SSQ
STRATEGIC STUDIES QUARTERLY
SUMMER 2020 VOL. 14, NO. 2

SSQ Strategic Studies Quarterly

https://www.af.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

SUMMER 2020 VOL. 14, NO. 2

POLICY FORUM

FEATURE

PERSPECTIVES

BOOK REVIEWS

IN MEMORIAM

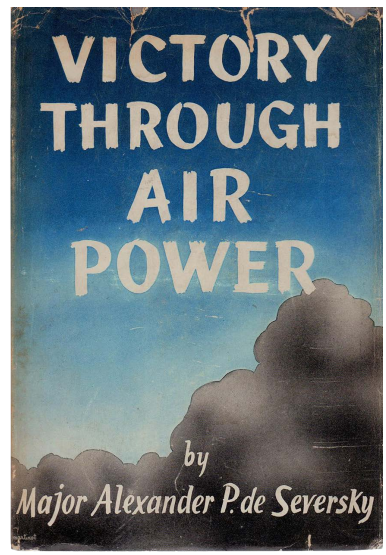
Victory through Space Power

Although the US Space Force was officially established on 20 December 2019, its future is hardly set. Nominally autonomous in its current form, it is still subordinate to the Department of the Air Force. With a minuscule personnel structure and budget relative to its peers on the DOD organizational chart, the Space Force would be rather easy to excise completely should future administrations wish to do so.

The last word is far from written, though one of the most recent is from Lt General Stephen Kwast, USAF, retired. Writing in *Politico* on 17 January 2020, he asserts that the US is "losing the race to space with China" because the Air Force is "trapped in an industrial-age mindset" that prevents it from effectively applying innovative solutions to the entirely unique domain of outer space, or even building the right kind of equipment.[1] He contends that the effectiveness of the Space Force—indeed its survival—depends on complete separation from the US Air Force.

Kwast is not anti-Air Force. He is simply making an argument that has echoed throughout America's military history: it is the same one proclaimed in the need for an independent Air Force. That is, to maximize effectiveness, any operational area understood to be a unique war-fighting domain must have a fully autonomous service dedicated to it. Until it does, its value to the nation will be stunted.

The first words on air-domain-based service separation from ground forces were voiced during World War I, culminating in 1919 with the inauguration of Britain's independent Royal Air Force. The essential arguments became widely known through the works of Italian theorist Giulio Douhet and America's William "Billy" Mitchell in the 1920s. But as far as American airpower goes, the most comprehensive assessment was published in January of 1942, just weeks after the bombing of Pearl Harbor. *Victory Through Air Power*, by Army Air Forces major Alexander de Seversky, was an instant best seller.[2] It was immediately picked up by the popular Book-of-the-Month Club and sent out to its 400,000 subscribers. The work was considered so vital that Walt



Disney financed and released a Technicolor feature film of the same name starring De Seversky in 1943. The public—previously isolationist, but now thrust into a global conflict and anxious regarding America's prospects—had quickly acquired a voracious appetite for treatises on military strategy and capabilities.

The situation appeared dire. *Victory Through Air Power* assured its readers that America, trailing its enemies in both armaments and military technology but possessing the indomitable character of its population, would surely prevail *so long as it embraced airpower* as its guiding principle. Unlike today, where no place on the earth is more than minutes away from a military strike, America was then protected by vast oceans. It would have the time to build a military juggernaut that would take the fight to Europe and Asia only if led by armadas of aircraft.

Willpower, technology, and industrial might were not enough to ensure victory, however. In the preface to the Book-of-the-Month Club edition, the editors point out that "Major de Seversky's main conclusion is that the wisest air strategy cannot be expected from managers of the war who are primarily navy-minded and army-minded; they must be primarily air-minded; and this calls for a radical change in the very organization of war-management: specifically, for an Air Force wholly separate from the Army and Navy, but working closely with them as a third independent arm" (xii–xiii).[3] To be sure, they stated, "the war cannot be won by either side unless a total supremacy in the air is achieved, and this idea itself must achieve total supremacy in the minds of our war-managers or we shall never do all that is needed, and all that we can do, to bring the war to a successful close for our side" (xiv). This argument resonates today with General Kwast's position on space power.

De Seversky begins his treatise by asserting that the US had fallen behind the other great powers in military aeronautics and preparation, "whether measured by the yardstick of military performance [or] by the yardstick of planned strategy, tactics, and organization" (4). If the status of space power is not perfectly analogous to that of De Seversky's assessment of airpower, it is certainly approaching that precipice if Kwast is correct in apprising the rapid rise of China and the resurgence of Russia. For De Seversky, a slumbering America was jolted from its insular malaise. He cautions, "As dramatically as though it had been staged by providence as a warning to the American people, our own entry into the war, on December 7, 1941, was signalized by a humiliating defeat through enemy air power. . . . The terrific danger was exposed for all to see" (4). Due to range limitations of enemy aircraft, the US population and industrial base were

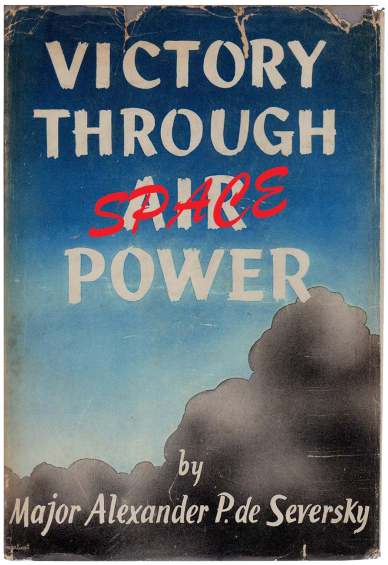
not threatened directly and had time to recover. Such is simply not the case today. Debilitating loss of space capabilities from a surprise attack; direct assaults with ballistic and cruise missiles; cyber strikes; or, in the near future, space-based weaponry could be anticipated within minutes. At the very least, America's military is currently ill-equipped to fight effectively beyond its borders without robust space support.

