* of June 2020 reemphasizes that the DOD will "deter aggression in space" and "support US leadership in space traffic management."[12] While both documents state that STM must keep peace and foster prosperity, a widespread presumption is that the Space Force should focus on space security and STM on economic prosperity. Unfortunately, hostile and legitimate RPOs can be indistinguishable. An adversary can use dual-use spacecraft to hide coorbital ASAT attacks under the guise of peacetime maneuvers in proximity of our critical satellites. Currently, international law does not prevent a nation from stalking another country's satellites.[13]

Western and international STM systems, still in their early stage of development, can be designed to resolve the threat of dual-use proximity

operations while providing economic prosperity.[14] This article first describes why STM is needed to protect against the proximity threat. Then it proposes two core measures: self-defense zones and bodyguard spacecraft for STM to deter and defend against the proximity threat. Third, it designs a dual-track approach to pursue Western and international STM in parallel. Next, it identifies incentives to attract China and Russia to participate in STM. Finally, it recommends a strategy for STM that maintains peace and supports economic prosperity.

The Necessity for Space Traffic Management

In 2018, the Long Term Sustainability (LTS) Working Group of the Committee on the Peaceful Uses of Outer Space (COPUOS) tried to establish voluntary "measures for the safe conduct of proximity space operations."[15] Russia blocked adding these RPO measures to the 21 guidelines developed by the working group over the prior eight years.[16] Finally, in June 2019, Russia endorsed the 21 guidelines, but RPO rules were not included. While these guidelines will help avoid accidental collisions of functional satellites with space debris, they will not prevent satellites from being deliberately threatened or disabled by robotic spacecraft.

Even if Russia and China agreed to reconsider RPO measures, there is another problem. COPUOS has long focused only on guidelines for commercial safety, not military security. Taking advantage of this tradition, Russia and China could steer RPO guidelines toward helping commercial operators avoid accidental collisions but leaving the option of using proximity operations to threaten critical US military satellites. This threat could be a powerful instrument for executing their asymmetric strategies to counterbalance the more superior US military capabilities in space. For example, in its 2019 document *China Military Power*, the US Defense Intelligence Agency states, "PLA [People's Liberation Army] writings emphasize the necessity of 'destroying, damaging, and interfering with the enemy's reconnaissance . . . and communications satellites,' suggesting that such systems, as well as navigation and early warning satellites, could be among the targets of attacks designed to 'blind and deafen the enemy.'"[17]

Such an attack would be most damaging if it is the fateful opening of a war in space or on Earth. China could pre-position and maintain multiple dual-use robotic spacecraft arbitrarily close to our critical satellites. Even more worrying is that this threat will grow. Sometime in the latter half of the 2020s, China will have the capability to pre-position dozens of cheap RPO small satellites (smallsats[18]) close to dozens of our satellites, such as the Global Positioning System (GPS). Although these spacecraft are slow-

moving, they will be able to legally pre-position during peacetime and get unreasonably close. After "legitimately" setting up this threatening posture, China would have an advantage in a crisis, such as one involving Taiwan. If the US intervenes, China could disable critical satellites so quickly that we would not have enough time to defend them. The disabling could severely degrade US war-fighting capabilities. Furthermore, knowing an intervention could fail, the US might decide not to intervene in the first place and would risk its credibility among allies.[19] The US could prevent such a threat scenario and outcome by creating and enforcing a more comprehensive STM regime that provides timely warning and prevention.

Already, "rumors have been circulating for years that the Chinese Communist Party (CCP) has developed small satellites with robotic arms that could be used as anti-satellite weapons." The rumors indicate that "some of the smaller satellites are lighter than 22 pounds, yet have a triple-eye sensor to gauge the shapes of targets and can adjust their speed and rotation, allowing them to grab objects within a distance of six inches, using a single robotic arm."[20] Considering their significant research and development in RPOs and smallsats,[21] China as well as Russia can likely deploy a few attackers in the first half of the 2020s and then, in the second half of the decade, dozens of inexpensive smallsats capable of RPOs to mount a simultaneous proximity attack. These proximity ASATs would have a cost ratio (e.g., millions each for ASATs versus hundreds of millions each for a victim's satellites) highly favorable to the attacker. It would be even more favorable to the attacker if one includes the high cost to the victim of losing the services provided until its satellite capability is fully replaced. Constellations of even dozens of satellites could still be vulnerable. For example, the 32 GPS III satellites, which will replace the current GPS by 2025, cost about half a billion dollars each.[22] Dozens of cheap, robotic ASATs could defeat most of these 32 satellites, degrading or eliminating a critical service needed in peacetime and wartime.

It is important to note that existing space treaties focus heavily on commercial and not military space. For example, the Liability Convention allows up to three and a half years for compensation after a satellite is damaged.[23] While this may be satisfactory for settling commercial disputes, compensation is not a key goal in military space. The military objective is the survival of national technical capabilities in space for peacetime and wartime operations. Therefore, the Department of Defense has a great interest in STM to prevent hostile proximity operations.

Moreover, unless the DOD plays a more active role in steering STM to deal with satellite security, not just safety, the Consortium for Execution

of Rendezvous and Servicing Operations (CONFERS)—an industry-led initiative—may reinforce the wrong notion that STM should focus on commercial activities only. CONFERS aims to "leverage best practices from government and industry to research, develop, and publish non-binding, consensus-derived technical and safety standards that servicing providers and clients for on-orbit servicing operations would adopt."[24] It recommended design and operational practices on 1 February 2019 that echo Space Policy Directive-3 in stating that "specific techniques [for spaceflight safety] may include passive safe orbits, safety zones, and keep-out spheres or volumes for RPO and OOS [on-orbit servicing] activities."[25] While mention of safety zones sounds promising, Brian Weeden, executive director of CONFERS, is noncommittal toward self-defense zones.[26] Another concern is that an adversary may not follow the CONFERS "non-binding, consensus-derived" standards during times of crisis or war. Like COPUOS, this industry-led initiative is likely to favor economic prosperity over military security. The DOD must pursue self-defense zones and bodyguard spacecraft within expanded STM to deter and defend against the RPO threat.

Deterring and Defending against the Proximity Threat

While RPO spacecraft cannot be banned, their dual-use threat can be eliminated by prohibiting close proximity operations without prior consent. A self-defense zone can be used for timely alert to indicate whether a spacecraft is too close, and a bodyguard spacecraft can provide the necessary protection.

Self-Defense Zones

The first and most important space treaty, the 1967 Outer Space Treaty, does not mention self-defense zones or similar measures. However, Article IX of the treaty says, "States Parties to the Treaty shall be guided by the principle of cooperation and mutual assistance and shall conduct all their activities in outer space, including the Moon and other celestial bodies, with due regard to the corresponding interests of all other States Parties to the Treaty."[27] It also states that the principle of due regard should be used to prevent "potentially harmful interference."[28]

More than five decades later, in May 2020, NASA released the Artemis Accords concerning the Moon, which propose the following:

> Avoiding harmful interference is an important principle of the Outer Space Treaty which is implemented by the Artemis Accords. Specifically,

> via the Artemis Accords, NASA and partner nations will provide public information regarding the location and general nature of operations which will inform the scale and scope of "Safety Zones." Notification and coordination between partner nations to respect such safety zones will prevent harmful interference, implementing Article IX of the Outer Space Treaty and reinforcing the principle of due regard.[29]

The Artemis Accords raise a question far closer to home: should safety zones or self-defense zones be implemented on the Moon, but not in Earth orbits, especially when the latter is far more urgent to the well-being of humankind? This same logic should propel the United States and other nations to establish zones for Earth orbits. Moreover, these zones should be set up and enforced during peacetime to establish precedent and prevent ambiguity before a crisis.[30]

Space Policy Directive-3 directs the Department of Commerce to oversee STM but mandates that "the Secretaries of Defense, Commerce, and Transportation . . . shall develop space traffic standards and best practices, including technical guidelines, minimum safety standards, behavioral norms, and orbital conjunction prevention protocols related to pre-launch risk assessment and on-orbit collision avoidance support services."[31] These standards and practices should be specific, transparent, and unambiguous so that space users can easily understand and comply with STM regulations. Directive-3 also recommends that the United States establish a process for "transiting volumes used by existing satellites" (the legal description of self-defense zones).[32] To make the process enforceable, one needs to specify the shapes and sizes of the zone. For example, each zone in geosynchronous Earth orbit (GEO) altitude could be spheres with a 50 km radius.[33] Regardless of the actual size of each zone, the DOD must proactively engage STM efforts now and not wait. It must use its knowledge to ensure equitable standards and practices for all space users.

Bodyguard Spacecraft

Many, including Weeden, suggest creating a resilient satellite architecture.[34] Such an architecture is a good strategy but faces three major challenges. First, achieving resilience will take time. Replacing all vulnerable and critical satellite constellations will not occur until the 2030s. Particularly acute is the vulnerability of US legacy constellations composed of too few (e.g., a dozen), expensive (e.g., $1 billion a satellite), and large (e.g., the size of a school bus) satellites. Examples include the GEO-based Space Based Infrared System (SBIR) satellites for early warning or Advanced Extreme High Frequency (AEHF) satellites for communications

in a nuclear-disrupted environment. Both systems have the vulnerable attributes of number, cost, and size. Because these satellites are critical for early warning and nuclear deterrence, the DOD should establish zones and bodyguards to protect them in the 2020s.[35]

Second, if the deployment of resilient constellations of proliferated smallsats is delayed, zones and bodyguards would still be needed well into the 2030s. Third, according to Christopher Scolese, National Reconnaissance Office (NRO) director, the NRO would continue to operate a mix of satellites of many sizes. There will be "some number of large satellites to address questions that only they can."[36] The need for legacy-style satellites will likely go beyond the NRO and the 2020s.[37] US legacy-style large and expensive satellites have a poor cost-exchange ratio for the cheap attackers and need to be defended with equally cheap bodyguard spacecraft.

The US should start a crash program now to develop smallsat bodyguards capable of defending against the mid-2020 ASAT threat.[38] The program should take advantage of its smallsat development, such as the low-cost Blackjack satellites.[39] The US has already indicated that smallsats will cost far less than $1 billion each. SpaceX indicates a cost of $1 million each;[40] Bank of America, $5 million;[41] Planet Labs, merely $100,000–200,000;[42] and Morgan Stanley, $500,000.[43] DARPA envisioned that the cost of each smallsat under its Blackjack program, including launch, would be less than $6 million.[44]

In November 2018, the *Economist* reported that Erwin Duhamel, then head of security strategy at the European Space Agency, "observes that officials in several places are now studying the idea of defending important satellites with 'bodyguard' spacecraft."[45] On 25 July 2019, France announced that it would implement bodyguard spacecraft to protect its critical satellites in 2023.[46] The United States has not made any public statement about whether it will use bodyguard spacecraft to protect critical satellites against robotic ASATs. It is currently deficient in defining self-defense zones and deploying bodyguard spacecraft in time to counter the emerging robotic threat.[47]

Protecting Satellites without Escalation

The US should design and operate self-defense zones and bodyguard spacecraft to protect our satellites without escalating any potential conflict. The following three guidelines will help. First, Article 51 in the UN Charter says that "nothing in the present Charter shall impair the inherent right of individual or collective self-defence."[48] The right of self-defense is never in contention but of great concern is the misuse of self-

defense for offensive purposes. The US can minimize this problem by announcing that it will use self-defense zones to protect its satellites and declaring that its spacecraft, including bodyguards, will follow the same rules to respect another country's zones. We should also design all bodyguard weapons, including robotic arms, for short-range defense—to disable an invader inside our zone. Such short-range bodyguards would be adequate for self-defense within our zones but could hardly be used to attack other countries' satellites from outside their self-defense zones.

Second, to maintain crisis stability, we need to allow invaders to retreat from our zones without harm to any satellites. Thus, as soon as another country's spacecraft of any kind enters our zones without prior consent, we should immediately broadcast the incursion and demand immediate retreat. If the intrusion continues, the bodyguard should initially take defensive actions that will cause only temporary or reversible damage to invaders. Each bodyguard should host a suite of selected countermeasures—such as electronic jamming, laser dazzling, and decoys—to disable the invader without permanent harm. A bodyguard might also capture an invader and move it out of our zone.

Finally, once appropriate reversible countermeasures have been exhausted, a bodyguard would disable the invader without creating excessive debris. For example, a bodyguard can use its robotic arms to bend antennae, solar panels, or sensors of the invading robotic spacecraft to disable it with little or no debris.

The Legality of Zones in Western and International STM

The first fundamental space event—*Sputnik 1* circling the Earth—violated international air law that, at the time, extended a nation's sovereignty vertically into outer space over its territory.[49] Fortunately, this law was unable to restrain progress in rocketry that launched humankind into the space age. Today, should we consider only candidate solutions that meet all existing laws, or should we also consider solutions that are far more effective? The foreword to the 2002 United Nations collections of its treaties and principles on outer space states, "As is appropriate to an environment whose nature is so extraordinary, the extension of international law to outer space has been gradual and evolutionary—commencing with the study of questions relating to legal aspects, proceeding to the formulation of principles of a legal nature and, then, incorporating such principles in general multilateral treaties."[50]

This passage reflects the development of space law as gradual and evolutionary to ensure its relevance in guiding solutions to the challenges of

the space environment. While our forefathers could not possibly know what new threats or opportunities would look like, they made provisions for us to amend articles to better manage the contemporaneous space environment. For example, Article XV of the Outer Space Treaty states, "Any State Party to the Treaty may propose amendments to this Treaty. Amendments shall enter into force for each State Party to the Treaty accepting the amendments upon their acceptance by a majority of the States Parties to the Treaty and thereafter for each remaining State Party to the Treaty on the date of acceptance by it."[51]

As to the proximity threat, Weeden claims that the space zones proposed in the past are "unlikely to have a strong legal footing."[52] Some argue "that a 'keep out' area like a safety zone could run afoul of the Outer Space Treaty's prohibition on appropriating space for one nation's sovereign use."[53] That is to say, space zones do not comply with Article II of the Outer Space Treaty where "outer space, including the moon and other celestial bodies, is not subject to national appropriation by claim of sovereignty, by means of use or occupation, or by any other means."[54] However, the author has argued that "while the owner of the satellite does not have sovereignty over the self-defense zone, the United States can propose, according to Article IX of the 1974 Convention on Registration of Objects Launched into Outer Space, that this Convention be amended to automatically include the self-defense zone in the registration of the satellite to be launched or, retroactively, already launched into space."[55]

When the International Telecommunications Union assigns an orbital slot to a GEO satellite, the satellite owner does not have sovereignty over the slot. One can argue that Article II does not consider an assigned slot as claiming sovereignty because Article II must yield to a law of higher order—the law of nature. A law of physics dictates that two physical objects, such as satellites, cannot occupy the same spot at the same time. Consequently, when a law of man (i.e., Article II) conflicts with a law of nature, the former has no choice but to be waived.

Similarly, if a robotic spacecraft can legally pre-position itself arbitrarily close to a satellite before commencing an attack, the defender cannot possibly be fast enough to exercise Article 51's inherent right of self-defense within its legally assigned slot. It is illegal either for the attacker to stay so close or for the defender to exercise satellite self-defense. Clearly, the former should yield. Claiming that self-defense zones violate Article II disregards the purpose of the Outer Space Treaty—that being the peaceful use of space. As zones are needed to deter the proximity threat and keep the peace, Article II must yield.

For those who still believe that Article II should continue to reign supreme, they can seek comfort from the concept of warning zones, which are designed to meet existing laws. In June 2020, Michael Cerny et al. made an important observation of US maritime operational practices since at least 2006. They drew on Heinegg's analysis that zones merely served to protect vessels and found that

> much like Article I of the OST [Outer Space Treaty], international maritime law does not recognize any situation during which freedom of navigation on the high seas can be limited. However, warning zones are neither operational nor exclusionary, and instead "merely serve to protect the naval vessels from attack or from illegal activities. . . ."[56] Although these zones are historically established during wartime or national emergency, it is generally accepted that these zones can be established during peacetime under international law to protect naval vessels.[57]

Cerny et al. then applied the concept of maritime warning zones to space:

> The declaration of the zone itself is *not* understood—either implicitly or explicitly—to grant any right to the declaring state that it does not already possess. Instead, much like certain similar zones in the maritime domain, warning zones in space serve an information gathering function: "trespass" *per se* is not restricted, but can, upon meeting certain thresholds, provide increasingly certain evidence of hostile intent which would justify preemptive use of force in self-defense. Warning zones, would, therefore, provide an important—indeed, essential mechanism for clarifying intent, reducing the propensity for miscalculation by either side, improving signaling by both parties, and enhancing stability in crises (emphasis in original).[58]

Astutely, they took advantage of the boundary of a self-defense zone as a clear threshold for action against "increasingly certain evidence of hostile intent," which would justify preemptive self-defense. They conclude that "the unilateral establishment of warning zones around United States satellites presents a potential solution to the threat of co-orbital ASATs without violating Articles I and II of the OST [Outer Space Treaty]."[59]

Moreover, Rebecca Reesman and Andrew Rogers report that "to reduce the chance of collisions and to make the intent of nearby objects clear, the ISS [International Space Station] has a nominal approach ellipsoid around it in space. This ellipsoid extends four kilometers in front and behind the ISS path and two kilometers above, below, and beside it. The ISS also has defined a 200-meter 'keep-out' zone; external vehicles are only permitted to fly in this zone with approval and within a defined

approach corridor."[60] Thus, zones have already been used in space to keep at least one satellite (i.e., the ISS) safe.

In retrospect, in the past four decades, space zones of different names—such as self-defense zone, keep-out zone, safety zone, and, most recently, warning zone—have been proposed. A self-defense zone is intended to provide a timely warning for initiating legitimate preemptive self-defense while a keep-out zone restricts traffic in an area to prevent potential attackers. A safety zone is established to keep other spacecraft at a distance to avoid collisions, while a warning zone serves to provide "increasingly certain evidence of hostile intent"[61] to justify preemptive self-defense. In any case, as the proximity threat is fast approaching, pragmatists would simply contend that we are far better off to have any of these roughly similar zones in the interim than to wait for the perfect zone at some future time.

Some space planners and experts with very different ideologies are also uncomfortable with self-defense zones. Why? They want to protect America's right to conduct close-up inspections of Chinese and Russian satellites, one satellite at a time. This is something the United States has done since 2016.[62] The desire of policy makers to preserve America's freedom of action to conduct close one-on-one inspections, however, comes with high risk. This policy unwittingly validates China's and Russia's right to threaten our key satellites at close range with an unlimited number of hostile robotic spacecraft. The Pentagon should study whether the US needs close inspection and if it would be acceptable to forgo inspections of another country's satellites from less than 50 km unless requested to do so.[63]

A compromise is limiting the number of simultaneous close inspections. The US could continue one-on-one inspections closer than 50 km but abide by a less stringent but still useful rule: no state should have more than one spacecraft close to any other state's satellites without prior consent.[64] This policy would allow continued close-in space inspections but deprive China and Russia of the right to simultaneously attack more than one of our key satellites at close range. Of course, the United States will observe the same rule toward another country's satellites. The key is to be willing to negotiate and agree to a threshold number of spacecraft in close proximity that applies to all countries equally and fairly. However, regardless of how low the threshold number of close-in spacecraft, we will always need bodyguard spacecraft to protect our critical but vulnerable satellites.

The Legality of Bodyguard Spacecraft in Western and International STM

The legality of bodyguard spacecraft hinges on the legality of preemptive self-defense. As far back as 1842, Secretary of State Daniel Webster viewed preemption as legal, provided certain conditions are met. Subsequently, jurists like Roberto Ago in 1980 came to a similar set of conditions: necessity, proportionality, and immediacy.[65] Thus, "pre-emptive self-defense against space stalkers is necessary because the US cannot defend with, as Ago stated, 'measures not involving the use of armed force.' It is proportional because . . . the pre-emption is not allowed to go beyond what is needed to disable this attack. It must take place immediately, as the attack is ready and can be imminent."[66] Cerny et al. also argue similarly.[67] Thus, bodyguards should be allowed to exercise the "inherent right"[68] of self-defense stipulated in Article 51 of the UN Charter, even preemptively when Ago's conditions are met.

A Dual-Track Approach to International STM

China and Russia want STM to focus on commercial servicing and not self-defense zones or bodyguards. The latter would prevent them from taking advantage of lax commercial STM rules, which do not prohibit stalking another country's satellites. Worse yet, they are not alone in their distaste for zones and bodyguards. Many others in the West feel the same. There is a powerful camp with similar views, including Weeden:

> We need to keep the military security discussion separate from the commercial servicing and RPO discussions. . . . We've got 35 companies as members . . . working on best practices and standards for commercial servicing. Their biggest concern is they're going to get lumped in with all the military stuff and all of their investment and insurance is going to evaporate. I think if we do have space traffic management, it has to be explicitly for commercial, civil activities[,] . . . but I don't think we should try and make civil space traffic management that applies to military space traffic.[69]

This view can have serious unintended consequences. Weeden does not want military security measures, such as zones and bodyguards, included in international STM. He considers that "the RPO [proximity] threat is misunderstood and overblown" and that STM contributes little to military security, such as preventing a proximity attack.[70] In a 2020 Global World Foundation report, Weeden and Victoria Samson state that "warning time of such a [Chinese RPO satellite's] close approach would likely

be at least hours (for LEO) or days (for GEO), unless the attacking satellite was already in a very similar orbit."[71] Indeed, attacking RPO spacecraft are slow-flying. This warning of days long is useless because international rules currently allow the attacking satellite to remain arbitrarily close to its target for indefinite periods. Their statement of "unless the attacking satellite was already in a very similar orbit" would likely mean that if the attacking satellite were in the same orbit and arbitrarily close to our satellite as the current rules allow, there would be insufficient warning to take legitimate and timely actions to prevent the attack. Thus, the STM cannot simply focus on commercial and civil activities and should include a "military security discussion," such as how to prohibit the attacker from getting so arbitrarily close during peacetime or crisis. A key motivation of the States Parties to agree to the Outer Space Treaty was their recognition of "the common interest of all mankind in the progress of the exploration and use of outer space for peaceful purposes."[72] Clearly, using STM to prevent the proximity threat is for a critical peaceful purpose.

Viewing the proximity threat as overblown and STM as applying exclusively to commercial matters—and thus being dealt with only in international fora—aids China and Russia in two ways. First, they hope the West will continue to be ambivalent about the new danger. Second, they seek continued negotiations on STM matters via international fora, where agreements are typically made by consensus and they will have far better control of the outcome as they can just say no to the ones they do not like.

Unfortunately, if the West continues to think and negotiate as China and Russia expect, it will inadvertently live in the shadow of the proximity threat indefinitely. Naturally, China and Russia would encourage those in the West who want commercial activities in STM and not zones and bodyguards. Moreover, there are those in the West who consider any negotiation to be a failure without an agreement. Given such thinking, China and Russia would have the upper hand in any final STM framework. They could stall any deliberation or decision on zones and bodyguards, as they have already done in agreeing to the 21 guidelines only, further delaying rules that resolve the proximity threat. The status quo remains, as does the proximity threat. Since the current approach will have dire consequences, the US must devise a new tactic to create international STM that provides economic prosperity while preventing threatening proximity operations.

Pursuing an International STM

The United States should take the lead on a dual-track approach by proposing and pursuing a Western STM regime and an international

STM regime in parallel. The West's initial negotiating position will propose zones and bodyguards in both STM regimes. While negotiations may compromise in the details of the zones and bodyguard rules, they will never forsake zones and bodyguards or accept STM that cannot prevent the proximity threat.

Western countries will decide rules for Western STM while all participating countries will determine rules for international STM. International agreements are typically made by consensus of the negotiating State Parties. Thus, the key disadvantage of pursuing only international STM, as the United States is currently doing, is allowing China and Russia to block those Western measures they dislike. The West is forced to choose either having international STM that subjects itself to the proximity threat or reverting to the status quo, which allows arbitrarily close pre-positioning and makes proximity threat possible in the first place. Under a dual-track approach, Western STM can be completed quickly, setting up a fair model for international STM. The latter would follow in due time without any deadline.

The DOD's Defense Space Strategy refers to "allies" 32 times in the summary report of 18 pages. One needs to cite only a few passages to see how heavily the DOD relies on its allies and partners. First, "the strategy . . . moves with purpose and speed across four lines of effort," including "cooperat[ing] with allies, partners, industry, and other US Government departments and agencies."[73] Second, "in cooperation with allies and partners, DoD will . . . deter aggression in space . . . and support US leadership in space traffic management."[74] Third, "the United States has long maintained a robust and prolific arrangement of alliances and partnerships built on trust, common values, and shared national interests. This approach creates an important advantage for the United States and its allies and partners."[75]

The first point says that cooperation with allies and partners is one of four major efforts for the DOD to deal with space security issues. The second point makes explicit that deterring aggression in space and taking leadership in STM are two key national security issues. The third point implies that to negotiate more effectively with China and Russia, the United States needs to speak with one voice with other Western countries. In fact, it is easier to form a united front with our allies and partners because we share common ideologies, values, and interests. The dual-track approach capitalizes on these three points to wean China and Russia from the use of proximity threat.

Under the dual-track approach, China and Russia would clearly reject both STM proposals at the start. However, in the interim, before international STM is agreed upon, the West must develop and conduct its space operations according to its own STM. China and Russia would have little control over the Western STM agreement. The West could choose the type of zone—self-defense, keep-out, safety, or warning—garnering the most support from Western signatories. From the start, the West must design STM that keeps the peace and creates economic prosperity for all countries. This altruistic pursuit will attract adherents to Western and international STM. It should be noted that China and Russia are not prevented from doing space business with Western clients—the choice is theirs. However, it is common practice for companies doing business in a foreign country to abide by local regulations and laws. Thus, there is precedent for abiding by traffic rules in the vicinity of Western satellites. In doing so, China and Russia would not enter Western self-defense zones without prior consent.

Most importantly, the West can protect its critical space assets with zones and bodyguards through either international or Western STM. China and Russia could no longer pose the proximity threat whether they join either or neither STM agreement. In the case of joining an international STM regime that includes zones and bodyguards, they are prevented from posing a proximity threat by the zones and bodyguards already in international STM. In the case of joining Western STM that also includes zones and bodyguards, they are similarly prevented from posing a proximity threat. Even if they do not participate in either STM regime but try to get close to Western satellites, the West will still have zones to provide warning, the right of self-defense, and bodyguards to block the threat. Thus, under the two-track approach, the West would have no proximity threat and be far better off than under the current one-track approach. Its satellites would be protected regardless of the outcome of international STM.

A creative mind might wonder whether it would be far more straightforward to offer only Western STM. However, if just Western STM is offered, China and Russia would rightly complain of unfair treatment since they have no voting rights in Western STM. The dual-track approach offers international STM wherein decisions will be made by all participating State Parties. In the current approach, China and Russia have the right to disagree with what the West proposes. In the proposed international STM, the West has the right to disagree with theirs. The key difference between the current and recommended approaches is that,

without an international STM agreement, the current approach will revert to the status quo of no zones or bodyguards. In the suggested approach, the West will have Western STM—with zones and bodyguards—to fall back on.

Essentially, our current approach is to negotiate with China and Russia in not threatening our satellites and trust them to keep their word so that we can relieve the need to protect these critical assets. However, even the best-negotiated outcome could not meet the desired purpose of keeping our satellites safe. Assume that the US were able to offer enough incentives to China and Russia to deter the proximity threat. They can still conduct a successful and damaging proximity attack at the opening of a war if such an attack would yield national security benefits greater than the costs of breaking the STM agreement. China and Russia could also withdraw from the agreement in advance (e.g., one-year prior notice in the Outer Space Treaty[76]) knowing that the West could not possibly ready a defense in the span from the countries' withdrawal to the proximity attack.

Under the dual-track approach, the West can negotiate international STM under the auspices of the United Nations, just as it now does under the single-track approach. Negotiating Western STM is a multilateral effort for establishing the rules for companies and countries, domestic or foreign, to conduct space operations for Western clients and countries. In the context of the Western space market and STM, China and Russia plan to serve Western clients for economic benefits, and the West plans to attract them for the same. Once the proposed approach makes the inclusion of zones and bodyguards nonnegotiable, the proximity threat is solved provided that zones and bodyguards are implemented. Interestingly, with military-security measures (zones and bodyguards) included in STM, the negotiation will rightly focus on economic terms. Specifically, these are the benefits for China and Russia doing business in the Western space market and the benefits for the West having China and Russia participating in the Western market. Now, the negotiation becomes similar to a trade agreement or foreign direct investment where the result is more likely than the current approach to be a positive-sum game or agreement.

Phil Schneider, president of Schneider Consulting, comments that "historically, most foreign companies investing in the United States have been primarily driven by a need to create or enhance access to new markets and customers."[77] Similarly, the key incentive for China and Russia is to keep, or gain far more, access to the Western space market. While

China, Russia, and the West are competitors in the global space market, the West is endowed with ample incentives to attract countries to participate in the Western space market. Since the rules in Western and international STM are essentially the same, once China and Russia agree to abide by Western STM rules, it would be a small step for them to join with international STM. Thus, the dual-track approach uses Western STM as a necessary detour, which ironically can make international STM more likely than the current single-track approach. The West's key incentives for attracting Chinese and Russian companies to the Western space market are to enhance a competitive environment, including breaking domestic monopolies to be more innovative for the long term. Doing so will help Western space firms maintain or expand their global market share. Moreover, the global space industry will produce cheaper or better space products and services. Regulatory barriers imposed on Chinese and Russian firms should focus on national security and not on protecting domestic firms' market shares for the short term. The West should add considerable incentives to attract China and Russia to join international STM: a common STM system will sow harmony and cooperation and reap peace and prosperity for all countries.

Incentives for China and Russia to Join STM

Better access to the West's space business and technical expertise is the greatest incentive for China and Russia to join Western or international STM. Morgan Stanley assessed that the revenue from the global space market or economy was $339 billion in 2016 and projects it to be $1.1 trillion annually by 2040.[78] The company lists 11 market segments (table 1). Three of these are selected for analysis here: space launch, satellite manufacturing, and satellite internet services. The first segment selected, space launch, is projected to account for only 1 percent of the global space economy by 2040. However, Chinese and Russian space launch services have been the strongest suit in their space business. If they lose their edge in their strongest sector, they may fare even worse in other space sectors. Second, satellite manufacturing, although projected at just 1.7 percent, is chosen for analysis because it is the most important capability that determines their competitiveness in the global space market. Finally, satellite internet services (i.e., access) is chosen for the obvious reason that, at 34.6 percent, it will be the largest segment of the global space economy by 2040.[79]

Table 1. Global space economy by 2040

Segment	Annual revenue in $billion	Annual revenue in %
Consumer TV	100.0	8.6
Consumer broadband	80.0	6.9
Mobile satellite services	20.0	1.7
Earth observation services	30.0	2.6
Ground equipment	215.0	18.6
Satellite manufacturing	20.0	1.7
Satellite launch	11.0	1.0
Government spending	180.0	15.6
Insurance	0.8	0.1
Internet services	400.0	34.6
Space freight transportation	100.0	8.6
Total	*1,156.8*	*100*

Source: Morgan Stanley, *Space: Investment Implications of the Final Frontier* (New York: Morgan Stanley, 12 October 2017), 10, http://www.fullertreacymoney.com/.

Space Launch Industry

Goldman Sachs reported that the US share of global launch revenues during 2006–16 was 19 percent.[80] It jumped to 47 percent largely because SpaceX quickly captured numerous launches due to its success in substantially reducing launch costs.[81] By 2040, Morgan Stanley projects that SpaceX will have about 60 percent of the global launch market.[82] At worst, it will likely form a duopoly with Arianespace, a subsidiary of ArianeGroup. The combined launch share of SpaceX and Arianespace will continue to be far higher than that of Russia and China together. SpaceX's achievement in space launch hinges on its pathbreaking innovation and willingness to risk huge sums of capital. Its prowess forces both longtime and start-up competitors, including those from China and Russia, to strive for lower prices and better quality.

Goldman Sachs indicates that while Chinese Long March rockets are competitively priced with low failure rates, regulatory barriers prevent US components from flying on Chinese rockets.[83] Since nearly all European satellites contain US components, Chinese rockets cannot be used to launch US or most European satellites.[84] This regulatory barrier exemplifies how strongly the United States can control access for China and Russia to the Western space market and technical know-how.

As to the next-generation launch vehicle, Russia counts on *Soyuz-5*—expected to make its inaugural flight in 2022. However, Vitaly Egorov,

Open Space group manager on Facebook and PR specialist for a Russian aerospace company, remarks, "The Soyuz-5 is described as Russia's best bet at ensuring the nation's triumphal return to the launch vehicle market. It should be noted, however, that nearly all its launch characteristics are inferior to the current market leader [SpaceX], the Falcon-9. The only thing yet unknown is the price tag."[85]

For almost two decades, the United States has relied on the Russian RD-180 engine in Atlas 5 to power national security space launches. *Foreign Policy*'s Pentagon correspondent Lara Seligman recently noted that "the Defense Department is racing toward a congressionally mandated deadline of December 2022 to fly the first all-American rocket, powered by domestically produced engines, for US national security space launches."[86] Further, former Russian deputy prime minister Dmitry Rogozin, now the general director of Roscosmos, believes that it is not worth the effort for Russia to try to elbow SpaceX and China aside in the market for launch vehicles.[87] His statement signals that Russia may give up competing in the market to launch Western satellites.

The US and the West in general will continue to dominate the launch market. China's and Russia's launches of Western satellites will decline due to their deteriorating competitiveness with Western launch providers, such as SpaceX and Arianespace, and the US restrictions limiting or banning their launches.

Satellite Manufacturers

Goldman Sachs's list of 13 key satellite manufacturers worldwide includes nine in the US, two in Europe, and one in Brazil, with the China Great Wall Industry Corp. rounding out the list. Russia does not show up. This list is consistent with Goldman Sach's statement that "most satellites are built in the US or Europe."[88] The company also notes that "the nascent Chinese satellite manufacturers have yet to prove their technology over a meaningful period of time. Their oldest satellite is six years old, according to the Union of Concerned Scientists database as of June 2016."[89] Thus, China has too little experience in satellite manufacturing to capture sizable business from the West unless it focuses on some niche segments such as small satellites, where low prices can be a more critical consideration than better performance and reliability.

Egorov further indicates that "Russian navigation and telecommunication satellites suffer from low reliability and have an operational life twice as short as their European or American counterparts. Russia remote sensing satellites are 5–10 times the mass of their rivals and lag in terms of

data quality."[90] While Russia will try to get into satellite manufacturing,[91] like China, it would need to partner with Western companies to access market and technical know-how.

Satellite Internet Service Market

Besides continuing space launch services, SpaceX plans to launch its Starlink small satellites (smallsats) in low Earth orbits (LEO) to provide global internet services. Morgan Stanley reports that SpaceX plans to launch 11,943 smallsats to LEO in two phases: from 2019 to 2024 and from 2029 to 2032 for full capacity—at a cost of $10 to $15 billion.[92] In October 2019, SpaceX asked the International Telecommunication Union to arrange a spectrum for 30,000 Starlink satellites. These are in addition to about 12,000 already approved by the US Federal Communications Commission.

Thus, SpaceX's 42,000 Starlink satellites alone (i.e., 12,000 plus 30,000) will be five times the total spacecraft previously launched by all countries. SpaceX's spectrum request for so many satellites will make it difficult for countries such as China and Russia to play catch-up. According to an Institute of Defense Analyses report by Irina Liu et al., representatives of several Chinese companies expressed concern that their companies will be unable to obtain the spectrum they need because the large international companies have already claimed the most desirable spectrum bands for satellite communications. Thus, Chinese commercial companies are forced to use the less desirable V and Q bands that are more expensive to operate in.[93]

According to SpaceX, while about 720 LEO satellites will serve the US, merely doubling the number will serve most of the world. Thus, business economics favors LEO satellite constellations that serve the world versus a region. For this reason, the lion's share of the satellite internet access segment will likely go to only a few providers worldwide. These providers can afford the huge capital expenditure, offer good services at competitive prices, and are able to capture a big chunk of the market first.

China is unlikely to be one of these few big winners. Its best bet is to form a partnership with major Western providers. Moreover, even China's lower-cost advantage is losing its edge in providing space products and services. Liu et al. relay that "our interviews and searches on Chinese job websites showed that salaries in the space sector in China are lower, but not significantly so, than those in the United States." Moreover, "salaries for talented Chinese researchers are rising. . . . Although skilled labor in the commercial space industry in China will likely remain cheaper than in the West for the foreseeable future, somewhat cheaper labor may be insuf-

ficient to make the Chinese commercial space sector competitive with that of the West."[94]

Choosing Incentives or Intransigence

The two-track STM approach can protect satellites against the proximity threat whether or not China and Russia agree to join Western or international STM. The two countries would have to recalculate their strategic choices. Should they forgo the Western space market and voice their objection to the "unreasonable rules"? Or should they face the new reality that they can no longer mount an effective proximity threat and, instead, focus on maximizing the benefits from participating in the lucrative Western space market? Without a proximity threat, they would have far less bargaining power with the West and might acquiesce to Western STM, leading to better economic, social, and diplomatic relations.

China and Russia may well prefer to participate in the Western space market and follow the Western rules of zones and bodyguards. On 22 May 2020, the *Wall Street Journal* reported that "China broke with more than a quarter-century of tradition by not issuing an economic growth target for 2020, a stark acknowledgment of the challenges facing the world's second-largest economy."[95] In the aftermath of the Covid-19 pandemic, China and Russia will likely view participation in the lucrative Western space market as essential to resuming robust economic growth.

China and Russia lag behind the West, particularly the US, in technical know-how, the ability to innovate rapidly, the attitude of risk-taking, and financial resources to launch another constellation to compete with SpaceX or other early movers into the huge ($400 billion per year as shown in table 1) satellite internet service realm. Their best alternative is to form partnerships or working relationships with Western space service providers to share Western business and profits. They could support Western providers by building and managing ground facilities and services, including acting as distribution partners.[96] In their domestic markets, they will continue to retain unbeatable home-court advantages including financial and other subsidies by their governments. They also have some competitive advantages in countries friendly toward them.

China and Russia could manufacture components or even smallsats, in addition to launching them, if the West did not set up barriers to ban or limit their services in these areas. Smallsats would be an excellent vehicle to provide space services, other than launching, to the West. Making smallsats is relatively low tech, and they will be widely used in many applications worldwide. Moreover, since smallsats are intended for more frequent re-

placement and upgrade, they generally have a far shorter service life than large satellites. Buyers would be more willing to trade lower quality for a lower price for smallsat components or even for whole smallsats. The fact that the service life of Russia's satellites is around 50 percent that of their US counterparts might not be a deal breaker in the case of smallsats.[97]

Another potential business for China and Russia is Earth imaging and monitoring provided by companies such as Planet Labs, which operates the largest constellation of Earth observation satellites in the world. GIS Geography states that Earth remote sensing has hundreds of applications from farming to retail to urban growth.[98] Robbie Schingler, co-founder and chief strategy officer of Planet Labs, notes, "We run into challenges with predatory pricing by Chinese backed companies."[99] Although China or Russia would have some catching up to do, each has the potential—aside from strong subsidies from their governments—to compete in these services for Western and other commercial clients where lower cost is still a primary consideration.

Perhaps the most intriguing Western space segment for China and Russia is the use of robotic spacecraft to refuel, repair, move, and upgrade satellites in orbit and remove space debris. Getting into these peaceful services would be a perfect outcome for all parties involved. Their long-held claim of never wanting to pose a proximity threat would be vindicated by their joining international STM with zones and bodyguards or acquiescing to Western STM with the same measures. Accepting zone/bodyguard rules is the tradeoff China and Russia will make to acquire sales and technical expertise in the new space age—what they have been eagerly seeking from the West. Moreover, once they work more closely with the West in space, a better understanding of and improved relations with each other will enhance the atmosphere for reducing threats and keeping the peace.

In sum, the West, especially the United States, will continue to have most of the huge and growing global space market in the next two decades. Under the current single-track approach, China and Russia can continue to threaten the West with the proximity threat. In contrast, under the proposed dual-track approach, the West will have deterred the proximity threat from them and, later, North Korea, Iran, and others, regardless of whether these countries agree to STM rules. Most satisfying and important, the proposed approach harnesses the West's tremendous market power in space to create effective international STM.

Pursuing an International STM for Peace and Prosperity

STM is currently under early development. After analyzing the characteristics of the proximity threat and identifying the opportunities to counter it, three main actions emerge for developing a space traffic management regime that will resolve the proximity threat and provide economic prosperity for the West, China, Russia, and the rest of the world.

First, Western countries should unilaterally enact self-defense or warning zones to provide timely alerts that another country's spacecraft have entered the zones without prior consent. Collecting intelligence of hostile intent and preparing for self-defense would then be legitimately initiated to prevent invaders from reaching our vulnerable but critical satellites. Moreover, the US should pre-position bodyguard spacecraft, equipped with countermeasures, inside the zones to protect satellites.

Second, the United States should lead the way to pursue both Western STM and international STM in parallel, each having self-defense zones enforced by bodyguard spacecraft. Regardless of whether other countries, particularly China and Russia, join either STM arena, Western satellites will be protected by zones and bodyguards against the proximity threat. China and Russia will no longer be able to threaten the West's satellites in that way.

Third, during the next two decades, the global space market will grow to an annual revenue of $1.1 trillion by 2040, and the West will continue to have the preponderance of the global market. The West should offer China and Russia adequate economic incentives regarding access to the Western space market and technical know-how. Since the zones and bodyguards embodied in our proposed STM arenas will render their proximity threat ineffective, these countries are now left with two choices. One choice is not to participate in the Western market and voice their objection to the "unreasonable rules" of Western STM. The second choice is to participate in the Western space market and, as a common business practice, naturally follow Western STM, including rules for self-defense zones and bodyguard spacecraft.

China and Russia may well prefer the latter: participation in the lucrative Western space market despite the need to follow Western STM embedded with zone/bodyguard rules. Once they are willing to abide by these rules under Western STM, it could be a small accommodation to join the international STM because both STM systems will have essentially identical rules. Finally, countries will have fair international STM that resolves the proximity threat, helps keep the peace, and aids prosperity for all in the emerging second space age. SSQ

Brian G. Chow

Dr. Chow is an independent policy analyst with over 25 years as a senior physical scientist specializing in space and national security. He holds a PhD in physics from Case Western Reserve University and an MBA with distinction and a PhD in finance from the University of Michigan. This piece is a continuation of the author's previous *Strategic Studies Quarterly* articles "Stalkers in Space: Defeating the Threat," Summer 2017, and "Space Arms Control: A Hybrid Approach," Summer 2018. Contact him at brianchow.sp@gmail.com.

Notes

1. U.S.-China Economic and Security Review Commission (USCC), *2015 Report to Congress of the U.S.-China Economic and Security Review Commission*, 114th Cong., 1st sess. (Washington, DC: Government Printing Office, November 2015), 16, https://www.uscc.gov/. See also Brian Chow, "Avoiding Space War Needs a New Approach," *Defense News*, 16 September 2015, https://www.defensenews.com/.

2. USCC, *Report to Congress*, 321.

3. The White House, "Remarks by Vice President Pence on the Future of the U.S. Military in Space," 9 August 2018, https://www.whitehouse.gov/.

4. They are Daniel R. Coats, director of national intelligence ("satellites . . . intended to advance counterspace capabilities"): see Daniel R. Coats, *Statement for the Record: Worldwide Threat Assessment of the US Intelligence Community* (Washington, DC: Office of the Director of National Intelligence, 13 February 2018), 13, https://www.dni.gov/; Office of the Secretary of Defense ("dual-use technologies in space that could be applied to counterspace missions" in 2018 and 2019; "orbiting space robots" in 2020): see Department of Defense, *Annual Report to Congress: Military and Security Developments Involving the People's Republic of China 2018* (Washington, DC: Department of Defense, 2018), 40, https://media.defense.gov/; and the DOD's same-titled 2019 and 2020 reports, respectively, 51, https://media.defense.gov/; and 65, https://media.defense.gov/; Lt Gen Robert Ashley, director of the Defense Intelligence Agency ("enemy robot satellites"): see Patrick Tucker, "Pentagon Intelligence Chief: Russia and China Will Have Weapons in Space 'In the Near Future,' " *Defense One*, 27 June 2018, https://www.defenseone.com/; Dr. Yleem Poblete, former assistant secretary of state for arms control verification and compliance ("very abnormal behavior by a declared 'space apparatus inspector' . . . a very troubling development"): see Yleem D. S. Poblete, "Assistant Secretary Poblete Addresses the Conference on Disarmament," U.S. Mission to International Organizations in Geneva, 14 August 2018, https://geneva.usmission.gov/; the Air Force's National Air and Space Intelligence Center (NASIC) ("counterspace capabilities . . . to . . . grapple with a satellite"): see National Air and Space Intelligence Center, *Competing in Space* (Wright-Patterson AFB, OH: NASIC, December 2018), 24, https://media.defense.gov/; Defense Intelligence Agency ("robotic technology for satellite . . . but . . . used for military purposes"): see Defense Intelligence Agency, *Challenges to Security in Space*, January 2019, 10, https://www.dia.mil/; Gen John Hyten, commander, US Strategic Command ("kamikaze satellites"): see David Axe, "Pentagon Admits Plan to Launch 1,300 Satellites Might Not Prevent Chinese or Russian Attacks," *Daily Beast*, 10 April 2019, https://www.thedailybeast.com/; Dr. Scott Pace, executive secretary, National Space Council (quoted Vice President Pence's concern: "maneuver their satellites into close proximity of ours"): see "National Space Council's Pace on International Space Cooperation, Europe, China," 12 October 2019, https://www.youtube.com/; Gen John Raymond, the first chief of space operations, United States Space Force ("'threatening behavior' in outer space"): see Sandra Erwin, "Raymond Calls Out Russia for 'Threatening Behavior' in Outer Space," *SpaceNews*, 10 February 2020, https://spacenews.com/; and Dr. Christopher Ford, assistant secretary of state for international security and non-proliferation ("concerns about Russia's ability to operate in close proximity to other satellites"): see Christopher Ford, "Whither Arms Control in Outer Space? Space Threats, Space Hypocrisy, and the Hope of Space Norms," *New Paradigms Forum*, 6 April 2020, http://www.newparadigmsforum.com/.

5. Michael Gordon, "U.S., Russia Hold Talks on Space Security," *Wall Street Journal*, 27 July 2020, https://www.wsj.com/.

6. Sandra Erwin, "U.S. SPACECOM nominee Dickson says countries must be held accountable for actions in space," *SpaceNews*, 28 July 2020, https://spacenews.com/.

7. *Spacecraft* and *satellite* are the same and used interchangeably. However, in this article, *spacecraft* tend to represent vehicles with high maneuverability and that frequently move considerable distances while *satellites* tend to stay in, and provide services from, the same assigned orbits.

8. The well-accepted Kessler Syndrome theory says that collisions between space objects could cause a chain reaction of further collisions, increasing the debris density to such a high level that it renders space activities in certain orbital ranges infeasible. Donald J. Kessler and Burton G. Cour-Palais, "Collision Frequency of Artificial Satellites: The Creation of a Debris Belt," *Journal of Geophysical Research* 83, no. A6 (1 June 1978): 2637–46, http://www.castor2.ca/.

9. Permanent Representative of the Russian Federation and the Permanent Representative of China to the Conference on Disarmament, "Draft Treaty on the Prevention of the Placement of Weapons in Outer Space, the Threat or Use of Force against Outer Space Objects," 12 June 2014, CD/1985, https://www.fmprc.gov.cn/.

10. The White House, Space Policy Directive-3, National Space Traffic Management Policy, 18 June 2018, https://www.whitehouse.gov/.

11. The White House, *Space Traffic Management*.

12. Department of Defense, *Defense Space Strategy Summary* (Washington, DC: DOD, June 2020), 8, https://media.defense.gov/.

13. Brian G. Chow, "Stalkers in Space: Defeating the Threat," *Strategic Studies Quarterly* 11, no. 2 (Summer 2017): 96–102, https://www.airuniversity.af.edu/.

14. This article uses a broad definition of the *West* (i.e., Western countries), which includes NATO member countries; Australia, New Zealand, Japan, South Korea, Taiwan, Israel; and the nonaligned Austria, Finland, Sweden, and Switzerland. See "Western countries," fact-index.com, accessed October 2020, http://www.fact-index.com/. It also builds on the self-defense zone/bodyguard spacecraft portion of a research report by Brian G. Chow, "Commercial Space: Space Controls and the Invisible Hand," which first appeared on the website of its sponsor, the Nonproliferation Policy Education Center, on 16 June 2019, http://npolicy.org/.

15. Committee on the Peaceful Uses of Outer Space, "Draft Guidelines for the Long-Term Sustainability of Outer Space Activities," A/AC.105/2018/CRP.21, 27 June 2018, Guideline 8 on p. 7, https://www.unoosa.org/.

16. Committee on Peaceful Uses, "Draft Guidelines," 7; and the Committee on the Peaceful Uses of Outer Space, "Guidelines for the Long-term Sustainability of Outer Space Activities," A/AC.105/2018/CRP.20, 27 June 2018, https://www.unoosa.org/.

17. Defense Intelligence Agency, *China Military Power: Modernizing a Force to Fight and Win* (Washington, DC: Defense Intelligence Agency, January 2019), 43, https://www.dia.mil/.

18. NASA defines a *smallsat* as a satellite with a mass less than 180 kilograms and about the size of a large kitchen refrigerator. CubeSats are smallsats and a class of nanosatellites that use a standard size and form factors. The basic unit or 1U measures 10x10x10 cms. A CubeSat takes different sizes such as 1, 1.5, 2, 3, 6, and 12U with a mass of 1-10 kilograms. See Elizabeth Mabrouk, ed., "What Are SmallSats and CubeSats?," NASA, 6 August 2017, https://www.nasa.gov/.

19. For further discussion on the US's horrific choices between warfighting with degraded space support and no intervention when necessary, see Brian G. Chow, "Space Arms Control: A Hybrid Approach," *Strategic Studies Quarterly* 12, no. 2 (Summer 2018): 117, https://www.airuniversity.af.edu/.

20. Joshua Philipp, "China Is Branding Anti-Satellite Weapons as 'Scavenger Satellites,' " *Epoch Times*, 5 May 2019 and updated 6 May 2019, https://www.theepoch times.com/.

21. Jacqueline Feldscher, "Chinese Small Satellites Threaten U.S. Industry," *Politico*, 26 April 2019, https://www.politico.com/.

22. Wayne Rash, "The Future of GPS," Hewlett Packard Enterprise, 16 July 2019, https://www.hpe.com/. The price of the first 10 satellites was estimated at $577 million each. The Air Force expected the remaining 22 satellites to cost $7.2 billion or $327 million each. However, the Government Accounting Office estimated the cost at $12 billion or $545 million each. See Dan Elliott, "US Air Force Set to Launch 1st Next-Generation GPS Satellite," *Air Force Times*, 16 December 2018, https://www.usatoday.com/story/.

23. Article X of the Convention on International Liability for Damage Caused by Space Objects says that a claim should be filed not later than one year following the date of the occurrence of the damage. Article XIV states, "If no settlement of a claim is arrived at . . . within one year from the date on which the claimant State notifies the launching State . . . , the parties concerned shall establish a Claims Commission." Article XV adds six months to appoint the Chairman of the Commission if the claimant State and the launching State cannot agree on the choice of the Chairman. Finally, Article XIX states that the Commission shall give its decision or award no later than one year from the date of its establishment. Thus, the total amount of time is up to three and a half years. See United Nations Office for Outer Space Treaties, *United Nations Treaties and Principles on Outer Space and Related General Assembly Resolutions* (New York: United Nations, 2002), 16–18, https://www.unoosa.org/.

24. Todd Master, "Consortium for Execution of Rendezvous and Servicing Operations (CONFERS)," Defense Advanced Research Projects Agency, accessed 30 April 2020, https://www.darpa.mil/.

25. CONFERS, "CONFERS Recommended Design and Operational Practices," 1 February 2019, https://www.satelliteconfers.org/.

26. "Transcript of a Discussion between Dr. Brian Chow and Dr. Brian Weeden on Space Zones and Bodyguards for Proximity Operations," moderated by Henry Sokolski, co-sponsored by the Nonproliferation Policy Education Center (NPEC) and the American Bar Association (ABA) Standing Committee on Law and National Security, Washington, D.C., 2 March 2020, http://npolicy.org/; ABA Standing Committee on Law and National Security and NPEC, "Working Smarter with America's Spacefaring Allies," Workshop Report, March 2020, 17, https://www.americanbar.org/; and Secure World Foundation, "SWF Discusses the Pros and Cons of 'Guardian' Satellites and Space Zones of Control," 2 March 2020, https://swfound.org/.

27. United Nations, *Treaties and Principles*, 6.

28. United Nations, 6.

29. NASA, "The Artemis Accords," May 2020, https://www.nasa.gov/.

30. Chow, "Stalkers in Space," 96, 100, 102.

31. The White House, *Space Traffic Management*.

32. The White House, *Space Traffic Management*.

33. While 50 km is a reasonable choice, the DOD should determine actual threshold distance after consultation with other agencies domestic and foreign. The use of specific numerical values in guidelines is common. For example, the Inter-Agency Space Debris Coordination Committee (IADC) Space Debris Mitigation Guidelines use 25 years as the post-mission orbital lifetime limitation. See also Inter-Agency Space Debris Coordination Committee, "IADC Space Debris Mitigation Guidelines," IADC-02-01 Revision 1, September 2007, 9, http://www.unoosa.org/.

34. American Bar Association and NPEC, "America's Spacefaring Allies," 19.

35. Brian G. Chow, "Nuclear Vulnerability: In-Orbit Bodyguards Would Help Protect NC3 Satellites from Attacks," *SpaceNews*, 1 April 2019, https://spacenews.com/.

36. Sandra Erwin, "Scolese: NRO Advancing Space Technology, Developing Tactics to Defend Satellites," *SpaceNews*, 3 December 2019, https://spacenews.com/.

37. For example, the Air Force's constellation of Next-Generation Overhead Persistent Infrared (Next-Gen OPIR) missile warning satellites will be deployed by 2029. However, it seems unlikely that they will be resilient enough against a proximity threat because the constellation has the three nonresilient attributes discussed earlier (number, cost, and size). See Sandra Erwin, "Northrop

Grumman Gets $2.3 Billion Space Force Contract to Develop Missile-Warning Satellites," *Space News*, 18 May 2020, https://spacenews.com/.

38. The author recommended in another article that "the United States should quickly initiate a crash program to develop cheap but effective bodyguards by using small satellites." Brian G. Chow, "Two Ways to Ward off Killer Spacecraft," *Defense One*, 30 July 2019, https://www.defenseone.com/.

39. The Defense Advanced Research Projects Agency plans to launch the first three experimental Blackjack satellites in late 2020 and early 2021, and as many as 20 satellites by 2022. Sandra Erwin, "DARPA to Begin Launching Blackjack Satellites in Late 2020," *SpaceNews*, 11 May 2020, https://spacenews.com/.

40. Morgan Stanley reported a cost of $10–$15 billion for SpaceX's deployment of 11,943 Starlink smallsats. This amounts to roughly $1 million per satellite for manufacturing and launch. Morgan Stanley, *Space: Investment Implications of the Final Frontier* (New York: Morgan Stanley, 12 October 2017), 9, 22, http://www.fullertreacymoney.com/.

41. According to the Organisation of Economic Co-operation and Development, "Satellites typically cost about US$1bn [billion] to build and launch but the rise of small satellites (CubeSats) means this cost has now been driven down to US$5mn [million] by some companies like Terra Bella/Skybox Imaging." Further, "the speed of manufacturing satellites is increasing too; in early 2015 it took Planet Labs just nine days to build two CubeSats." See Felix Tran et al., "Thematic Investing: To Infinity and Beyond – Global Space Primer," Bank of America Merrill Lynch Transforming World Thematic Research, 2017, 62, http://newspaceglobal.com/.

42. While Planet Labs reported a cost as low as $10k for each Dove cubesat, others estimated the cost to be $100-200k excluding launch cost. See Stack Exchange, "Cost of an Earth Observation Cubesat Satellite like Planet Labs' Doves?," *Space Exploration Beta*, 2019, https://space.stackexchange.com/.

43. Morgan Stanley, "Space: Investing in the Final Frontier," 24 July 2020, https://www.morganstanley.com/.

44. Sandra Erwin, "DARPA Sees Clear Path to Faster, Cheaper Space Technology," *SpaceNews*, 1 March 2018, https://spacenews.com/.

45. "It Will Soon Be Possible to Send a Satellite to Repair Another," *The Economist*, 24 November 2018, https://www.economist.com/.

46. "France's New Space Defense Strategy," *SatelliteObservation.net*, posted by gosnold, 27 July 2019, https://satelliteobservation.net/.

47. Brian G. Chow and Henry Sokolski, "Growing U.S. Satellite Vulnerability: The Silent 'Apocalypse Next,'" *SpaceNews*, 22 August 2018, https://spacenews.com/.

48. Repertory of Practice of United Nations Organs, Charter of the United Nations, chap. VII, art. 51, accessed 26 June 2020, https://legal.un.org/. (Henceforth, UN Charter.)

49. Jason Krause, "The Outer Space Treaty Turns 50. Can It Survive a New Space Race?," *ABA Journal*, 1 April 2017, https://www.abajournal.com/.

50. United Nations, *Treaties and Principles*, v.

51. United Nations, 8.

52. Chow and Weeden, "Transcript."

53. For example, see American Bar Association Standing Committee on Law and National Security and Nonproliferation Policy Education Center, "Future Space Controls and the Invisible Hand," Workshop Report, September 2019, 23, https://www.americanbar.org/.

54. United Nations, *Treaties and Principles*, 4.

55. Chow, "Stalkers in Space," 95.

56. Wolff Heintschel von Heinegg, "Current Legal Issues in Maritime Operations: Maritime Interception Operations in the Global War on Terrorism, Exclusion Zones, Hospital Ships and Maritime Neutrality," *International Law Studies* 80 (2006): 213–14, https://digital-commons.usnwc.edu/. Quoted in Michael Cerny et al., "Countering Co-Orbital ASATs: Warning Zones in

GEO as a Lawful Trigger for Self-Defense," Nonproliferation Policy Education Center, 30 June 2020, 31, http://npolicy.org/.

57. Cerny et al., 31.

58. Cerny et al., 14.

59. Cerny et al., 32.

60. Rebecca Reesman and Andrew Rogers, *Getting in Your Space: Learning from Past Rendezvous and Proximity Operations* (El Segundo, CA: The Aerospace Corporation, May 2018), 5, https://aerospace.org/.

61. Cerny et al., "Warning Zones," 14.

62. Brian Weeden and Victoria Samson, eds., *Global Counterspace Capabilities: An Open Source Assessment* (Broomfield, CO: Global World Foundation, April 2019), 3–5, https://swfound.org/.

63. As said earlier, while 50 km is a reasonable choice, the DOD should determine the actual threshold distance after consultation with other domestic and foreign agencies.

64. Brian G. Chow and Henry Sokolski, "Priority-One for Space Policy Should Be to Protect U.S. Satellites," *SpaceNews*, 12 October 2019, https://spacenews.com/.

65. Roberto Ago, "Addendum: Eighth Report on State Responsibility by Mr. Roberto Ago, Special Rapporteur—the Internationally Wrongful Act of the State, Source of International Responsibility (part 1)," extract from the *Yearbook of the International Law Commission*, 1980, vol. II(1), Document:-A/CN.4/318/Add.5-7, United Nations, 69–70, https://legal.un.org/.

66. Brian Chow, "China's New Space Threat and the Justification of US Pre-emptive Self-Defense," *Space Review*, 18 January 2016, https://www.thespacereview.com/.

67. Cerny et al., "Warning Zones," 23–26.

68. UN Charter, chap. VII, art. 51.

69. Chow and Weeden, "Transcript."

70. Chow and Weeden, "Transcript."

71. Brian Weeden and Victoria Samson, eds., *Global Counterspace Capabilities: An Open Source Assessment* (Broomfield, CO: Global World Foundation, April 2020), 1–8, https://swfound.org/.

72. United Nations, *Treaties and Principles*, 8.

73. Department of Defense, *Defense Space Strategy Summary*, 1.

74. Department of Defense, 2.

75. Department of Defense, 5.

76. United Nations, *Treaties and Principles*, 8.

77. Phil Schneider, "Financial Incentives for Foreign Business," *Area Development*, 2010, https://www.areadevelopment.com/.

78. Morgan Stanley, *Investment Implications*, 11, 26.

79. For more detailed analysis of the global space market, see Chow, "Commercial Space," 4–16.

80. Goldman Sachs, *Space: The Next Investment Frontier*, pt. 8, Profiles in Innovation series (New York: Goldman Sachs, 4 April 2017), 3, http://www.fullertreacymoney.com/.

81. Bank of America Merrill Lynch reported that "on a per kilogram of thrust basis launch costs have fallen from around US$10,000 per kg/LEO in 1967 when Saturn V launched to around US$2,600 per kg/LEO in 2016 with the Falcon 9 v1.2 (Full Thrust). . . . This could fall even further with Falcon Heavy, which is due for testing from late 2017 with a cost of only US$1,400 per kg/LEO, implying launch costs will have fallen by a factor of 10x with its introduction. Elon Musk believes that 'when upper/second stage & fairing are reusable launch costs will drop by a factor of more than 100x.'" Tran et al., "Global Space Primer," 32.

82. Morgan Stanley, *Investment Implications*, 25.

83. Goldman Sachs, *Next Investment Frontier*, 29.

84. Goldman Sachs, 29. See also Peter B. de Selding, "European Satellites Still Heavily Dependent on U.S. Parts," *SpaceNews*, 29 January 2015, https://spacenews.com/.

85. Vitaly Egorov, "Commercial Alternatives: The Issues and Challenges of the Russian Space Industry – Part III," *SpaceWatch.GLOBAL*, March 2018, https://spacewatch.global/.

86. Lara Seligman, "The New Space Race," *Foreign Policy*, 14 May 2019, https://foreignpolicy.com/.

87. TASS, "Russian Deputy PM Sees No Reason for Competing with Musk on Launch Vehicles Market," 17 April 2018, http://tass.com/.

88. Goldman Sachs, *Next Investment Frontier*, 13.

89. Goldman Sachs, 29.

90. Vitaly Egorov, "The Issues and Challenges of the Russian Space Industry – Part I," *SpaceWatch.Global*, March 2018, https://spacewatch.global/.

91. Brian Wang, "China Main SpaceX Competitor as Russia Is Giving Up," *Next Big Future*, 5 August 2018, https://www.nextbigfuture.com/.

92. Morgan Stanley, *Investment Implications*, 22.

93. Irina Liu et al., *Evaluation of China's Commercial Space Sector*, IDA Document D-10873 (Washington, DC: Institute of Defense Analyses, September 2019), 76, https://www.ida.org/.

94. Liu et al., 89.

95. Jonathan Cheng, "Beijing Scraps GDP Target, a Bad Sign for World Reliant on China Growth," *Wall Street Journal*, 22 May 2020, https://www.wsj.com/.

96. China and Russia can serve as distribution partners for Western space companies as others are doing. For example, Hughes Network Systems invested $50 million in OneWeb in 2015, and another $50 million in 2020 into the consortium to purchase OneWeb out of bankruptcy protection. In exchange, Hughes still plans to be a distribution partner for OneWeb capacity for connectivity to customers with networks for government and business sites, cellular backhaul needs, and community Wi-Fi hotspots. See Caleb Henry, "Hughes Network Systems to invest $50 Million in Revived OneWeb," *SpaceNews*, 28 July 2020, https://spacenews.com/.

97. Roger McDermott, "Russia's Military Exploitation of Outer Space," *Eurasia Daily Monitor* 17, no. 47 (8 April 2020), https://jamestown.org/. See also Egorov, "Russian Space Industry – Part I."

98. GIS Geography, "Planet Labs Imagery: The Entire Earth, Everyday," 26 April 2020, https://gisgeography.com/. See also GIS Geography, "100 Earth Shattering Remote Sensing Applications & Uses," 5 March 2020, https://gisgeography.com/

99. Sandra Erwin, "Ex-Im Bank to Step Up Support for Space Industry Challenged by Chinese Competitors," *SpaceNews*, 9 July 2020, https://spacenews.com/.

PERSPECTIVE

Missing: Legal Frameworks for Chemical Security

Richard T. Cupitt
Mary C. Vecellio

Abstract

In recent years, state and non-state actors have broken the taboo against the use of chemical weapons. Yet evidence suggests that the national legal frameworks for chemical security, as required of all UN member states by United Nations Security Council Resolution 1540 (2004), remain persistently underdeveloped. Worse, the international community has yet to generate a widely accepted set of international standards for chemical security. To provide a baseline on national implementation of the chemical security obligations under Resolution 1540, the authors led a research team that first identified key practices for chemical security laws and regulations from a review of more than 30 national, regional, and industry codes of conduct and guidance. They then extracted more than 600 laws and regulations identified by the 1540 Committee for analysis. After comparing these measures against key practices derived from the codes and guidance, the authors generated a composite index score for each UN member state and created a choropleth map to provide new insights into the status of 1540 implementation, from geographic clusters to unexpected outliers. Finally, they offer several potential determinants for further research.

*****

In the aftermath of the widespread use of chemical weapons during the First World War, many countries committed to not use such weapons again in the 1925 Geneva Protocol.[1] In the following decades, despite their use by a few governments and non-state actors against domestic and foreign targets, a strong international norm against chemical weapons emerged.[2] With the end of the Cold War, the international community further formalized this norm into a robust regime against chemical weapons by establishing the 1993 Chemical Weapons Convention (CWC), which now has 193 state parties and created the Organisation for the Prohibition of Chemical Weapons (OPCW). More recently, however, state and non-state actors' use of chemical weapons threatens to under-

mine the chemical weapons nonproliferation regime at its foundation. The chemical weapons attacks during the Syrian civil war brought opportunities for international cooperation, resulting in Syria acceding to the CWC and destroying much of its chemical warfare agents. At the same time, the attacks brought moments of division, such as grappling with the Syrian government's role in continued attacks and a contested Security Council vote on the OCPW's responsibility to determine attribution.

Furthermore, perpetrators of chemical weapon attacks are employing new agents and tactics, from sophisticated chemical weapons for assassinations in Malaysia and the United Kingdom to attacks, virtual and physical, on chemical facilities.[3] The use of novel agents in these most recent attacks have even prompted CWC state parties to add new chemicals, the families of novichocks and carbamates, to the schedules of chemicals controlled under the CWC for the very first time.[4]

Unfortunately, these challenges to the nonproliferation regime and the norms that underpin it are not isolated or infrequent. Of the 517 events involving chemical, biological, radiological, or nuclear terrorism from 1990 to 2017 in the Profiles of Incidents Involving CBRN and Non-State Actors (POICN) database, more than 400 involve chemical terrorism occurring in at least 59 countries on six continents.[5] Thus, the international community has much more to do to secure and prevent the illicit use of these chemicals.

The global community knows little about how national systems are implemented and enforced, beyond evidence that illicit actors can and have exploited them.[6] However, with funding from Global Affairs Canada, the Henry L. Stimson Center began a project to explore the national legal frameworks for chemical security in all 193 UN member states with the intention to develop a compendium of laws, regulations, or their equivalent that include specific obligations to secure toxic chemicals of proliferation concern.[7] The project also sought to identify a set of emerging chemical security standards by reviewing open source literature and then evaluating national legal measures against key elements of those standards.

All UN member states are required to have effective legal measures and other controls in place for chemical security under legally binding obligations of United Nations Security Council Resolution (UNSCR) 1540 (2004). However, the OPCW has not yet developed an international code of conduct or guidance on chemical security. Without OPCW guidance, countries determine on their own how they should implement their chemical security obligations. In contrast, the International Atomic Energy Agency (IAEA) has produced a code of conduct and guidance for

nuclear and radiological security.[8] The lack of internationally accepted chemical security standards and practices has contributed to a global disarray of national systems to secure toxic chemicals, their precursors, and related facilities. This article first identifies key practices and standards of chemical security applicable to UNSCR 1540. It then generates a composite index score to evaluate each UN member state and provides insight into each state's implementation. Finally, the article recommends areas for more research into compliance with UNSCR 1540.

Chemical Security Practices and Standards: Is There Guidance?

Unlike the relatively clear and internationally accepted IAEA standards and recommendations regarding the security of nuclear and radiological sources, the security of chemical weapons–related material has no such guidance. The OPCW does not outline, much less detail, explicit international standards and best practices for securing chemical weapons–related materials, facilities, or equipment. As a body designed to implement the Chemical Weapons Convention (CWC), the OPCW had to focus initially on the dismantlement of declared chemical weapons programs and abandoned chemical weapons and the monitoring of production and movement of scheduled chemicals to prevent the re-emergence of state programs.[9] Only since 2017 has the ambit of its work shifted to include securing chemicals of proliferation concern.[10] This does not mean, however, that other (though less globally authoritative) bodies have not produced codes, guidance, or sets of effective practices for securing chemicals of proliferation concern. We identified and reviewed over 30 sources related to chemical security, varying greatly in purpose and scope.[11]

Our review of these resources identified five primary documents detailing how to address and implement a range of chemical security measures in chemical facilities and laboratories:

- US Department of Homeland Security's Chemical Facility Anti-Terrorism Standards (CFATS)
- US Department of Homeland Security's *Risk-Based Performance Standards* [RBPS] *Guidance: Chemical Facility Anti-Terrorism Standards*
- Responsible Care© Security Code of Management Practices (chemical industry initiative in the United States)

- European Responsible Care© Security Code Guidance and Best Practice for the Implementation of the Code (chemical industry initiative in Europe)
- National Research Council, *Promoting Chemical Laboratory Safety and Security in Developing Countries*

These sources are considered foundational to our research because of their level of specificity regarding essential elements of a strong security system for chemical weapons–related materials. For example, CFATS establishes 18 risk-based performance standards that identify which areas of a facility's security system are examined.[12] The RBPS guidance document accompanying CFATS offers detailed recommendations to assist high-risk chemical facilities in selecting and implementing appropriate protective measures and practices to meet the 18 performance standards outlined in CFATS.[13]

Both the US and European Responsible Care Security Codes add value because they are the primary chemical industry initiatives on securing high-risk chemical materials.[14] The purpose of these codes is to help "protect people, property, products, processes, information, and information systems by enhancing security, including security against a potential terrorist attack, throughout the chemical industry value chain."[15] Notably, the chemical industries in the United States and Europe established these security codes. Though supported by governments, these codes are solely implemented and monitored by countries' chemical industry.

Finally, the National Research Council's *Promoting Chemical Laboratory Safety and Security in Developing Countries* offers guidance for laboratories in the developing world to implement safe and secure practices in handling and storing hazardous chemicals. It includes information on how to develop administrative structures and support systems to delineate responsibility and accountability in a chemical laboratory. It also describes how to establish a safety and security management system and outlines the types of hazards and risks in chemical laboratories.[16] These sources and many others provide a strong understanding of chemical security common standards and best practices currently being discussed and implemented in facilities and laboratories around the world.

Emerging Chemical Security Standards and Effective Practices

Based on these five primary documents, we extracted the following 21 common effective practices for securing chemical weapons–related materials, facilities, and equipment:

- Training for relevant stakeholders
- Registration/inventory of chemical materials
- Registration/inventory of licenses
- List of controlled chemicals, technologies, and equipment of concern
- Awareness-raising for relevant stakeholders
- Physical security measures
- Access controls
- Inspector authority/system
- Registration system
- Background checks
- Supply chain verification practices
- Security guards
- Proliferation-resistant chemistry practices
- Defining criminal offences and violations
- Imprisonment as penalty provisions
- Fines as penalty provisions
- Other penalty provisions (e.g., search and seizure, suspension of license)
- Incident reporting
- Threat reporting
- Risk-based security approaches
- Authorization/licensing system

Nuanced Understandings of Chemical Security Standards and Effective Practices

It is important to note that many of these practices are understood differently throughout the literature. For instance, the most common chemical security standard is the provision of training. However, the type of training and intended stakeholder vary across sources. Many sources recommend training all personnel in contact with chemical materials and equipment, including facility employees, contractors, service providers, value chain partners, transport staff, scientists, and students.[17] Others also encourage training stakeholders who research and/or regulate chemical materials. During the Global Summit on Chemical Safety and Security in 2016, the deputy director of the OPCW noted that it had trained safety

officers, researchers, policy makers, and legal officers who addressed chemical safety and security concerns.[18]

Some references discuss training specific topics, such as security vulnerability assessments, security awareness, potential hazards, and standard operating laboratory procedures. For example, the *Code of Conduct for the Practice of Chemistry in the Middle East and North Africa* recommends a "program of effective, qualified, mandatory training that covers safety, security, and environmental responsibilities."[19] Other sources encourage training all chemical personnel to watch for suspicious activities or persons.[20]

Moreover, discussions on implementing an inspections system occur throughout the literature, but in two different ways. One way is to inspect personnel, vehicles, equipment, and materials as they enter a chemical facility's premises.[21] The second and more common way is for a chemical facility (public or private) or laboratory to have regular third-party or independent inspections to assess security vulnerabilities or overall compliance with company policies or national regulations. Both the American and European chemical industry initiatives embodied in the Responsible Care© Security Code and the European Responsible Care© Security Code encourage chemical companies to implement third-party verification and to use external auditors and inspectors to monitor security threats for evolving threats.[22] Similarly, from a laboratory perspective, sources agree that an effective compliance system should have a program for regular inspections of all science, engineering, safety, and security practices at facilities.[23]

Additionally, we found that reporting incidents and suspicious activities (e.g., theft, diversion, fraud, facility breach, material or equipment tampering, cyber sabotage) is considered an essential practice for a robust chemical security framework, along with reporting credible security threats.[24] For example, the US chemical industry's Responsible Care© Security Code differentiates between incident and threat reporting requirements. US chemical companies affiliated with the American Chemistry Council are required to evaluate, respond to, investigate, report, communicate, and take corrective action for security incidents. They are also required to relay security threats—specifically physical and cybersecurity threats—to law enforcement personnel as appropriate.[25]

Ultimately, the extensive literature on chemical security demonstrates that stakeholders everywhere are considering common elements for a strong chemical security system. We extracted these common 21 best practices for comparison against national legal framework requirements world-

wide. We sought to determine whether the current legislation reflects these emerging standards.

Comparing Legal and Practical Standards

The 1540 Committee has collected a trove of information on the measures taken by each UN member state to implement the resolution. Collated in a "1540 Matrix" for each member state, this data primarily includes a range of laws, regulations, decrees, and other legal measures on the nonproliferation of nuclear, chemical, and biological weapons and their means of delivery. The 1540 Committee's matrix has more than 300 data fields to characterize implementation by each member state. At least 10 of these fields address implementation efforts to secure chemicals and/or facilities of proliferation concern.[26] Information in each member state's matrix is derived primarily from information submitted by that member state directly to the 1540 Committee. However, the committee can supplement the national reports by using any official government source produced by that member state, such as an official legal gazette, ministerial websites, or submissions to international or regional organizations.

Using the 1540 matrices, we began our research on national implementation of nonproliferation, including efforts related to chemical security, by searching all the names of the legal measures identified in the 10 data fields in the 193 1540 matrices.[27] This search generated an initial list of 643 national legal measures related to chemical security across all UN member states. Next, the research team introduced context into the textual analysis to refine our findings, narrowing the number of relevant chemical security measures to 43 found in 32 UN member states. It is worth noting that there is a considerable discrepancy between the much larger number of measures listed in the 1540 Committee matrices—643—and the 43 measures we identified as having explicit textual requirements to secure chemical weapons–related materials and facilities. This difference likely emanates from several sources, including some error by the research team, the 1540 Committee, and certainly from the authors' stricter textual requirements. The number of states (32) we identified with chemical security measures in place, however, correlates more closely to the number of states (55) the 1540 Committee identifies with physical protection requirements.[28] We also understand that the 1540 Committee's current Group of Experts will soon issue revised matrices with considerably fewer relevant legal measures for chemical security than it previously identified.

We compared each chemical security law to the index of 21 chemical security practices to evaluate if and how each law or regulation complies.

A composite score could range from 0 to 21. Based on our analysis, all 43 measures' composite index scores range from 3 to 18, indicating the number of chemical security standards a single law/regulation incorporates. Similarly, at the state level, we developed a composite index of 22 chemical security practices ranging from 0 (for a state where the authors could not identify any relevant law/regulation) to 22. The state-level composite index range differs from the measures level due to the need to count whether a UN member state has a law that requires securing chemical weapons–related materials/facilities. All measures evaluated against the security elements had to have this requirement to be considered relevant. However, only 32 states had relevant laws. Therefore, when analyzing national implementation efforts across the globe, we included an additional (22nd) chemical security element related to if/whether the state has a law requiring chemical security. Using the state-level composite index, we mapped the low to high scores.[29]

Evaluating National Legal Measures

Only a few of the 21 common practices identified seemed to be represented in the 43 chemical security laws and regulations identified, either at the individual measures level or, therefore, at the state level (see figs. 1 and 2, respectively). For instance, of the 32 states with relevant chemical security legislative frameworks, less than half (15) incorporate a national registry of chemicals, only 11 have a physical security requirement, and just four include background checks (fig. 1).

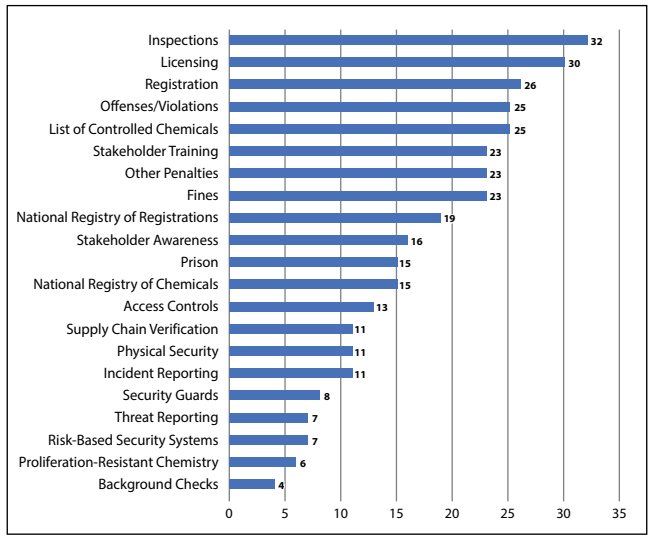


Figure 1. 21 Chemical security elements—individual measures

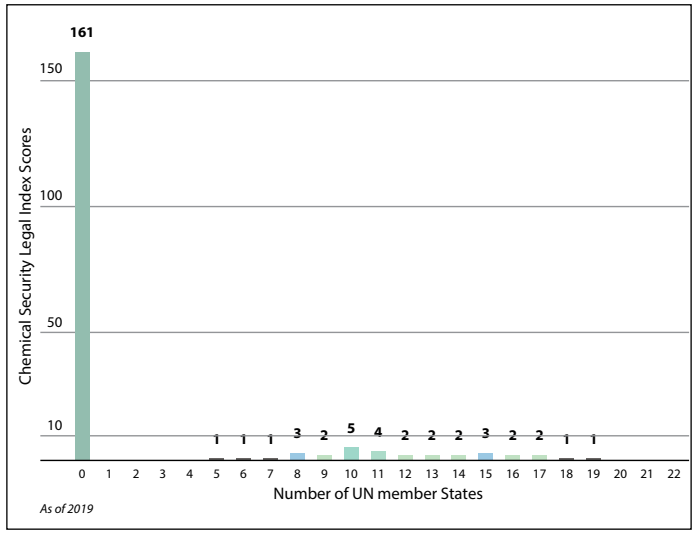


Figure 2. Chemical security elements within national legal frameworks

Among those 32 states with measures in place, the distribution of states ranges across the possible composite index scores in a near normal curve with an average composite index score of 11.8, with a low score of 5 and a high score of 19 (fig. 3). Ultimately, not one national legal framework includes every chemical security practice commonly discussed in the literature. Given the more than 4,000 declarable and inspectable facilities across 80 CWC state parties (which in itself may not reflect all facilities of concern as only 137 states have CWC-implementing legislation in place), the number of weak links in the worldwide chain of national legal chemical security frameworks is disturbingly high.[30]

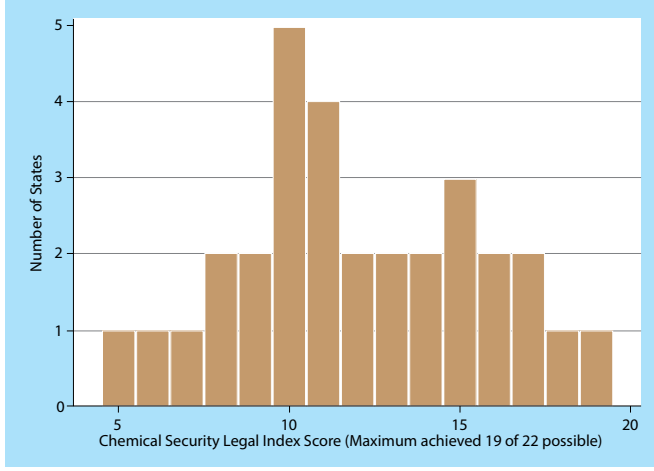


Figure 3. Chemical security elements in national legal framework

Nonetheless, the picture is not completely bleak. A simple test of the composite index scores of states with no CWC-declarable facilities and those with one or more declarable facilities suggests that states with declarable facilities under the CWC have a statistically higher composite index score than those states that have not declared such facilities.[31] In other words, our data analysis indicates that UN member states with CWC-declarable facilities tend to implement more chemical security practices in their national legal frameworks than states without such facilities. Perhaps the governing bodies of states with declarable facilities are more aware of the types of chemical facilities they have and the risks they pose. However, the relatively low mean index score for those states with relevant chemical facilities also indicates that almost all states, with or without these types of facilities, have considerable room for improvement.

The choropleth world map with state-level composite index scores exemplifies the usefulness of alternative data visualization (fig. 4).[32] The map clearly indicates that low levels of integration of effective chemical security practices in national legislation are not regionally determinant compared to, for example, UNSCR 1540 implementation. Variation exists even among European Union members, which one might not expect given the EU's legal harmonization. The choropleth map also shows that some relatively low-capacity states—such as Cuba, Indonesia, and Uganda—have more chemical security elements in their laws and regulations than do many high capacity states—such as Australia and Spain. The difference suggests that the international community will need a nuanced approach to understanding the determinants of chemical security legal frameworks. In many respects, the chemical security elements found most commonly in the 43 measures we identified seem closely linked to common elements for chemical safety, such as licensing and inspection of operators.

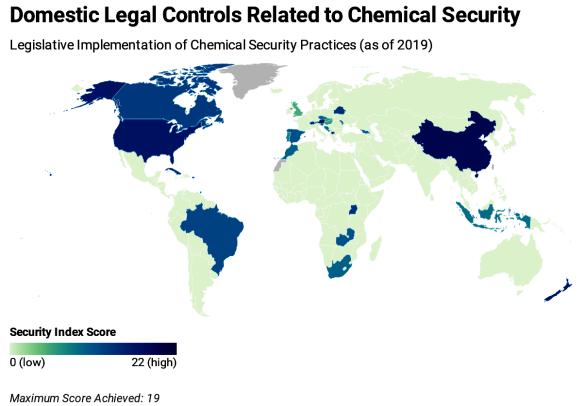


Figure 4. Composite chemical security index score

Given the limited correlation between the 21 chemical security practices identified in the literature and the 43 relevant chemical security laws and regulations, we sought to further verify if the chemical security practices are truly emerging standards. To better understand the context in which these chemical security standards were identified and used as variables to evaluate national legal frameworks, we compared the inauguration and amendment dates of laws with the lowest number of security measures against laws with the highest number of security measures. We also checked the publication dates of the chemical security literature we reviewed to determine if a relationship might exist between the number of security elements in legislation and when the legislation was established and/or updated. As multilateral discussions of chemical security have increased in recent years at the OPCW and other forums, one might expect that more recently adopted measures would align more closely with the emerging standards.[33]

As an exploratory effort, we selected laws based on each measure's low (3–6) or high (14–18) cumulative index score (again, most scores appear somewhere in the middle given the near normal distribution). We determined that some of the lowest-scoring chemical security measures in the dataset had been established with no new amendments in more than a decade—well before the increased use of chemical weapons or the rise in chemical security discussions. Legislation that exemplifies this trend include the following (table 1):

Table 1. Lowest-scoring measures and enactment/amendment dates

Country	Title of Legal Text	Year Enacted/ Amended	Composite Index Score
Hungary	Act LXXIV on the Management and Organization for the Prevention of Disasters and the Prevention of Major Accidents Involving Dangerous Substances of 1999	1999	5
Slovakia	Act No. 163/2001 Coll. of 5 April 2001 on Chemical Substances and Chemical Preparations, as amended in 2008	2001/2008	6
United Kingdom	Carriage of Dangerous Goods and Use of Transportable Pressure Equipment Regulations 2007	2007	4
United States	Public Law 109-294, an Act Making Appropriations for the Department of Homeland Security for the Fiscal Year Ending September 30, 2007, and for Other Purposes of 2006	2006	3

Meanwhile, some of the highest-scoring chemical security measures in the dataset were either adopted or amended in the last three years (table 2).

Table 2. Highest-scoring measures and enactment/amendment dates

Country	Title of Legal Text	Year Enacted/ Amended	Composite Index Score
Austria	Federal Law No. 145/1998 on the Transport of Dangerous Goods, as amended in 2018	1998/2018	14
Austria	Foreign Trade and Payments Act 2011, as amended in 2019	2011/2019	14
Liechtenstein & Switzerland	Federal Act on Protection Against Dangerous Substances and Preparations (Chemicals Act) of 15 December 2000, as amended in 2017	2000/2017	14
New Zealand	Hazardous Substances and New Organisms Act 1996, as amended in 2018	1996/2018	16
United States	Title 6, Chapter 1, Part 27 Chemical Facility Anti-Terrorism Standards of 2014, amended in 2019	2014/2019	14
United States	Department of Defense Instruction 5210.65, *Security Standards for Safeguarding Chemical Agents*, 19 January 2016	2016	15

Interestingly, after checking the publication dates of the resources on chemical security, we found that more than half were published within the last seven years (2012 to 2019). Given the trends between when laws are established/amended and how many security elements are included, we hypothesize that chemical security best practices are, in fact, *emerging*. They have been discussed regularly in literature for a few years, but their actual legislative implementation is still relatively new. It may be more likely that these types of standards will be found in newly written and recently updated laws. States that establish and amend chemical laws and regulations now appear to be thinking about security aspects more acutely and are determining more legislative measures to protect such high-risk materials from evolving threats.

Granted, outliers exist. China's State Council Order No. 591 on the Safety and Management of Hazardous Chemicals is the highest-scoring measure in the dataset with a cumulative index of 18 chemical security best practices, but it was last amended in 2011. Also, the United States' Title 33—Navigation and Navigable Waters, Part 105, is one of the lowest-scoring measures in the dataset with a cumulative index of four chemical security best practices, yet it was adopted in 2018. These outliers may exist due to the nature of the laws themselves. The Chinese State Council Order focuses on managing high-risk chemical materials, making it more likely

that significant consideration was given to incorporating security elements. Similarly, the Navigable Waters Act emphasizes securing maritime facilities, in which controls on toxic chemicals play only a small part.

Nonetheless, it appears that newer and updated legislation tends to include more chemical security elements than do older laws and regulations. This might prove a fruitful area for future research on how external shocks, such as chemical weapons use, and multilateral discussions, such as recent special meetings at the OPCW, affect developments in national implementation of international obligations and norms.

Essentially, our datasets show a limited relationship between chemical security best practices identified from the literature and their application in current chemical security laws and regulations worldwide, either at the national level or in individual measures. Despite this finding, we believe that these common practices may ultimately form emerging international standards for securing chemical weapons–related materials, facilities, and equipment. Based on the evidence when comparing the chemical security literature and the timing of those measures with the highest and lowest scores, these standards are relatively new and, thus, perhaps have not yet percolated into revisions of older laws and regulations.

Conclusions and Recommendations

It is essential to keep in mind that measuring a country's chemical security infrastructure in a field that lacks clearly determined international standards requires analyzing several variables, including chemical facility culture, physical protection equipment available, and national legal frameworks. Ultimately, legislation is an integral piece of a larger puzzle to begin the process of understanding global chemical security implementation practices.

Additionally, understanding why particular laws or regulations are adopted (or not) is itself a field of legal studies replete with its own controversies where determining intent is notoriously difficult.[34] Despite the very real threats from state and non-state actors (specifically, using chemical weapons, seeking chemical weapons–related materials, and targeting chemical facilities), the evidence suggests that a worryingly small number of UN member states have *any* legislation with an explicit requirement to secure chemical weapons–related materials and facilities. Determining why a state adopts a specific law or regulation for securing toxic chemicals—especially those of proliferation concern—would best be served by a series of in-depth case studies, as most are about current or recent public policy.[35] While a valuable avenue for future research, such projects go well beyond

our purpose here, that is, attempting to create a baseline of what legal measures all states now have or do not have as compared to emerging chemical security standards and best practices.

This study reveals that many states appear to have no chemical security measures in their national legal frameworks and suggests a greater determinant: the lack of authoritative international standards. Not only might states need more guidance on what to do before they act, but if they act now, their efforts may result in national systems out of alignment when global standards do eventually emerge. Although states may wish to—and likely should—update their existing frameworks to incorporate more of the chemical security elements already identified in the literature, they must do so in careful consideration of ongoing multilateral discussions on the topic.

Understanding precisely why states adopt or amend specific legislation related to chemical security is challenging and requires further study. Nonetheless, we can offer several potential vectors for future research. As noted earlier, states with more declarable facilities under the CWC have significantly higher scores statistically than states with no such facilities. However, a hypothesized relationship is not apparent, as states with more facilities might also face more domestic resistance to implementing costly security measures. One also might expect that states that have experienced terrorist incidents involving a weapon of mass destruction would be more likely to adopt chemical security measures than those that have not been subject to such attacks.

In the case of the United States, for example, the naming and timing of the adoption of the main regulation, the Chemical Facility Anti-Terrorisms Standards, emerged from an increased threat awareness. The US recognized that it faced a substantial threat from terrorists interested in causing mass casualties and mass disruption and that certain chemical facilities and their products were particularly vulnerable to terrorists.[36] Yet Japan, which suffered the atrocious chemical terrorism attacks by members of Aum Shinrikyo in 1995, does not thus far have a law or regulation with an explicit obligation to secure chemicals of proliferation concern.

The impact of international obligations might be another interesting avenue of exploration. Most of the measures in the dataset relate to dangerous or hazardous goods; however, only a few refer to international environmental or safety treaties or conventions, such as the Stockholm Convention on Persistent Organic Pollutants or the Convention concerning International Carriage by Rail. At least seven states that include chemical security obligations in their national legal measures specifically refer to the

CWC (Bosnia and Herzegovina, Cuba, Indonesia, Morocco, North Macedonia, Morocco, and Spain) even though the convention does not include obligations to secure chemicals of proliferation concern. Disturbingly, not one of the 43 measures refers directly to UNSCR 1540—unlike a range of nonproliferation measures in other risk areas, such as the European Union's export control regulations. As noted above, however, the absence of the resolution does not necessarily mean that it has not influenced a chemical security–related measure. India's Weapons of Mass Destruction Act of 2005, for example, does not mention resolution 1540 but was purposefully modeled on it. We suspect that the resolution has had more of an effect in policy areas where international standards exist, such as in the nuclear security field where states more clearly understand what they need to do to implement the obligations of UNSCR 1540.

Moreover, it is worth considering how UN member states incorporate chemical security into their national legal frameworks. Some states create completely new laws that emphasize managing chemical safety and security, such as China's State Council Order No. 591 on the Safety and Management of Hazardous Chemicals. Other states can build from existing and ancillary legislation, as demonstrated by the United States' Title 33—Navigation and Navigable Waters, Part 105. That China scored so high and the United States scored so low indicates that where there are standalone laws, they appear to be more comprehensive in their chemical security practices as opposed to ancillary legislation. Therefore, if national governments already have the political will (that can be sustained during the legislative process) to develop or enhance their chemical security legal infrastructures, then creating a comprehensive measure on chemical security would likely be the most effective approach. However, drafting and enacting new legislation takes extensive time and effort. Thus, if national governments are not in a position to augment their chemical security legislative infrastructures, then amending an existing (though ancillary) law may be an easier approach to start implementing some chemical security elements. Though amending existing legislation may not reap the most comprehensive chemical security benefits, it is a way to start a national conversation about chemical security.

Additionally, given the recent chemical attacks in Malaysia, the United Kingdom, and Syria, the risks and threats posed by controlled chemicals and nefarious actors may necessitate that national governments act sooner rather than later to implement security standards. Amending ancillary legislation could be an efficient path forward in the presence of many competing priorities in legislative bodies. Implementing chemical security

into national legal frameworks is critical, but how states do it will largely depend on their national realities.

Finally, from the small group of UN member states that have laws or regulations with an explicit chemical security requirement, a significant variation is also evident in how states approach chemical security laws and regulations. Unlike the standards outlined by the IAEA for nuclear and radiological security, the chemical security field lacks such internationally accepted and authoritative guidance. Given the few chemical security measures and laws in place that incorporate the emerging standards identified here, we believe there is an urgent need for action. The international community must determine what constitutes standards for securing chemical weapons–related materials and facilities (through the OPCW) and how to best incorporate them into national legal frameworks. Ultimately, we recommend that CWC state parties that have developed strong national legal frameworks for chemical security demonstrate leadership in this global security arena by beginning a sustained dialogue on what chemical security and its associated legal frameworks should look like. Once states parties signal their interest in and prioritize chemical security practices, the OPCW can take that as a cue to begin creating international standards in this space. The OPCW is the main international body that needs to develop internationally accepted standards for chemical security. Though doing so sounds like a lofty and contentious road to travel, CWC state parties have cooperated during fractious times and despite divisive issues—as demonstrated by the recent (June 2020) additions to the CWC Schedules, which were the first since their establishment in 1993.

The contributions of civil society, academia, and industry have fostered a more robust and frequent discussion of potential chemical security standards and effective practices in the OPCW and elsewhere. If the OPCW does develop widely accepted standards, a similarly diverse set of stakeholders should have an important role in (1) raising awareness of the new standards and practices; (2) helping states in their implementation; and (3) further tailoring the guidance to the specific circumstances—such as threat and risk profiles—of different national legal jurisdictions, industry sectors, and modes of scientific collaboration. Developing international chemical security standards can help countries better protect their societies, industry, and environment from chemical terrorism and chemical warfare. SSQ

Richard T. Cupitt

Dr. Cupitt is a Senior Fellow and director of the Partnerships in Proliferation Prevention Program at the Henry L. Stimson Center. He has served as an expert for the United Nations Security Council, worked in the US State and Commerce Departments, and held numerous academic positions.

Mary C. Vecellio

Ms. Vecellio is a research associate with the Partnerships in Proliferation Prevention Program at the Stimson Center. Her research focuses on WMD nonproliferation with specific focus on United Nations Security Council Resolution 1540 assistance programs and chemical weapons–related material security.

Notes

For additional references, see the online bibliography at https://www.airuniversity.af.edu/.

1. United Nations, Office of Disarmament Affairs, *Protocol for the Prohibition of the Use in War of Asphyxiating, Poisonous or Other Gases, and of Bacteriological Methods of Warfare*, signed at Geneva, 17 June 1925, https://www.un.org/.

2. Richard M. Price, *The Chemical Weapons Taboo* (Ithaca: Cornell University Press, 1997); and Jonathan B. Tucker, *War of Nerves: Chemical Warfare from World War I to Al-Qaeda* (New York: Pantheon, 2006).

3. "OPCW Executive Council Condemns Chemical Weapons Use in Fatal Incident in Malaysia," Organisation for the Prohibition of Chemical Weapons, 10 March 2017, https://www.opcw.org/. See also David Bond, "Third Russian GRU Agent Linked with Novichok Attack on Skripals," *Financial Times*, 14 February 2019, https://www.ft.com/.

4. Laura Howes, "New Nerve Agents Added to the Chemical Weapons Convention," *Chemical & Engineering News*, 2 December 2019, https://cen.acs.org/policy/.

5. Markus K. Binder and Gary A. Ackerman, "Pick Your POICN: Introducing the Profiles of Incidents Involving CBRN and Non-State Actors (POICN) Database," *Studies in Conflict & Terrorism*, 2019, https://doi.org/10.1080/1057610X.2019.1577541.

6. Simon Marks, "Belgian Exporters Found Guilty of Sending Chemicals to Syria," *Politico*, 19 April 2019, https://www.politico.eu/.

7. The compendium also fulfilled one of the action items of the G7 Global Partnership's Chemical Security Sub-Working Group to implement its strategic vision. See Global Partnership Against the Spread of Weapons and Materials of Mass Destruction, Chemical Security Sub-Working Group, "Strategic Vision," accessed October 2020, https://www.gpwmd.com/.

8. International Atomic Energy Agency (IAEA), *Code of Conduct on the Safety and Security of Radioactive Sources* (Austria: IAEA, 2004), https://www-pub.iaea.org/. See also International Atomic Energy Agency (IAEA), *Guidance on the Import and Export of Radioactive Sources* (Vienna: IAEA, 2012), https://www-pub.iaea.org/.

9. See Organisation for the Prohibition of Chemical Weapons (OPCW), "Eliminating Chemical Weapons: Committed to Complete and Verifiable Destruction," accessed October 2020, https://www.opcw.org/.

10. Organisation for the Prohibition of Chemical Weapons (OPCW), "OPCW-Facilitated Report Recommends International Chemical Security Coordination," 7 December 2017, https://www.opcw.org/. See also OPCW, "Expert Workshop on International Chemical Security Coordination: Informal Summary," 7 December 2017, https://www.opcw.org/.

11. Sources included international guidelines and recommendations, like the *Hague Ethical Guidelines* and the UN's *Recommendations on the Transport of Dangerous Goods*; regional and country-specific regulations and policy plans, such as the United States Chemical Facility Anti-Terrorism Standards (CFATS) and the Code of Conduct for the Practice of Chemistry in the Middle East and Northern Africa; chemical industry initiatives, like the Responsible Care Codes

for both the United States and European Union; and chemical laboratory management programs. (For the full list, see the bibliography at https://www.airuniversity.af.edu/.)

12. Chemical Facility Anti-Terrorism Standards (CFATS), 6 CFR Part 27, 1 January 2012, Cornell Law School, Legal Information Institute, https://www.law.cornell.edu/.

13. US Department of Homeland Security, *Risk-Based Performance Standards Guidance: Chemical Facility Anti-Terrorism Standards* (Washington, DC: Department of Homeland Security, May 2009), 7, https://www.dhs.gov/.

14. American Chemistry Council, "Responsible Care Security Code," 6, https://responsiblecare.americanchemistry.com/. See also Cefic, "European Responsible Care Security Code Guidance and Best Practice for the Implementation of the Code," Adopted in 2010 (Brussels: Cefic, 2013), 4, https://www.opcw.org/.

15. Cefic, "European Responsible Care Security Code Guidance," 4.

16. National Research Council, *Promoting Chemical Laboratory Safety and Security in Developing Countries,* (Washington, DC: The National Academies Press, 2010), chap. 3, https://www.nap.edu/read/.

17. Sources include US Department of Homeland Security, *Chemical Sector Security Awareness Guide: A Guide for Owners, Operators, and Chemical Supply-Chain Professionals* (Washington, DC: Department of Homeland Security, September 2012), https://www.dhs.gov/; Organisation for the Prohibition of Chemical Weapons, *The Hague Ethical Guidelines* (The Hague: Organisation for the Prohibition of Chemical Weapons, n.d.), https://www.opcw.org/; and Lisa Moran and Tina Masciangioli, eds., *Chemical Laboratory Safety and Security: A Guide to Prudent Chemical Management* (Washington, DC: The National Academies Press, 2010), http://dels.nas.edu/.

18. Deputy Director-General Hamid Ali Rao, OPCW (speech, Global Summit on Chemical Safety and Security, Kielce, Poland, 18 April 2016), 3, Organisation for the Prohibition of Chemical Weapons, https://www.opcw.org/.

19. Code of Conduct for the Practice of Chemistry in the Middle East and Northern Africa, in *Compilation of Codes of Ethics and Conduct: A Collection of Codes of Ethics and Conduct (and Related Guidelines) Compiled for the Workshops on Ethical Guidelines for the Practice of Chemistry under the Norms of the Chemical Weapons Convention*, ed. Organisation for the Prohibition of Chemical Weapons (The Hague: Organisation for the Prohibition of Chemical Weapons, September 2015), 30–33, https://www.opcw.org/.

20. Committee on Reducing the Threat of Improvised Explosive Device Attacks by Restricting Access to Chemical Explosive Precursors, *Reducing the Threat of Improvised Device Attacks by Restricting Access to Explosive Precursor Chemicals* (Washington, DC: The National Academies Press, 2018), chap. 4, https://www.nap.edu/. See also Cefic, "European Responsible Care Security Code," 4.

21. US Department of Homeland Security, *Risk-Based Performance Standards Guidance*, 41.

22. American Chemistry Council, "Responsible Care Security Code," 4; see also Cefic, "European Responsible Care Security Code Guidance," 4.

23. Moran and Masciangioli, *Chemical Laboratory Safety and Security; and* National Research Council, *Promoting Chemical Laboratory Safety,* chap. 3.

24. US Department of Homeland Security, *Risk-Based Performance Standards Guidance*, 109. See also Committee (report), *Reducing the Threat of Improvised Device Attacks,* 76, 92.

25. American Chemistry Council, "Responsible Care Security Code," 4.

26. These include the eight fields for securing chemical weapons–related materials in production, use, storage, and transport. Each topic has two data points, one for a legal measure and another for evidence of enforcement, usually in terms of a legal penalty provision. See 1540 Committee, "Approved 1540 Committee Matrix Template," 2018, https://www.un.org/.

27. A special thanks to Jessica Hartnett, Joshua Quinn, Ashley Fischer, and Adam Wilson from the University of Minnesota Humphrey School of Public Affairs and Doug Seals and Jacquelyn Harms, Spring 2019 interns at the Stimson Center.

28. UN Security Council, *Final Document on the 2016 Comprehensive Review of the Status of Implementation of Resolution 1540 (2004),* S/2016/1038, December 2016, 69, https://undocs.org/.

29. The authors are currently developing an online searchable database of the 43 chemical security measures that includes both our assessments of these laws against the 21 chemical security elements as well as a compendium of the full text of each law, regulation, or similar measure collected. We view the compendium and their assessments more as a dictionary of use with an aim to create a baseline for a discussion of standards rather than to suggest these 21 elements are standards.

30. Organisation for the Prohibition of Chemical Weapons, "OPCW by the Numbers," accessed October 2020, https://www.opcw.org/.

31. The group mean was 1.06 versus 3.31 and the t-score was -3.35 (with 187 degrees of freedom given the 189 state parties to the CWC).

32. A choropleth map is a thematic map that has shaded or patterned areas proportionate to a statistical variable that represents an aggregate summary of an area's geographic characteristics. In this case, the choropleth map is used to visualize countries' statistical implementation of up to 21 chemical security best practices in national legal measures.

33. For example, see Chemical Security Sub Working Group of the G7, Global Partnership against the Spread of Weapons of Mass Destruction, and the Global Chemical Congress at INTERPOL.

34. See, for example, Antonin Scalia and Bryan A. Garner, *Reading Law: The Interpretation of Legal Texts* (St. Paul: Thomson/West, 2012); Victoria Nourse, *Misreading Law, Misreading Democracy* (London: Harvard University Press, 2016); and Neil MacCormick and Robert S. Summers, *Interpreting Statutes: A Comparative Study* (Abingdon, UK: Routledge, 2016).

35. Robert K. Yin, *Case Study Research and Applications: Design and Methods* (Los Angeles: SAGE, 2018).

36. United States Government Accountability Office, *Chemical Terrorism: A Strategy and Implementation Plan Would Help DHS Better Manage Fragmented Chemical Defense Programs and Activities,* GAO-18-562 (Washington, DC: Government Accountability Office, August 2018), https://www.gao.gov.

PERSPECTIVE

Nuclear-Weapon-Free Zones and Contemporary Arms Control

Elizabeth Mendenhall

Abstract

Nuclear-weapon-free zones (NWFZ) can offer increased nuclear security and stability for the "second nuclear age." This article surveys existing NWFZs and describes their goals and the role of nuclear-armed states in creating and maintaining the zones. Finally, it evaluates the prospects for creating three new NWFZs as a productive contribution to disarmament and nonproliferation including nuclear zero.

*****

Whether the "second nuclear age" is more dangerous than the Cold War is a hotly contested topic among scholars and practitioners of nuclear security. On one hand, a recent issue of the *Bulletin of the Atomic Scientists* warns of a new nuclear arms race as all nuclear states except North Korea actively modernize and upgrade their existing arsenals in competing efforts to alter the balance of military power.[1] The strategic situation is complicated by investments in new nonnuclear strategic weapons such as offensive cyber weapons, precision-strike missiles, antiballistic missile systems, anti-satellite weapons, and artificial intelligence technologies. Regional security environments are deteriorating, especially in and around the Asian continent, as opponents of the United States use "gray zone" strategies to push back against US influence and extended deterrence.[2] In the United States, the 2018 *Nuclear Posture Review* is clear in recommending investment in submarine-launched cruise missiles and low-yield nuclear weapons. These modernization efforts feed on one another as investments in the nuclear forces of one state spur adjustments and investments by others, connecting global and regional nuclear rivalries in a single dangerous dynamic.[3] Two dark predictions of the second nuclear age—increased proliferation and intensification of rivalries—seem to be coming true.

On the other hand, several objective metrics suggest that the second nuclear age has actually been more stable and secure than the first in terms of the risk of nuclear use. Political scientist Christopher Fettweis argues

that the current nuclear era is "better in most ways" and that this fact is "plain and irrefutable."[4] In addition to reductions in the frequency and intensity of many forms of violence, he points to the lack of leakage from nuclear weapon states—both intentional and as a result of theft—and the prevalence of "reverse nuclear proliferation." Anxiety about the supposedly heightened risks of the current nuclear era, he explains, is largely the result of golden age thinking and imperfect memories of the Cold War—as well as overhyped concern about a nuclear North Korea that he argues has generally been rational and restrained. The number of nuclear actors is the same as at the end of the Cold War (swapping South Africa for North Korea) and the number of warheads drastically lower. From this perspective, the nuclear security situation is better than before.[5]

These sharply differing assessments of the contemporary nuclear era suggest the importance of a renewed push for arms control. Arms control refers to any efforts to "limit the numbers, types, or dispositions of weapons."[6] International arms control agreements typically involve reciprocal, mutual constraints on weapons capabilities by at least two states. Nuclear-weapon-free zones (NWFZ) formalize this mutual constraint on a regional basis. They serve the purposes of nonproliferation and disarmament, thereby achieving security from nuclear weapons through institutionalized mutual restraint. A reassessment of regional NWFZs is especially warranted at this time because the current outlook for arms control is considered bleak, in part because of credibility issues with the Nuclear Non-Proliferation Treaty (NPT).[7] This article surveys existing NWFZs, with a special focus on their goals and the role of nuclear weapon states in creating and maintaining them. Finally, the article evaluates proposals for three new NWFZs as a productive contribution to disarmament and nonproliferation. New or expanded NWFZs can make a productive contribution to nuclear security and stability, but they may be most useful and feasible in a modified form.

Existing Nuclear-Weapon-Free Zones

Nuclear-weapon-free zones are a core part of the larger nuclear control regime. The concept of NWFZs pre-dates the NPT but is explicitly endorsed by it. Article VII formally defines the right for states to create regional NWFZs "to assure the total absence of nuclear weapons in their respective territories."[8] The idea of NWFZs received significant support at NPT review conferences throughout the 1970s and 1980s.[9] In its most basic version, a full *nuclear-weapon-free zone* is defined in United Nations General Assembly (UNGA) Resolution 3472 B (1975) as "any zone, rec-

ognized as such by the General Assembly of the United Nations, which any group of states, in the free exercise of their sovereignty, has established by virtue of a treaty or convention whereby: (a) The statute of total absence of nuclear weapons to which the zone shall be subject, including the procedure for the delimitation of the zone, is defined; [and] (b) An international system of verification and control is established to guarantee compliance with the obligations deriving from that statute."[10]

The same resolution also defines the "principal obligations" that nuclear weapon states have toward NWFZs and the states included in those zones. Nuclear states are required ("shall undertake") to participate in a treaty, convention, or protocol that obligates them to respect the "total absence of nuclear weapons in the zone," refrain from supporting any violations of the NWFZ, and "refrain from using or threatening to use nuclear weapons against the States included in the zone."[11] In other words, the nuclear weapon states are technically obligated to provide formal negative security assurances to the members of a NWFZ. Most NWFZs contain protocols for this purpose.

The creation of NWFZs has been generally supported by United Nations agencies and processes, although most individual NWFZs have been negotiated through regional initiatives and institutions.[12] A basic model for regional NWFZs emerged during the Cold War, so most extant agreements share similar design features. Depending on how one counts, nine NWFZs exist today. The zones share a core requirement of banning the deployment and use of nuclear weapons in a particular zone, but they vary significantly among other metrics: what activities they ban, how they calculate the covered zone, verification mechanisms, connection to international bodies like the International Atomic Energy Agency (IAEA) and NPT, and buy-in from extrazonal states—especially nuclear weapon states. There are two basic types of NWFZs: agreements covering the global commons and those covering groups of sovereign territorial states. Several states have declared a unilateral NWFZ within their territories, most notably Mongolia, but the prototypical NWFZ—and the one defined by the UNGA in 1975—involves groups of states.

The first NWFZs were established in global commons: Antarctica, outer space, and the international seabed (table 1). In each of these cases, the nuclear superpowers sought to cooperatively restrain themselves to avoid the expansion of the Cold War arms race into new parts of the planet, which could lead to dangerous and expensive new forms of competition.[13] At the times of negotiation, neither the United States nor the Soviet Union was strategically or financially committed to new nuclear

deployment configurations in these spaces. The Antarctic, Seabed Arms, and Outer Space Treaties were "agreements to maintain the status quo."[14] These three agreements are different than the archetypical regional NWFZ because their zones are coextensive with global commons, negotiations were initiated by the nuclear superpowers, and the treaties cover broader content and enjoy wider participation by the international community.

Table 1. Current nuclear-weapon-free zone agreements (in force)

				What is prohibited for nuclear weapons?					
Year in force	**Agreement**	**Zone**	**Type**	**Manufacturing**	**Testing**	**Storage/Basing**	**Transit**	**Waste**	**Peaceful Explosions**
1961	Antarctic Treaty	South of 60 degrees	Commons		X	X		X	X
1967	Outer Space Treaty	Global	Commons		X	X			
1968	Latin America NWFZ	Member territories and ocean areas	NWFZ	X	X	X			
1971	Seabed Arms Limitation	Seabed beyond 12 nm	Commons		X	X			
1986	South Pacific NWFZ	Member territories and ocean areas	NWFZ	X	X	X		X	X
1997	Southeast Asia NWFZ	Member territories and ocean areas	NWFZ	X	X	X		X	X
2009	African NWFZ	Member territories	NWFZ	X	X	X		X	X
2009	Central Asia NWFZ	Member territories	NWFZ	X	X	X		X	X

Regional NWFZs cover member state territories, including territorial seas and airspace (see table 1). Almost two-thirds of UN member states are also members of a regional NWFZ. The first regional NWFZ in Latin America "served as both a call and a blueprint" for additional NWFZs, and subsequent agreements explicitly worked from the model it created.[15] Later agreements added new features, such as prohibiting nuclear waste and peaceful explosions and requiring cooperation in environmental remediation of nuclear waste areas.[16] These agreements too were praised for their "strong message" and "demonstration effect" for other regions to establish their own NWFZs.[17] Proponents of arms control hoped that

NWFZs could grow incrementally and network together to create a zone of peace that would cover most or all of the planet.

Goals of Nuclear-Weapon-Free Zones

NWFZs are explicitly intended to contribute to the larger project of total global disarmament. The first regional NWFZ, the Treaty of Tlatelolco, was more strongly motivated by disarmament than nonproliferation as an objective.[18] The establishment of a NWFZ is a gradual, incremental approach to disarmament by slowly and painstakingly ruling out portions of the planet for nuclear deployment and use and by locking in the status quo after disarmament by regional states.[19] In the case of the Central Asia NWFZ, the agreement made permanent (indefinite) the prior relinquishment of Soviet nuclear weapons systems and infrastructure. Other former proto-nuclear states that are current members of the NWFZ include Brazil, Argentina, South Africa, and Libya. The protocols attached to the NWFZ also represent a formal request for negative security assurance from the five recognized nuclear powers, which, when acceded to, restrain the possibilities for nuclear use. Such legally binding negative security assurances are intended to provide more certainty and reliability than unilateral security assurances. Beyond the practical goal of reducing scenarios for nuclear use, the formal and public commitment to the goal of nuclear disarmament may provide a degree of prestige to nonnuclear states, which through their endorsement of and participation in the NWFZ are taking a principled stance against the deployment and use of nuclear weapons. This formal position also strengthens the norms against nuclear weapons possession and use both globally and regionally.

NWFZs are also intended to enhance regional security. In many cases, the pursuit of a NWFZ is described as a matter of urgency to stave off emerging tensions or risks.[20] Most basically, NWFZs promote dialog and enhance confidence among member states. But they are also designed to regulate the deployment of nuclear weapons.[21] The NPT did not prohibit the nuclear powers from stationing weapons in the territories of otherwise nonnuclear states, meaning that a region without a nuclear weapon state could still contain nuclear weapons.[22] Historically, when one nuclear weapon state establishes a military presence in a region, this invites competition from other nuclear weapon states, "thereby turning the region into a zone of tension and confrontation."[23] Even if tensions are low, the mere existence of a nuclear weapon in a region could be considered dangerous because of the risk of accidents or theft (although this risk may be overestimated). NWFZs are a way for member states to prevent nuclear

risks from spilling into their region by prohibiting basing, stationing, or installations managed by nuclear powers.

Verification mechanisms vary between NWFZ agreements, including "national technical means," special bilateral or regional organizations, reliance on IAEA safeguards, and combinations of all of these. The system created by the Tlatelolco Treaty—which relies on the IAEA, a new regional organization, and a bilateral commission between Argentina and Brazil—is often identified as a successful and useful model.[24] The addition of new mechanisms for monitoring and verification is a key advantage of NWFZs compared to simply relying on existing global nonproliferation regimes. In general, NWFZ agreements expand the set of potential violations, add functions associated with information sharing and consultation, and create procedures for complaints or potential violations.[25] Verification mechanisms are especially important for ensuring the durability of a NWFZ in the context of mistrust or conflict between regional actors.[26] In some cases, there are separate provisions for verification of the dismantling and destruction of existing nuclear devices and the conversion of nuclear production facilities. No regional NWFZ includes a mechanism for verifying the activities of extrazonal states. But NWFZs can place formal conditions on the relationship between nuclear weapon states and their nonnuclear regional allies. Even when states are formally under the "nuclear umbrella" of a nuclear ally, they can prohibit the basing of that ally's weapons in their territory through participation in the NWFZ agreement.

NWFZs also, somewhat obviously, serve the goals of nonproliferation by prohibiting the emergence of nuclear states in the regions they cover. Participation in NWFZs has been described as "one of the most practical steps that non-nuclear weapon states can take to help bolster the nonproliferation regime."[27] Arguably, NWFZs represent an alternate track from the NPT, a means to create a nonproliferation regime through bottom-up agreements originating in and largely covering the Global South.[28] Many proponents of the NWFZ envision a set of expanding and interlocking regional NWFZs that will cover progressively larger parts of the planet.[29] Because the knowledge of nuclear weapons technology is widespread, the main obstacle to horizontal proliferation is lack of motive as opposed to lack of means. There exists a substantial, and increasing, "nuclear overhang"—the gap between the number of states that could acquire versus have acquired nuclear weapons.[30] This overhang also represents the possibility of rapid horizontal proliferation. Although most states that do not have nuclear weapons also do not want them, it is possible that in the future this situation may shift due to changes in leadership or the regional

security environment. The establishment of the NWFZ is an attempt to formally and legally solidify the nonnuclear status of particular regions. Although a NWFZ cannot physically prevent any given state from acquiring nuclear weapons, it can raise the reputation costs of doing so and also—through verification mechanisms—help ensure that other regional actors are aware of any new horizontal proliferation.

Every part of the planet is physically vulnerable; "geography provides little protection in the nuclear age."[31] This reality is the basis of criticisms of NWFZs as "politically vacuous" and "worse than mere scraps of paper" because a "nuclear-free zone" on paper is not the same as a "nuclear-safe zone" in practice.[32] NWFZs can be understood as an attempt to rewrite the geography of nuclear strategy and risk by adopting a regional form of co-binding, or mutual institutional restraint, that constructs social and legal limits to the geography of nuclear weapons.[33] And at the very least, NWFZs can reduce potential nuclear risks emanating from the region itself. Successfully imposing a geographic barrier to nuclear threats and risks requires the participation of the nuclear weapon states. The basic challenge is that the geographic reach of nuclear technology is global, so prohibiting the basing of nuclear weapons in a region does not reduce, let alone eliminate, the region's vulnerability to nuclear use. The design of all regional NWFZs includes protocols to garner participation, and require a commitment, of the five recognized nuclear powers (United States, Russia, China, France, United Kingdom). Like the NPT, however, existing regional NWFZs do not formally include the nuclear outlier states—India, Pakistan, Israel, and North Korea. Thus, NWFZs cannot fully achieve regional security from nuclear weapons without the participation of nuclear-weapon states.

Role of Nuclear Weapon States

The role and participation of nuclear weapon states in NWFZs are varied, with decreasing support over time. The Antarctic, Outer Space, and Seabed Arms Limitation Treaties have been extremely successful in limiting the scope of nuclear deployment without generating significant controversy and concern related to possible violations. This accomplishment is largely because the nuclear weapon states were centrally involved from the beginning, and they willingly accepted limitations on their military forces to avoid a costly and strategically bereft arms race. The Seabed Arms Limitation and Outer Space Treaties also allowed for considerable leeway in terms of transit of nuclear weapons and militarization in general. Traditional NWFZs face a bigger challenge—getting extrazonal nuclear

weapon states to commit to restrictions on transit, deployment, and use of their existing nuclear weapons. The nuclear weapon states initially supported the creation of regional NWFZs, as evidenced by their formal encouragement, participation in protocols, and inputs during the negotiation phase. But in general, their support and participation have weakened over time as NWFZs cover larger portions of the planet and get closer to areas of strategic interest.[34] Additionally, the four nuclear outlier states have not been invited or attempted to participate in NWFZs. While the recognized nuclear weapon states often express support for NWFZs, their rhetoric usually exceeds their practical and formal commitments.[35]

As regional actors attempt to apply the NWFZ model to more challenging political and security environments, the required commitment from nuclear-armed states becomes greater. Regions uncovered by existing NWFZs include the territory of many nuclear and ally states under the umbrella of extended nuclear deterrence. NWFZs represent a fundamental challenge to the "very legitimacy of nuclear possession" and are therefore apparently incompatible with nuclear deterrence strategies.[36] Indeed, NWFZs are understood as a "fundamentally different security alternative" to nuclear deterrence.[37] Although the reliability and utility of nuclear deterrence theory and strategy have been increasingly questioned in the second nuclear age, it remains a lodestar for the military and grand strategy of existing nuclear weapon states and their allies.[38]

Current participation by nuclear weapon states in NWFZs centers on the ratification of protocols to each agreement (table 2). None of the NWFZs cover the central sovereign territory (the metropole) of an existing nuclear weapon state. But several NWFZ negotiations were motivated in part by a history of harmful and damaging nuclear weapons activity by nuclear-armed states, especially testing.[39] All the regional NWFZs include a protocol for nuclear weapon states, and some include another protocol for extrazonal states that control a territory in the region. The negative security assurance protocols commit the five recognized nuclear weapon states to abide by the dictates of the NWFZs, including not helping any member state to violate the agreement, not stationing or storing nuclear weapons in the zone, and not using or threatening to use nuclear weapons against states in the zone. When these protocols are not signed or ratified by the nuclear weapon states, the lack of buy-in undermines the effectiveness of NWFZs as regional security frameworks.[40]

Table 2. Ratifications of security assurance protocols

			Countries ratifying security assurance protocols				
Year in force	**Treaty name**	**Zone**	**United States**	**Russia**	**China**	**United Kingdom**	**France**
1968	Tlatelolco Treaty	Latin America NWFZ	X	X	X	X	X
1986	Rarotonga Treaty	South Pacific NWFZ		X	X	X	X
1997	Bangkok Treaty	Southeast Asia NWFZ					
2009	Pelindaba Treaty	African NWFZ		X	X	X	X
2009	Semipalatinsk Treaty	Central Asia NWFZ		X	X	X	X

Negative security assurances are a primary way that NWFZs contribute to disarmament, as opposed to just nonproliferation, and thereby enhance regional security. But ultimately, the granting of a negative security assurance is done at the discretion of the nuclear weapon state. The first regional NWFZ, Latin America, is the only one with full participation in its protocols. The South Pacific and African NWFZ protocols also prohibit use or threatened use against territories within the region that extrazonal states are internationally responsible for, such as Diego Garcia in the Indian Ocean (the location of a US/UK military base). The Bangkok Treaty protocol goes even further and commits nuclear weapon states not to use or threaten to use nuclear weapons "within the Southeast Asia Nuclear Weapon-Free Zone," which includes the continental shelves and exclusive economic zones (EEZ) of member countries. This added detail is one of the main reasons that none of the nuclear weapon states have ratified the Bangkok protocol.[41] When nuclear weapon states do ratify protocols, they often include interpretative declarations to clarify what they believe they are still able to do.

The United States has been particularly reticent to participate in protocols, although it has assisted in negotiation of this aspect of NWFZs.[42] Its only ratification of a NWFZ security assurance protocol, in Tlatelolco, was significant in part because the United States committed to denuclearize its territories in Puerto Rico, Guantanamo, and the US Virgin Islands.[43] By the time of the next agreement, Rarotonga, the United States had decided that it did not want to set a precedent of participation, as the number of NWFZs was apparently growing and participation "would po-

tentially undermine its policy of deterrence, and . . . limit its future ability to meet its security commitments worldwide."[44] Negative security assurances are also incompatible with the concept of extended deterrence, or nuclear umbrellas, a global strategy used by the United States to cement its network of alliances.[45] Although some countries under the US nuclear umbrella, such as Australia, have participated in NWFZs, in general these states are more hesitant to endorse strong calls for disarmament and the withdrawal of extended deterrence commitments.[46] This situation presents major obstacles to the effective functioning of existing NWFZs and raises serious doubts about the creation of new ones.

Nuclear weapon states—especially the United States and Russia—are particularly sensitive about potential barriers to transit of nuclear weapons through or across particular regions. Indeed, as the amount of the planet covered by NWFZs expands, nuclear weapon states have tended to lose enthusiasm over possible new restrictions on the movement of nuclear-armed delivery vehicles.[47] But four of the five regional NWFZs explicitly grant member states discretion over the transit of nuclear-armed ships and aircraft through their territories, and the other—Tlatelolco—has generated an informal consensus interpretation that member states also have this right. And since there is no verification mechanism for extrazonal states, and nuclear states rarely disclose whether their vessels are armed with nuclear weapons, it is unlikely that member states of NWFZs could effectively prohibit nuclear transit through all regional waters. Despite these practical realities, nuclear weapon states are unlikely to accept any agreement that would draw attention to or delegitimize the transit of nuclear-armed vehicles such as submarines.

This sensitivity about restrictions on nuclear transit is connected to broader concerns about maritime navigation. The Pelindaba, Bangkok, and Rarotonga agreements do include a formal provision stating that nothing in the treaties will "prejudice" the rights or exercise of the rights granted to states by the UN Convention on the Law of the Sea, including the principle of freedom of the seas and rights of innocent passage. But there are long-standing disagreements about the meaning of innocent passage, including about vessels transporting nuclear material.[48] So any reach of a NWFZ into maritime territory raises questions and concerns. Russian signature of the African NWFZ Protocol was delayed by uncertainty about whether the agreement would fully apply to the US base on the UK's Diego Garcia. The Southeast Asia agreement formally covers the EEZ and continental shelf ocean territory. The United States has also expressed concern about the coverage of the South Pacific NWFZ, which

includes large portions of the EEZ and high seas. Although this provision is understood as only an indication of "the optimal area of application," as opposed to a formally binding provision on ocean users, even the suggestion that such an outcome is desirable raises concerns for maritime nuclear weapon states.[49]

An Uncertain Future: Proposed Nuclear-Weapon-Free Zones

The existing NWFZs have a mixed, but positive, track record of helping to achieve nonproliferation and disarmament goals, especially those agreements that cover the global commons. Given the bleak outlook for unilateral, bilateral, or multilateral arms control among the nuclear weapon states, can geography-based prohibitions on nuclear weapons contribute productively to the arms control agenda? If the ultimate goal is coverage of the entire planet, the NWFZ model—in terms of the approach to negotiation and design of the instrument—will have to adjust to more challenging circumstances. New agreements may have to address currently deployed nuclear weapons by states that would prefer to maintain their nuclear forces and force structure, often as part of a nuclear deterrence strategy. Many experts suggest more tailored and limited NWFZs, moving away from the rigid twentieth-century idea of a "pristinely pure" NWFZ without any nuclear weapons–related activities.[50] The zonal element can be maintained and applied to other types of prohibitions and requirements, with the goal of increasing transparency and trust, limiting nuclear assets, and developing monitoring and verification practices. NWFZ territories could also be drawn creatively, such as within subzones of a country or countries.[51] Three potential NWFZs are currently on the table, with support from stakeholders and other proponents and varying levels of interest from the regional parties—the Middle East, the Arctic, and Northeast Asia.[52]

Each of these regions is subject to long-standing, ongoing, and/or emerging tensions among great and middle powers. These tensions are a central impediment to the negotiation of additional NWFZs, along with the power and prestige that incentivize nuclear weapon states to maintain their arsenals.[53] Historically, NWFZs have been established only after the "resolution of outstanding political and security issues."[54] These complicated regional security environments suggest that any successful NWFZ will probably need to be negotiated gradually and adopted incrementally, and its final design may need to depart from the dominant model of NWFZs in innovative ways. A NWFZ that achieves the goals outlined above will also require a robust and engaged monitoring, verification, and

compliance mechanism that can reliably make judgements about nuclear status, despite the many ways that states subject to inspection can delay or deny access, destroy evidence, conceal facilities, or provide incomplete or inaccurate reports.[55] These are challenges for the institutional design of new NWFZs, but the most proximate issue may be how to get the incremental and region-specific process started.

The Middle East NWFZ

The proposal with the most international attention is that of a Middle East weapon-of-mass-destruction-free zone (WMDFZ). The basic goal is to prevent a catastrophic regional war that uses WMDs. The proposed scope of the zone includes Iran, Israel, and all or most members of the Arab League.[56] The idea for a Middle East NWFZ was first proposed by Iran in the early 1970s and quickly taken up by Egypt.[57] There was little progress until a renewal of interest at NPT review conferences in the 1990s and 2000s. At the 2010 NPT Review Conference, participants endorsed a proposal to convene a conference in 2012 to move forward on the WMDFZ idea, with Finland appointed as the facilitator. The conference was cancelled, however, because states could not agree on preconditions for the meeting and because of a general decay of regional security conditions. The topic of a Middle East WMDFZ was again a focus of the 2015 NPT Review Conference, with strong support from Iran, among others. But no final document was adopted, and annual work meetings have failed to produce meaningful progress, in part because Israel and the United States did not attend. A key disagreement concerns the conditions of Israeli participation.

The Middle East is a challenging case for regional disarmament because it contains at least one nuclear power with strong incentives to retain nuclear forces and must also confront deep-seated animosities, mistrust, and tensions between regional actors. Support for a NWFZ in the region is broad but shallow; each regional actor imagines a version of the agreement that includes its preferred preconditions.[58] The main challenge is Israel; it has nuclear weapons but does not publically admit to having them, and it sees those weapons as an important power equalizer given its small population and territory and threatening regional security environment. Iran and Arab states pushing for a WMDFZ insist that as part of the process, Israel must accede to the NPT, submit to IAEA safeguards, and ultimately relinquish its nuclear weapons. These countries blame Israel's lack of meaningful participation in the WMDFZ project on the United States, which they accuse of applying a double standard and

shielding Israel from the nonproliferation regime.[59] They argue that Israel's nuclear weapons are bad for regional security and stability and that the real reason Israel perceives a need for nuclear weapons is to enforce its occupation of Palestine.[60]

In contrast, Israeli leaders believe that their nuclear weapons have had a stabilizing effect on the region, encouraging negotiated settlements and discouraging all-out war. Israel has also undertaken coercive counterproliferation measures against Syria, Iraq, and Iran. These counterproliferation measures, which include bombing and assassinations, seem to have had mixed, and sometimes definitively negative, results.[61] They certainly have not endeared regional states to Israel as a partner in nonproliferation. From Israel's perspective, its nuclear weapons serve as an insurance policy for the survival of the state and a deterrent against Iranian aggression. For Israel to even participate in a WMDFZ process would require holding a conference dealing with all regional security issues and establishing a "comprehensive peace" between Israel and its regional rivals.[62] Reaching this stage would entail normalization of diplomatic relations and the growth of commercial ties between Israel and states that do not currently recognize its existence.

The Middle East WMDFZ therefore seems to be stuck in a chicken-and-egg problem. Israel argues that regional security must come before a WMDFZ is possible, while the Arab countries and Iran argue that regional security is impossible without a WMDFZ.[63] Israel's preconditions—the achievement of regional peace and its own security—are viewed as a serious and shifting obstacle to the creation of a NWFZ.[64] If Israel were to meet the preconditions set by the Arab countries and Iran, namely joining the NPT and submitting to IAEA inspections, it would resolve a significant barrier to regional agreement: Israel's outlier status as an unrecognized nuclear weapon state.[65] But it is extremely unlikely that Israel would agree to modify its security strategy without substantial changes in the regional security environment that incentivize it to do so. And it is also highly unlikely that the United States would be willing to pressure Israel to relinquish its arsenal. Another complication is Iran's potential nuclear program. Although Iran has never produced a nuclear weapon, it has operated advanced fissile material production facilities and could arguably nuclearize in the future. The recent withdrawal of the United States from the Joint Comprehensive Plan of Action, and subsequent violations of the agreement by Iran, does not bode well for establishing the kind of regional security environment Israel insists is a necessary precondition.

The prospects of a Middle East WMDFZ are primarily impeded by the lack of two "crucial criteria": a common understanding of regional history and a productive relationship with the recognized nuclear weapon states.[66] Furthermore, the existence of an adjacent nuclear state—Pakistan—raises concerns about the possibility of rapid and facilitated proliferation.[67] In other words, the prevailing strategic landscape is difficult, complex, and durable. Israel does not have the option of swapping its own nuclear deterrent for the nuclear umbrella of the United States because the protocols of any WMDFZ or NWFZ would require the United States to formally agree not to use nuclear weapons in the region. Given its recent record of not participating in NWFZ protocols, there is no guarantee that the United States would agree to formal restraint in this historically volatile region. And other regional states, especially Iran, may not trust any commitments made by the United States. As a result, some commentators describe a Middle East WMDFZ as a "utopian dream" that will require "fundamental shifts in the basic positions of both sides."[68] The prospects of a WMDFZ therefore seem to depend on the success of the peace process as a whole.

In this situation progress is sure to be slow, but it may still be possible through an incremental approach. Although the prototypical NWFZ is negotiated and endorsed by all or most of the states in a given region, it is possible for regional proponents of a NWFZ to take steps toward that goal without regional consensus. Interested Middle Eastern states could perhaps join existing NWFZs, such as the African or Central Asian zones, as a demonstration of their commitment.[69] Informal, open-ended, and ongoing consultations (without preconditions) could also identify confidence-building steps that can be taken now, including information exchange, search and rescue exercises, communications network creation, and even coordinated accession to other multilateral frameworks.[70] Willing regional actors could draw on their past experiences with cooperative monitoring to construct bilateral or small multilateral monitoring and verification systems, which could be expanded or formally endorsed through a WMDFZ at a later date.[71] Each of these steps could improve the regional security environment in ways that make forming a Middle East NWFZ more possible.

A key component of any final WMDFZ agreement will be verification. Achieving transparency even incrementally will be challenging because the densely packed states of the region may fear that they are giving up information that could be used for targeting.[72] The IAEA has expressed support for the project, and Arab states have suggested using their inspection functions. Israel seems to prefer the creation of a regional verification

scheme. The model of the Tlatelolco Treaty has been identified as a useful precedent for establishing a regional-global linked verification system that puts special focus on the states that stoke the most concern about potential violations, while taking advantage of the resources, expertise, and credibility of the IAEA system.[73] A select group—comprising government officials and/or civilian experts—could begin determining the needs for regional verification and formulating options by drawing on the "rich menu of precedents" from existing NWFZ and other arms control agreements.[74] This effort could enhance the visibility of the WMDFZ project and get a "head start on the technical elements" of any final agreement.[75]

Given their connections to the region and technical expertise, the participation and support of the United States and other nuclear weapon states like Russia may be a key enabling condition for a Middle East NWFZ. Depending on assessments of feasibility and risk, the United States may determine that promoting institutionalized mutual restraint is a better option than, for example, formally extending the US nuclear umbrella to regional states. These nuclear weapon states, or other external powers such as the United Nations Security Council, could support the creation of a WMDFZ in several ways. For instance, they could offer incentives (economic or technological) for potential members, provide satellite and other data to support verification functions, or act as a mediator or arbitrator in cases of alleged noncompliance.[76] While it will be challenging to achieve the necessary level of trust and confidence between regional and external actors to make their participation effective, a good first step could include the offer of specific and practical forms of assistance.

The Arctic NWFZ

The idea of an Arctic NWFZ has been discussed by indigenous groups, academics, and civil society groups for several decades and has recently gained momentum as attention turns to the geopolitical implications of the melting ice cap. The Inuit Circumpolar Council passed a resolution calling for the creation of an Arctic NWFZ in 1986 and endorsed the idea again in 1998, and the Canadian Pugwash Group called for the same in 2007.[77] While the feasibility of an Arctic NWFZ is widely debated, many authors suggest that the idea is worth pursuing.[78] The clearest and most persuasive arguments for an Arctic NWFZ come from Adele Buckley, an active member of Canadian Pugwash. She argues that the presence of nuclear weapons in the Arctic is a "threat to global stability" and that an Arctic NWFZ can be part of an emerging cooperative security framework for the region.[79] Because the Arctic is currently experiencing major geo-

physical, ecological, and economic change, with an attendant increase in institution building, the near future may be an opportune time to invest in the idea of an Arctic NWFZ.

The main barriers to an Arctic NWFZ are the United States and Russia. The United States opposes any declaration of its own territory as nuclear free, while maintaining a ballistic missile defense system in subarctic Alaska. Russia operates an important naval base in the Arctic, and its nuclear-armed submarines regularly patrol in Arctic waters.[80] And although neither the US nor Russia bases intercontinental ballistic missiles or nuclear-armed bombers in the region, the Arctic represents an important potential route for both delivery systems. The nonnuclear Arctic littoral states of Canada, Denmark, and Norway are all members of the North Atlantic Treaty Organization (NATO); thus, they are technically committed to a collective defense strategy relying on nuclear weapons. The Arctic is therefore a very challenging case for a NWFZ, but proponents argue that now is the time to take "preventative measures" to reduce the risk of nuclear use as new scenarios for great power competition and conflict emerge along with the open water slowly replacing the multiyear ice cap.[81] Stakeholders are also interested in reducing the risk of nuclear pollution in environmentally sensitive ice-covered areas and preserving the rights of indigenous Arctic communities.

The prospects of an Arctic NWFZ depend almost entirely on the US-Russia relationship. Writing just before the Russian annexation of the Crimean Peninsula, Buckley argues that there is "room for change" in the positions of the United States and Russia—largely because the end of the Cold War has lessened the strategic imperatives for nuclear patrols in the Arctic.[82] Whether the end of the Cold War has softened the US-Russia rivalry sufficiently is a critical question for the prospects of an Arctic NWFZ. In the past several years, events such as Russia's invasion of Crimea and meddling in US elections have increased tension between the two nuclear superpowers. However, it has been noted that Arctic politics have been somewhat insulated from international politics as a whole.[83] Still, it is unlikely that either Russia or the United States would pursue the creation of a NWFZ in the Arctic, as their existing nuclear force structures and deployments include basing and transit through the region. However, Buckley argues that a NWFZ is possible through openness to a more limited version of the prototypical NWFZ and adoption of a gradual, incremental approach led by non-state actors and the nonnuclear weapon states of the region.

The final treaty design would likely encompass only a portion of the Arctic states' territories, perhaps only the regions within the Arctic Circle, because including the entire territory of member states would require complete disarmament on the parts of the United States and Russia. This plan would make the Arctic NWFZ unique among existing NWFZs because it would be the first to encompass only parts of the territories of member states. The Arctic Circle does include the Kola Peninsula, however, the location of the Russian Northern Fleet base. If the NWFZ included these facilities, the Arctic NWFZ would be unique for a second reason: it would be the first to "require the denuclearization of the Zone" as opposed to just prohibiting future nuclear basing or deployment or dismantling nuclear production facilities.[84] This is a major obstacle, as Russia has already expressed that its support for an Arctic NWFZ is contingent on such a zone not including the base on the Kola Peninsula, which hosts the majority of Russia's nuclear-armed submarines.[85] These delivery vehicles are especially critical for nuclear deterrence strategies. A reduction in the number, or shift in the basing, of Russia's nuclear-armed submarines would almost certainly require parallel and reciprocal cuts by the United States—unlikely in the medium term. A carved-out exception for the Kola Peninsula may be needed as a condition of possibility for an Arctic NWFZ.

A typical regional NWFZ would also require the US and Russia to provide one other negative security assurances and the three other recognized nuclear weapon states to provide these assurances for all regional member states. The idea of Russia and the United States issuing negative security assurances to one another is in complete contradiction to the prevailing strategy of nuclear deterrence—and therefore extremely difficult to achieve. But negative security assurances could be limited to the regions covered by the NWFZ, namely those north of the Arctic Circle.[86] They might also need to include a promise not to attack any remaining nuclear installations in the Arctic (that may be protected in carved-out exceptions) with conventional weapons, as doing so would have environmental and social impacts similar to using nuclear weapons against a conventional facility.[87] Although this approach would not completely denuclearize the Arctic or disarm member states in an Arctic NWFZ, it might still be a valuable check on the expansion of nuclear facilities and associated risks in the region.

Proponents of an Arctic NWFZ can move forward without waiting for the United States and Russia to agree to unilateral or bilateral disarmament. The lesson taken from previous NWFZs, especially in Central Asia, is that early efforts can eventually build momentum for an agree-

ment that would not have seemed possible during initial conversations.[88] Buckley argues that "the most likely successful path" to an Arctic NWFZ could be forged by the Arctic nonnuclear states, which could form the kernel of a NWFZ through multilateral agreement.[89] The basic idea is that initial cooperation among a limited regional group of nonnuclear states can contribute to confidence building, norm creation, and a learning process that eventually extends to nuclear weapons states. At the very least, an agreement among nonnuclear states can potentially restrict the deployment (and possibly the transit) of nuclear weapons in the region. Although only Denmark formally includes a NWFZ in its stated foreign policy objectives, many of the Arctic nonnuclear states have already fulfilled the typical requirements of a NWFZ agreement.[90] Denmark could initiate discussions and build consensus, with the goal of producing a formal agreement between willing states that could model cooperation, garner support within the UN General Assembly, and serve as a focal point for international pressure on the United States and Russia. Initiating these discussions with even a limited group of Arctic states could start to work out the relationship between NATO membership and future negative security assurances.[91] Such an agreement would lock in the nuclear-weapon-free status of much of the Arctic and could be designed to expand the zone covered as new members ratify. It could even create special protocol agreements for the United States and Russia to ratify one at a time, therefore bringing them incrementally into the fold of an Arctic NWFZ. A commitment by Denmark to a nuclear-weapon-free status could signal the US that it cannot base nuclear weapons in Greenland as it did during the early Cold War.[92]

Unilateral action by regional powers could assist in this process. Canada could unilaterally declare nuclear-weapon-free status, thereby outlawing the transit of radioactive material through its internal and territorial waters. Doing so may be contentious given the dispute over the status and ownership of the Northwest Passage, but arguably these narrow and ice-choked waterways are already a challenge for submarines and "very probably a de facto nuclear-weapon free zone" already.[93] Another Law of the Sea–related challenge concerns the Central Arctic Ocean, which retains a "high seas" status in international law. Although any collection of Arctic states cannot legally outlaw the deployment or transit of all nuclear weapons through this area, individual nuclear weapon states can agree to protocols prohibiting their own nuclear weapons in the Central Arctic Ocean.[94] However, verifying the cessation of typically clandestine nuclear-armed submarine patrols would present a special, perhaps insurmountable,

verification challenge. The US and Russia would likely reject any institutionalized restriction on nuclear transit through their own national waters and/or the high seas in the Arctic.

Despite the possibility of carve-outs and the leadership of nonnuclear states, the success of a potential Arctic NWFZ ultimately depends on the United States and Russia. The geographic advantage and sunk costs of existing Arctic nuclear facilities (including ballistic missile defense) makes any restructuring of nuclear forces a challenging endeavor. If a NWFZ and its protocols were to require any substantial changes, these would have to be worked out bilaterally so that the US and Russia could maintain their overall strategic postures relative to one another.[95] If either of the nuclear superpowers were willing to take unilateral measures to achieve at least partial compliance with the envisioned NWFZ, such actions could make an important contribution to the chances of reaching a final, binding, and meaningful NWFZ agreement.

The Northeast Asia NWFZ

Another potential NWFZ would be located in Northeast Asia, where several nuclear powers converge. At different times during the Cold War, the United States, the Soviet Union, and China all considered the possible utility of a regional NWFZ in Northeast Asia, especially centered on the Korean Peninsula.[96] The idea gained new momentum starting in the 1990s, when Track II diplomatic efforts got underway in Beijing, with guidance from Argentina. But optimism about the potential of a Northeast Asia NWFZ tends to wax and wane with saber-rattling and the resumption or failure of negotiations with North Korea over the status of its nuclear program. As a result, proposals for a Northeast Asia NWFZ often begin with an argument that current nonproliferation and disarmament strategies—including extended deterrence—are not working in this region.[97]

Shaped by more than a century of conflict and distrust among major actors, the regional nuclear security environment of Northeast Asia is complex. As the newest member of the nuclear weapons club, North Korea has strong incentives—including regional security, prestige, and domestic political control—to maintain its small nuclear weapons program. Nuclear powers Russia and China share borders with North Korea, while other regional powers like South Korea, Japan, and Taiwan are under the nuclear umbrella of the United States. A Northeast Asia NWFZ could serve various purposes, including nonproliferation for Japan and South Korea; disarmament of North Korea; and restraint of the deployment and/or use of nuclear weapons by China, the United States, and other nuclear

powers. Advocates of a Northeast Asia NWFZ describe the project as "an essential circuit-breaker in the downward spiral of mistrust in Northeast Asia."[98] Engines of this dangerous cycle could include the rise of China making the extended deterrence position of the United States increasingly untenable, the risk of North Korean nuclear weapons leakage or use, and the possibility of rapid proliferation by Japan and/or South Korea.

Most advocates of a Northeast Asia NWFZ assume that the full version is impossible in the current political environment and therefore propose more limited versions. The first type of limitation concerns membership. The 3 + 3 approach would include North Korea, South Korea, and Japan as nonnuclear powers making up the NWFZ, while China, Russia, and the United States would ratify protocols providing negative security assurances to NWFZ states.[99] The 2 + 3 approach would start with South Korea and Japan as nonnuclear states, with China, Russia, and the United States providing negative security assurances. The idea is that eventually North Korea would join at a later time as a nonnuclear state, a decision that would presumably be easier to make because of increased confidence in Japan's durability and South Korea's nonnuclear status.[100]

The second type of limitation concerns the territorial or technological scope of a potential NWFZ, found in proposals that include China and/or Russia as full member states. Track II negotiations throughout the 1990s and early 2000s downgraded their consensus proposal to minimize disruption to China's nuclear force structure and strategy and to protect the Russian nuclear bastion in the Sea of Okhotsk.[101] They also called for a limited NWFZ that only controlled tactical nuclear weapons and would give member states substantial flexibility to determine the overall number of weapons deployed.[102] Newer proposals tend to set aside the notion of full regional membership with limited scope in favor of blueprints that begin with an agreement between South Korea and Japan. In particular, it is suggested that South Korea and Japan design and implement a verification mechanism similar to the one adopted by Brazil and Argentina as part of their participation in the Latin America NWFZ. In so doing, South Korea and Japan could form the basis of an agreement that would expand in scope and membership over time.[103]

Current advocates of a Northeast Asia NWFZ recognize that the proposal may seem "excessively idealistic" but note that the history of stalled and failed negotiations do not suggest a more feasible alternative.[104] Like the proposed Arctic and Middle East NWFZs, the Northeast Asia zone concept relies on an incremental, confidence-building approach that creates the conditions of possibility for a full regional NWFZ. The hope is

that "embryonic security institutions" involving information exchange, communication networks, and administrative responsibilities would eventually generate trust and investment in the idea of collective regional security.[105] Ongoing diplomatic engagement between the US and North Korea as well as Japan and South Korea improves the prospects for reducing insecurity. At least one recent author believes that there is a real opportunity for the evolutionary emergence of a tacit regional settlement that includes a NWFZ.[106]

Although North Korea is a particularly recalcitrant, isolated, and entrenched nuclear weapon state, a regional NWFZ could provide two things the Kim regime has long demanded: "equal treatment under international law" and legally binding negative security guarantees.[107] These provisions would require the United States to pledge not to station or store nuclear weapons in South Korea and Japan but would not require total US disarmament. In exchange, North Korea would relinquish its nuclear weapons and materials and submit to inspections. But even if North Korea were willing to accept these terms, the United States is likely to balk at the request for a negative security guarantee, a retraction of its nuclear umbrella from key allies, and potential restrictions on the transit of nuclear-armed vehicles.[108] Unfortunately, US participation in the negative security assurance protocol is "indispensable" to the success of a Northeast Asia NWFZ.[109]

One option that would allow the United States to maintain its nuclear umbrella over South Korea and Japan involves the extension of a nuclear umbrella by China. Essentially, North Korea would participate in the NWFZ by replacing its domestic nuclear capacity with a nuclear security guarantee from China, thereby replicating the nuclear relationship between the United States and its regional allies.[110] Although this shift would require "radical reform" to Chinese nuclear doctrine, including the abandonment of its "no first use" nuclear pledge, it could serve Chinese interests by enhancing regional stability and promoting regional nonproliferation.[111] This approach—wherein China extends its nuclear umbrella over North Korea while North Korea dismantles its nuclear weapons program—may facilitate the inclusion of North Korea in a NWFZ. However, it would also be antithetical to the overall goal of a NWFZ by legitimating the use or threat of use of nuclear weapons by China and actually expanding the scenarios wherein nuclear use by China might occur.

At this stage, the United States can support a Northeast Asia NWFZ by continuing outreach to North Korea, managing alliance relationships, and dialoging with China about expectations for a future settlement.[112] Even-

tually, the United States can offer sanctions relief and incremental security guarantees in exchange for steps toward denuclearization and participation in the verification regime. Ideally, these incremental and iterative processes will shape the regional security environment in positive directions, thereby making the issuance of a negative security assurance to North Korea more thinkable. This movement can be facilitated by China, whose leverage and influence over North Korea is a key part of most proposals for a Northeast Asia NWFZ. A nonnuclear North Korea would remain a client state of China, which will have the same incentives for peace and restraint in its sphere of influence but fewer external threats to deal with.[113]

Whether a Northeast Asia NWFZ is feasible very much depends on domestic politics in South Korea and Japan. Both states are technologically and financially capable of rapid proliferation, and each has domestic constituencies who support proliferation as a response to the challenging regional security environment. The nonnuclear status of Japan and South Korea is in large part explained by the extension of the US nuclear umbrella, or positive security guarantees. The protocols of the Northeast Asia NWFZ would require the United States to remove the nuclear umbrella. The idea is that Japan and South Korea would accept the retraction of the nuclear umbrella and commit not to proliferate in exchange for negative security assurances from the US, China, and Russia. Confidence in these assurances would have to be high to garner domestic support in Japan and South Korea and to convince key stakeholders in government and industry.[114] It has been suggested that the buildup of conventional forces by South Korea and Japan could serve many of the same deterrence functions of the US nuclear umbrella, thereby making its retraction more palatable to the South Korean and Japanese defense establishments.[115]

Even proponents of a Northeast Asia NWFZ describe its prospects in restrained terms.[116] The NWFZ project reflects a liberal internationalism that has not taken root in the security policies of Northeast Asia; "all the regional players prefer the realist approach."[117] Continued missile testing by North Korea, and mixed messages about Japanese and South Korean proliferation by the US president, complicates the security calculations of regional actors. In this environment, it would be challenging to actualize some of the components of a Northeast Asia NWFZ. The extended nuclear deterrent of the United States would have to be withdrawn without stoking abandonment anxieties on the part of its allies. And if Japan and South Korea do formally commit to nonproliferation, they would be taking a risk that China would then lose interest in pressuring North Korea to relinquish its nuclear weapons.[118] Like the proposed Middle East and

Arctic NWFZs, the possibilities for a regional NWFZ seem closely tied to more general improvements in the regional security environment.

Conclusion

NWFZs have contributed positively to the overall arms control agenda, based in part on a learning process that accompanies incremental, progressive, institutionalized mutual restraint. Even when NWFZ agreements simply formalize the strongly held preferences of member states, they provide an accountability mechanism for states that may want to pursue proliferation in the future. And they have served to reorient the strategies and policies of nuclear weapon states. The Central Asia NWFZ created a "disarmament 'pocket' in a volatile region" and a historical break with the era of Soviet nuclear testing.[119] The African NWFZ formalizes and internationalizes the nonnuclear status of former proliferators Libya and South Africa. The Tlatelolco Treaty ensured that the Western Hemisphere would not be under the nuclear umbrella of the United States. But there has been a limit to how much existing NWFZs affect the strategies, policies, and force structures of nuclear weapon states, none of which have participated as members of a NWFZ. The negative security assurance protocols have been a central feature of existing NWFZs, yet the four other recognized nuclear powers were willing to ratify negative security assurance protocols when the United States has not, and will not. This suggests that these nuclear weapon states may not have perceived the protocols as a significant commitment or one that affects their ability to use nuclear weapons. Of course, the unrecognized nuclear powers of India, Pakistan, Israel, and North Korea are not asked or obligated to ratify the protocols. Consequently, existing NWFZs have made a limited contribution to the overall arms control agenda.

The prospects for near-term, full versions of NWFZs in the regions considered are not promising. Although regional stability can be a consequence of successful NWFZs, it is also an important precondition to their establishment. The Middle East zone is impeded primarily by longstanding disagreements about the causes of insecurity in the region and deep mistrust between Israel and Iran. The Arctic zone would require concerted (and coordinated) force structure and deployment changes by the US and Russia. The prospects of a Northeast Asia zone depend on fundamental shifts in the security strategies of a number of regional actors, including four nuclear weapon states. Although full versions of these proposed NWFZs are unlikely in the near term, the goal remains a valuable one. Incremental, gradual efforts toward a NWFZ can at least keep

the arms control agenda moving in the direction of progress. What is needed are "reasonable and practical ways to short-circuit the new, self-reinforcing worldwide nuclear arms race."[120] Regional-scale efforts may be more feasible because diplomats and policy makers can tailor and reshape the NWFZ to fit a regional security dynamic "with a familiarity and commitment unmatched by globally oriented institutions."[121] In a time when the international security environment discourages pursuit of arms control agreements, the interpersonal relationships between officials can be a crucial component of success.[122]

This approach to arms control also harnesses the leadership potential of nonnuclear democracies such as South Korea, Japan, Norway, Denmark, and Canada. The executives of these states could score domestic political points via the prestige associated with principled nonproliferation, which might also have the positive effect of increasing public concern about nuclear weapons. Another piece of low-hanging fruit in terms of moving the arms control agenda forward is US ratification of the remaining protocols, especially for the Africa, South Pacific, and Central Asia NWFZs.[123] Although ratification is highly unlikely during the Trump administration, it would bolster US credibility and the norms against nuclear use, with little strategic effect on the United States.[124]

After supporting the creation of NWFZs in the early decades of the Cold War (especially for global commons), the US strategic community cooled and then hardened its opinion toward NWFZs by the end of the twentieth century.[125] Although the establishment of new NWFZs may or may not serve US strategic interests at any given time, the need exists for attentiveness to shifting regional and international conditions that may alter the incentives and costs of pursuing institutionalized mutual restraint at the regional level. In the event problems with the theory and strategy of nuclear deterrence emerge or worsen, the extension of NWFZs could support an alternative route to strategic stability. Potential modifications in the design of new NWFZs suggest they could ensure, or even enhance, nuclear deterrence while still contributing to disarmament and nonproliferation.

The overall vision remains expanding the NWFZ system to include an interlocking set of zones covering progressively larger areas of the planet. The proposals for new NWFZs in the Arctic, Middle East, and Northeast Asia will be much more challenging, however, because they would directly impact nuclear weapon states—restricting their basing, deployment, and transit of nuclear weapons—and the terms of their security alliance relationships. To make the NWFZ idea more palatable for nuclear weapon states, many of the proposed designs use modified or limited versions of the

classic NWFZ model of Tlatelolco, with carve-outs and exceptions to accommodate existing nuclear force structures and to achieve compatibility with the strategy of nuclear deterrence. It is worth asking whether this departure from the NWFZ model would be important enough to undermine the utility of potential future NWFZs by diluting their overall meaning and effect.[126] Although flexibility in the NWFZ model can increase its usefulness for nonproliferation and limited disarmament in challenging regional security environments, too much flexibility may guarantee that NWFZs will never be an effective means of reaching global zero. SSQ

Elizabeth Mendenhall

Elizabeth Mendenhall is an assistant professor in the Departments of Marine Affairs and Political Science at the University of Rhode Island. She received her PhD in international relations from Johns Hopkins University in 2017. This article received the 2020 Gen Larry D. Welch Deterrence Writing Award (Junior Division) from US Strategic Command.

Notes

1. Benjamin Zala, "How the Next Nuclear Arms Race Will Be Different from the Last One," *Bulletin of the Atomic Scientists* 75, no. 1 (January 2019): 41, https://doi.org/10.1080/00963402.2019.1555999.

2. Ulrich Kühn, "Deterrence and Its Discontents," *Bulletin of the Atomic Scientists* 74, no. 4 (July 2018): 251, https://doi.org/10.1080/00963402.2018.1486613.

3. Zala, "Next Nuclear Arms Race," 41; and John Mecklin, "Introduction: The Wasteful and Dangerous Worldwide Nuclear Modernization Craze," *Bulletin of the Atomic Scientists* 75, no. 1 (January 2019): 1–2, https://doi.org/10.1080/00963402.2019.1555973.

4. Christopher J. Fettweis, "Pessimism and Nostalgia in the Second Nuclear Age," *Strategic Studies Quarterly* 13, no. 1 (Spring 2019): 25, https://www.airuniversity.af.edu/.

5. Fettweis, 17, 32, 41n97.

6. Joseph S. Nye, "Arms Control and International Politics." *Daedalus* 120, no. 1 (Winter 1991): 145.

7. Zala, "Next Nuclear Arms Race"; Joe Goldman, "Argentina, Brazil Open to Inspections," *Bulletin of the Atomic Scientists* 47, no. 4 (May 1991): 8–10; Adam M. Scheinman, "Low-Hanging Fruit: Ratify Protocols for Nuclear-Weapon-Free Zones," *Bulletin of the Atomic Scientists*, 28 June 2018, https://thebulletin.org/; J. Dhanapala, "Disappointment in the Third World," *Bulletin of the Atomic Scientists* 46, no. 6 (August 1990): 30–31; Anne Harrington de Santana, "The Strategy of Non-Proliferation: Maintaining the Credibility of an Incredible Pledge to Disarm," *Millennium: Journal of International Studies* 40, no. 1 (September 2011): 3–19, https://doi.org/10.1177/0305829811413312; and Nina Tannenwald, "Justice and Fairness in the Nuclear Nonproliferation Regime," *Ethics & International Affairs* 27, no. 3 (2013): 299–317, https://doi.org/10.1017/S0892679413000221.

8. United Nations, The Treaty on the Non-Proliferation of Nuclear Weapons (NPT), art. VII, assigned 1 July 1968, entered into force 5 March 1970, accessed August 2020, https://www.un.org/.

9. Edmundo Fujita, *The Prevention of Geographical Proliferation of Nuclear Weapons: Nuclear-Weapon-Free Zones and Zones of Peace in the Southern Hemisphere,* UN Institute for Disarmament Research Paper no. 4 (New York: United Nations, 1989), 37.

10. United Nations General Assembly (UNGA) Resolution 3472B, Comprehensive study of the quantities of nuclear-weapons-free states in all its aspects, sec. I, General Assembly Thirtieth Session, 2437th plenary meeting, 11 December 1975, https://www.un.org/.

11. UNGA Res. 3472B, sec. II(2)(c).

12. Sebastian Brixey-Williams, "The Ban Treaty: A Big Nuclear-Weapon-Free Zone?," *Bulletin of the Atomic Scientists*, 21 June 2017.

13. Fujita, *Prevention of Geographical Proliferation*; James A. Barry, "Seabed Arms Control Issue, 1967–1971: A Superpower Symbiosis?," *International Law Studies* 61 (1980): 572–85; Jacob Darwin Hamblin, *Oceanographers and the Cold War: Disciples of Marine Science (*Seattle: University of Washington Press, 2005); and James Clay Moltz, *The Politics of Space Security: Strategic Restraint and the Pursuit of National Interests* (Stanford, CA: Stanford University Press, 2008).

14. George Downs, David Rocke, and Peter Barsoom, "Is the Good News about Compliance Good News about Cooperation?" *International Organization* 50, no. 3 (Spring 1996): 389.

15. Christopher Dunlap, "Disarmament over Deterrence: Nuclear Lessons from Latin America," *Bulletin of the Atomic Scientists*, 1 August 2018, https://thebulletin.org/.

16. Ryan A. Musto, "North Korea Might Not Denuclearize, but the US Senate Should," *Bulletin of the Atomic Scientists*, 25 October 2017, https://thebulletin.org/.

17. Fujita, *Prevention of Geographical Proliferation*, 29; and Togzhan Kassenova, "The Struggle for a Nuclear-Weapon-Free Zone in Central Asia," *Bulletin of the Atomic Scientists*, 22 December 2008, https://thebulletin.org/.

18. Dunlap, "Disarmament over Deterrence."

19. Fujita, *Prevention of Geographical Proliferation*, 18; and Sonia Fernandez-Moreno, "Nuclear-Weapon-Free-Zones: Past Lessons and Future Prospects" (presentation, Carnegie International Nonproliferation Conference, Washington, D.C., 7 April 2009), https://carnegieendowment.org/.

20. Claudia Baumgart and Harald Müller, "A Nuclear Weapons-Free Zone in the Middle East: A Pie in the Sky?," *Washington Quarterly* 28, no. 1 (December 2004): 45–58, https://doi.org/10.1162/0163660042518125; Janene Sawers, "Nuclear Weapon-Free Zones as a New Deterrent?," *Peace Brief* 24 (Washington, D.C.: United States Institute of Peace, 28 April 2010), 4; Adele Buckley, "An Arctic Nuclear-Weapon-Free Zone: Circumpolar Non-Nuclear Weapons States Must Originate Negotiations," *Michigan State International Law Review* 22, no. 1 (2013): 173, https://digitalcommons.law.msu.edu/; and Harald Müller et al., *Nuclear Weapon-Free Zone in Europe: Concept – Problems – Chances* (Frankfurt am Mein: Peace Research Institute Frankfurt, 8 May 2015), 3, https://www.bmeia.gv.at/.

21. Scott Parrish, "Nuclear-Weapon-Free Zones and Maritime Transit of Nuclear Weapons," in *The Oceans in the Nuclear Age: Legacies and Risks, a Law of the Sea Institute Project,* eds. David D. Caron and Harry N. Scheiber, expanded ed. (Leiden: Brill-Nijhoff, 2014), 338.

22. Ernie Regehr, "A Nuclear-Weapon-Free Zone and Cooperative Security in the Arctic," Disarming Arctic Security briefing paper, The Simons Foundation, 14 October 2014, 6.

23. Fujita, *Prevention of Geographical Proliferation*, 5.

24. F. H. Hammad and Adel M. Ali, "Principles of Establishing a Middle East Weapons of Mass Destruction Free Zone (MEWMDFZ) Monitoring and Verification System," IAEA-SM-367/9/07 (Vienna, Austria, 2001); and Michael Hamel-Green, *Implementing a Japanese-Korean Nuclear Weapon Free Zone: Precedents, Legal Forms, Governance, Scope and Domain, Verification and Compliance, and Regional Benefits*, Austral Special Report 10-02A (Melbourne, Australia: Nautilus Institute, September 2010), http://nautilus.org/.

25. Xia Liping, "Nuclear-Weapon-Free Zones: Lessons for Nonproliferation in Northeast Asia," *The Nonproliferation Review* 6, no. 4 (December 1999): 83–92, https://doi.org/10.1080/10736709908436781.

26. Hamel-Green, "Japanese-Korean Nuclear Weapon Free Zone," 9.

27. Parrish, "Maritime Transit of Nuclear Weapons," 340.

28. Mansour Salsabili, "Fixing a Process in Jeopardy," *Bulletin of the Atomic Scientists*, 21 June 2013, https://thebulletin.org/; and Jennifer Knox, "Haves, Have-Nots, and Need-Nots: The Nuclear Ban Exposes Hidden Fault Lines," *Bulletin of the Atomic Scientists*, 3 July 2017, https://thebulletin.org/.

29. Fujita, *Prevention of Geographical Proliferation*; John E. Endicott, "Limited Nuclear-Weapon-Free Zones: The Time Has Come," *Korean Journal of Defense Analysis* 20, no. 1 (March 2008): 20, https://doi.org/10.1080/10163270802006305; and Heinz Gartner, "A Neutral State's Perspective on the Ban—and a Compromise," *Bulletin of the Atomic Scientists*, 15 August 2017, https://thebulletin.org/.

30. Daniel Deudney, "Hegemony, Nuclear Weapons, and Liberal Hegemony," in *Power, Order, and Change in World Politics*, ed. G. John Ikenberry (Cambridge: Cambridge University Press, 2014), 220.

31. Gerald Segal, "Nuclear Strategy: The Geography of Stability," *Political Geography Quarterly* 5, no. 4, suppl. 1 (October 1986): S42, https://doi.org/10.1016/0260-9827(86)90056-X.

32. Segal, S42.

33. Regehr, "Cooperative Security in the Arctic"; and Daniel Deudney, "Geopolitics as Theory: Historical Security Materialism," *European Journal of International Relations* 6, no. 1 (2000): 77–107, http://www.unice.fr/.

34. Fujita, *Prevention of Geographical Proliferation*, 39.

35. Kassenova, "Struggle for a Nuclear-Weapon-Free Zone"; William Potter, "Nuclear-Weapon-Free-Zones: Past Lessons and Future Prospects" (panel moderator, Carnegie International Nonproliferation Conference, Washington, D.C., 7 April 2009), https://carnegieendowment.org/; Leon V. Sigal, "Political Prospects for a Nuclear-Weapons-Free Zone in Northeast Asia: Political Prospects for an NWFZ," *Pacific Focus* 26, no. 1 (April 2011): 22–36, https://doi.org/10.1111/j.1976-5118.2011.01054.x.

36. Brixey-Williams, "Ban Treaty?"

37. Sawers, "Zones as a New Deterrent?," 2.

38. Fettweis, "Pessimism and Nostalgia"; and Daniel Deudney, "The Great Debate: The Nuclear-Political Question and World Order," in *The Oxford Handbook of International Security*, eds. Alexandra Gheciu and William C. Wohlforth (Oxford: Oxford University Press, 2018), 333–49, https://doi.org/10.1093/oxfordhb/9780198777854.013.22.

39. Fujita, *Prevention of Geographical Proliferation*.

40. Sawers, "Zones as a New Deterrent?," 3.

41. Nuclear Threat Initiative (NTI), "Southeast Asian Nuclear-Weapon-Free-Zone (SENWFZ) Treaty (Bangkok Treaty)," 30 March 2019, https://www.nti.org/.

42. NTI, "(Bangkok Treaty)."

43. Musto, "North Korea Might Not Denuclearize."

44. Fujita, *Prevention of Geographical Proliferation*, 39.

45. Gartner, "Neutral State's Perspective on the Ban."

46. Knox, "Haves, Have-Nots, and Need-Nots."

47. Fujita, *Prevention of Geographical Proliferation*, 38.

48. K. Hakapää and E. J. Molenaar, "Innocent Passage—Past and Present," *Marine Policy* 23, no. 2 (March 1999): 131–45, https://doi.org/10.1016/S0308-597X(98)00032-3; and Tullio Treves, "Navigation of Ships with Nuclear Cargoes: Dialogue between Flag and Coastal States as a Method for Managing the Dispute," in Caron and Scheiber, *Oceans in the Nuclear Age*, 217–38.

49. Fujita, *Prevention of Geographical Proliferation*, 27.

50. Endicott, "Limited Nuclear-Weapon-Free Zones," 21–22.

51. Fernandez-Moreno, "Nuclear-Weapon-Free-Zones."

52. This article does not evaluate proposals for reviving the Indian Ocean "zone of peace" idea despite some mild and intermittent support by regional states and scholarly literature. Even the apparent height of India's support for the zone of peace concept—during its leadership in the

nonaligned movement—has been called into question such that there is a low level of optimism about its current prospects. For information on this topic, see Yogesh Joshi, "Whither Non-Alignment? Indian Ocean Zone of Peace and New Delhi's Selective Alignment with Great Powers during the Cold War, 1964–1979," *Diplomacy & Statecraft* 30, no. 1 (January 2019): 26–49, https://doi.org/10.1080/09592296.2019.1557414; and A. V. Kupriianov, "The Indian Ocean as a Zone of Peace: An Outdated Concept or Format for the Future?," *Outlines of Global Transformations: Politics, Economics*, *Law* 12, no. 1 (April 2019): 204–19, https://doi.org/10.23932/2542-0240-2019-12-1-204-219.

53. Fernandez-Moreno, "Nuclear-Weapon-Free-Zones."

54. Parrish, "Maritime Transit of Nuclear Weapons," 340.

55. Joseph F. Pilat and Nathan E. Busch, "WMD Monitoring and Verification Regimes: Lessons from Iraq," *Contemporary Security Policy* 32, no. 2 (August 2011): 401–31, https://doi.org/10.1080/13523260.2011.590363.

56. Baumgart and Müller, "Nuclear Weapons-Free Zone in the Middle East?," 57.

57. Sameh Shoukry, "Nuclear-Weapon-Free-Zones: Past Lessons and Future Prospects," presented at the Carnegie International Nonproliferation Conference, Washington, D.C., 7 April 2009; and Mansour Salsabili, "Time for a Nuclear Pause Agreement?," *Bulletin of the Atomic Scientists*, 7 August 2018.

58. Zachary Davis and Warren H. Donnelly, "A Nuclear Weapons Free Zone in the Middle East: Background and Issues," CRS Issue Brief (Washington, D.C.: Congressional Research Service, 1992), 3.

59. Tannenwald, "Justice and Fairness," 312.

60. Baumgart and Müller, "Nuclear Weapons-Free Zone in the Middle East?"; and Gawdat Bahgat, "A Mideast Nuclear-Weapons-Free Zone: Pie in the Sky," *Middle East Policy* 22, no. 3 (September 2015): 27–35, https://doi.org/10.1111/mepo.12140.

61. Mālfrid Braut-Hegghammer, "Revisiting Osirak: Preventive Attacks and Nuclear Proliferation Risks," *International Security* 36, no. 1 (Summer 2011): 101–32, https://doi.org/10.1162/ISEC_a_00046.

62. Bahgat, "Mideast Nuclear-Weapons-Free Zone," 33.

63. Marc Finaud, "A Third Way toward a WMD-Free Middle East," *Bulletin of the Atomic Scientists*, 18 December 2013, https://thebulletin.org/.

64. Shoukry, "Nuclear-Weapon-Free-Zones."

65. Salsabili, "Time for a Nuclear Pause Agreement?"

66. Bahgat, "Mideast Nuclear-Weapons-Free Zone," 29.

67. Anupam Srivastava and Seema Gahlaut, "Curbing Proliferation from Emerging Suppliers: Export Controls in India and Pakistan," *Arms Control Today* 33, no 7 (September 2003): 12–16, https://www.jstor.org/; and David Albright and Corey Hinderstein, "Unraveling the A. Q. Khan and Future Proliferation Networks," *Washington Quarterly* 28, no. 2 (Spring 2005): 109–28, https://doi.org/10.1162/0163660053295176.

68. Baumgart and Müller, "Nuclear Weapons-Free Zone in the Middle East?," 45, 55.

69. Gartner, "Neutral State's Perspective on the Ban."

70. Finaud, "Third Way toward a WMD-Free Middle East."

71. Hammad and Ali, "Monitoring and Verification System."

72. Baumgart and Müller, "Nuclear Weapons-Free Zone in the Middle East?," 51.

73. Hammad and Ali, "Monitoring and Verification System."

74. Baumgart and Müller, "Nuclear Weapons-Free Zone in the Middle East?," 53.

75. Baumgart and Müller, 57.

76. Baumgart and Müller, 55; and Sawers, "Zones as a New Deterrent?," 3.

77. Regehr, "Cooperative Security in the Arctic," 2.

78. Regehr, 7.

79. Buckley, "An Arctic Nuclear-Weapon-Free Zone," 167.

80. Michael Wallace and Steven Staples, *Ridding the Arctic of Nuclear Weapons: A Task Long Overdue* (Ottawa, Ont.: Rideau Institute, 2011), https://rideauinstitute.ca/.

81. Adele Buckley, "Toward a Nuclear-Free Arctic," *SGI Quarterly*, April 2009, https://pugwashgroup.ca/.

82. Buckley, "Arctic Nuclear-Weapon-Free Zone," 184.

83. Michael Byers, "Crises and International Cooperation: An Arctic Case Study," *International Relations* 31, no. 4 (December 2017): 375–402, https://doi.org/10.1177/0047117817735680; and Jon Rahbek-Clemmensen, "The Ukraine Crisis Moves North. Is Arctic Conflict Spill-over Driven by Material Interests?," *Polar Record* 53, no. 1 (January 2017): 1–15, https://doi.org/10.1017/S0032247416000735.

84. Vladimir Rybachenkov, "An Arctic Nuclear Weapons Free Zone—a View from Russia" (text of remarks, seminar, "Arctic Nuclear Weapons Free Zone—Challenges and Opportunities," Danish Institute of International Studies, Copenhagen, 26 September 2012, 4–5, https://www.armscontrol.ru/.

85. Alexander Sergunin and Valery Konyshev, "Russia in Search of Its Arctic Strategy: Between Hard and Soft Power?," *Polar Journal* 4, no. 1 (January 2014): 81, https://doi.org/10.1080/2154896X.2014.913930.

86. Regehr, "Cooperative Security in the Arctic," 9.

87. Tazrian Alam, "An Arctic Nuclear Weapons Free Zone in the Northwest Passage," NATO Association of Canada, 7 October 2014, http://natoassociation.ca//.

88. Buckley, "Toward a Nuclear-Free Arctic."

89. Buckley, "Arctic Nuclear-Weapon-Free Zone," 190.

90. Adele Buckley, "Arctic Nuclear Weapon Free Zone" (presentation, Canadian Pugwash, University of Toronto, 30 March 2014).

91. Regehr, "Cooperative Security in the Arctic," 15.

92. Nikolaj Petersen, "The H. C. Hansen Paper and Nuclear Weapons in Greenland," *Scandinavian Journal of History* 23, no. 1–2 (June 1998): 21–44, https://doi.org/10.1080/03468759850116007.

93. Buckley, "Toward a Nuclear-Free Arctic."

94. Regehr, "Cooperative Security in the Arctic," 10.

95. Regehr, 12.

96. Peter Hayes and Michael Hamel-Green, "The Path Not Taken, The Way Still Open: Denuclearizing the Korean Peninsula and Northeast Asia," *Asia-Pacific Journal* 7, no. 50 (December 2009).

97. Sigal, "Political Prospects"; and Joshua Shifrinson, "Security in Northeast Asia: Structuring a Settlement," *Strategic Studies Quarterly* (Summer 2019): 23–47, https://www.airuniversity.af.edu/.

98. Tianjiao Jiang, "From Nuclear Hedging to Korea-Japan Nuclear Weapons Free Zone: Japan's Nuclear Options," *Copenhagen Journal of Asian Studies* 34, no. 1 (2016): 87, https://rauli.cbs.dk/.

99. Hiromichi Umebayashi and Tatsujiro Suzuki, "A Northeast Asia Nuclear Weapon-Free Zone at the New Stage of the Development in Global Nuclear Disarmament and Nonproliferation," *Korea Observer* 47, no. 4 (Winter 2016): 963–76, http://www.iks.or.kr/.

100. Jiang, "From Nuclear Hedging"; and Chung-in Moon, "Time May Be Right for a Northeast Asia Nuclear-Weapon-Free Zone," *Bulletin of the Atomic Scientists*, 25 August 2016, https://the bulletin.org/.

101. Endicott, "Limited Nuclear-Weapon-Free Zones."

102. Endicott, 15.

103. Hamel-Green, "Japanese-Korean Nuclear Weapon Free Zone."

104. Moon, "Time May Be Right."

105. Endicott, "Limited Nuclear-Weapon-Free Zones."

106. Shifrinson, "Security in Northeast Asia."

107. Moon, "Time May Be Right"; and Hamel-Green, "Japanese-Korean Nuclear Weapon Free Zone."

108. Hayes and Hamel-Green, "Path Not Taken."

109. Jiang, "From Nuclear Hedging," 85.

110. Philip Bobbitt, "Only China Can Solve the North Korea Problem—by Inviting It to Come underneath Its Own Nuclear Umbrella," *UnHerd,* 19 February 2018, https://unherd.com/.

111. Christian Conroy, "China's Nuclear Parasol," *The Diplomat*, 26 January 2014, https://the diplomat.com/.

112. Shifrinson, "Security in Northeast Asia," 34–36.

113. Shifrinson, 33.

114. Jiang, "From Nuclear Hedging," 96; and Sigal, "Political Prospects," 29.

115. Eric Gomez, "Revisiting the Value of the U.S. Nuclear Umbrella in East Asia," *War on the Rocks*, 6 March 2018, https://warontherocks.com/.

116. Jiang, "From Nuclear Hedging"; and Sigal, "Political Prospects."

117. Jiang, 99.

118. Sigal, "Political Prospects," 28.

119. Kassenova, "Struggle for a Nuclear-Weapon-Free Zone."

120. Mecklin, "Introduction," 1.

121. Endicott, "Limited Nuclear-Weapon-Free Zones," 21.

122. Zala, "Next Nuclear Arms Race," 42.

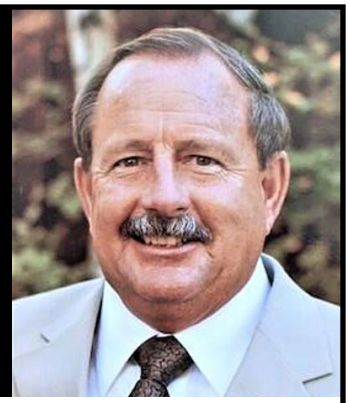
123. Sawers, "Zones as a New Deterrent?," 4; Adam M. Scheinman, "Low-Hanging Fruit: Ratify Protocols for Nuclear-Weapon-Free Zones." *Bulletin of the Atomic Scientists*, 28 June 2018, https://thebulletin.org/.

124. Scheinman, "Low-Hanging Fruit"; and Musto, "North Korea Might Not Denuclearize."

125. Fujita, *Prevention of Geographical Proliferation*, 36.

126. Regehr, "Cooperative Security in the Arctic."

In Memoriam
Stephen Chiabotti
1950–2020



We pay tribute to author, scholar, and strategist Dr. Stephen Chiabotti, Colonel, USAF, retired. Steve was a founding member of the *Strategic Studies Quarterly* (*SSQ*) Contributing Editors group and the longest-serving member of the group to date—11 years. Every month over those years, he provided insightful evaluations of *SSQ* submissions, adding greatly to the quality of *SSQ* content. Steve also published with *SSQ* on several occasions, including his most recent article in our Conventional Deterrence special edition, "Clausewitz as Counterpuncher" (https://www.airuniversity.af.edu/).

Colonel Chiabotti was a Distinguished Graduate of the United States Air Force Academy (1972), earning a degree in physics. He later earned a doctorate in history from Duke University. Colonel Chiabotti served 30 years active duty in the Air Force and became commandant of the USAF School of Advanced Air and Space Studies (SAASS)—an institution dedicated to educating strategists. Afterward, he taught strategy at SAASS as a civilian faculty member for 16 years.

Air University and *Strategic Studies Quarterly* have lost an intellectual giant and dedicated servant whose influence will be sorely missed. Steve's legacy will continue in the lives of his students, colleagues, and friends for generations.

Mission Statement

Strategic Studies Quarterly (*SSQ*) is the strategic journal of the Department of the Air Force, fostering intellectual enrichment for national and international security professionals. *SSQ* provides a forum for critically examining, informing, and debating national and international security matters. Contributions to *SSQ* will explore strategic issues of current and continuing interest to the larger defense community, and our international partners.

Disclaimer

The views and opinions expressed or implied in *SSQ* are those of the authors and should not be construed as carrying the official sanction of the Department of the Air Force, the Department of Defense, Air Education and Training Command, Air University, or other agencies or departments of the US government.

Comments

We encourage you to e-mail your comments, suggestions, or address change to
StrategicStudiesQuarterly@au.af.edu

Article Submission

The *SSQ* considers scholarly articles between 5,000 and 15,000 words from US and international authors. Please send your submission in Microsoft Word format via e-mail to

StrategicStudiesQuarterly@au.af.edu

Strategic Studies Quarterly (SSQ)
600 Chennault Circle, Building 1405
Maxwell AFB, AL 36112–6026
Tel (334) 953–7311

View and Subscribe to *Strategic Studies Quarterly* at
https://www.airuniversity.af.edu/SSQ/

Free Electronic Subscription

Like *SSQ* on Facebook at https://www.facebook.com/StrategicStudiesQuarterly

Strategic Studies Quarterly (*SSQ*) (ISSN 1936-1815) is published by Air University Press, Maxwell AFB, AL.