Given the situation in 1942, more armaments and better technology could overwhelm the enemy in time, but that was only part of De Seversky's victory equation for America. He notes, "Whether it utilizes these potentialities, or once more allows itself to trail along imitatively, depends on how quickly and thoroughly we cleanse our air power from the accretions of conservatism, timidity, and astigmatic leadership" (5). What follows is an argument for an independent Space Force directly lifted from and supported by extensive direct quotes from *Victory Through Air Power*, slightly altered by me [*substituting space-relevant words*] to accentuate common themes from then to now. In doing so, like De Seversky for airpower, "I hope to contribute toward that emancipation of American [*space*] power. . . . Above all I hope to convey the sense of [*space*] power as a dynamic, expanding force, the growth of which must be anticipated by courageous minds. It happens to be a force that eludes static, orthodox minds no matter how brilliant they may be" (5).

Modifying De Seversky for Space 2020

De Seversky's observations about resistance to change remain pervasive. From his perspective, "throughout history, new weapons have imposed new tactical principles upon the science of war making. Throughout history, too, some nations have been quicker than others to recognize and apply those principles. In this, as in every other field of human endeavor, new ideas have had to contend with inertias of habit and enshrined precedent. The very military leaders who should be most keenly alert to detect and utilize new instrumentalities are often muscle-bound by traditional thinking. Only that can account for the stubborn reluctance of our own higher authorities . . . to project their thinking into the future, beyond current events" (121).

Today, our most pressing need is for a "separate and independent [*Space Force*], organized as an equal partner in the great triumvirate of our land, sea, [*and air*] services. . . . A new weapon functioning in a different element calls for a specialized organization to meet its needs efficiently and expeditiously" (254). But this view is far from universal. It had been opposed at the highest levels of the Department of the Air Force until recently, with



the secretary and service chief voicing opposition in testimony to Congress as late as 2018.[4] Air Force arguments against separating space lacked historical acumen and were precisely the same as those of the Army against separating the Air Force in 1947. According to De Seversky, "There is a kind of desperation in the illogic of some of the arguments against freeing [*space*] power from its present subordination to the older services. . . . The simple truth is that a separate [*Space Force*] is not a wonder-working device to guarantee automatic military ascendancy, but *merely the minimal precondition for successful modern warfare*" (emphasis added) (255).

For those who argue that a separate Space Force would violate the principle of unity of command, De Seversky provides a unique rejoinder: "An overall command of all our forces would be a futile anomaly unless [*space*] had at least equal representation and authority in its councils" (257). This is a clear acknowledgment that *specialization* within a broader organization *enhances* cooperation and efficiency—it does not detract from it. De Seversky adds an additional factor when he offers the example of a painter who uses primary colors to achieve every possible combination of hue and tint. Innovation through creativity, critical to prevailing in an ever-changing battlespace, is lost without differentiation. Consequently, "an overall command would be stymied if it did not have the primary military services at its disposal" (257).

Setting up a separate and equal Department of the Space Force can be done easily and quickly. This is "a necessary and preliminary step. A High Command can be set up almost overnight when decided upon; it requires only the selection of the proper persons and their investment with the proper authority. But the full development of an organization to provide us with [*space*] power will take some time and should be started without delay" (258). In this passage, De Seversky presages the current status of the US Space Force. Where Kwast differs is in his concern that the Space Force as currently constituted is too small and subordinated to the Department of the Air Force. His fear is that this will be the *final* reorganization plan. As a step in the right direction, it is laudable. As an end, it is lamentable.

"Another set of objections rests on a misunderstanding of the scope of a separate [*Space Force*]. It assumes that the new Department would arbitrarily gather in and control anything that [*operates to, in, and from space*]. Those who proceed from this false premise, those who advance such arguments, are like Don Quixotes fighting windmills. [For example,] no matter how military aviation may be organized, neither the sea nor the land service can or should be stripped of airplanes logically and tactically a part of its operations. Just as the Navy has Marines and other adjuncts which, in a literal sense, are land troops—just as the Army has transport boats and other sea-going auxiliaries—so both the Army and the Navy would continue to possess airplanes for their own specific purposes" (258). The same argument applies to space. The Army, Navy, and Air Force will retain such space assets necessary to their primary purpose—to command the land, sea, and air, respectively. If command cannot be achieved, then to deny command of those domains to adversaries is critical. Note, too, whereas the Marine Corps (as a functional service) is appropriately situated as an *adjunct* to sea power, De Seversky insists the Air Force (as a *domain*-based service) will be suboptimally utilized as a mere adjunct to land power. Likewise, the Space Force should not be a mere adjunct to airpower.

"The distinction that needs to be made is between *integration* and *co-ordination*. The U.S. Marines, for instance, are integrated with the Navy. . . . The Army coastal batteries, though intended to help the fleet in warding off assaults on our shores, are merely co-ordinated with the Navy" (emphasis in original) (258). Further, "in like manner, certain [*space power*] auxiliaries should reasonably be integrated with the land, sea, [*and air*] forces, as part and parcel of those military branches. That, however, does not affect the broader question of a self-sufficient [*Space Force*], developed to conduct [*space*] warfare against an enemy [*space*] power, and also capable of co-ordinated effort with the surface forces" (258–59). As is evident in this supremely cogent point, it is perfectly reasonable that a force dedicated to *securing* command of its unique domain should have the ability to *challenge* command in adjacent domains. A ground force does not want opposing air forces flying freely over it and so rightfully acquires surface-to-air missiles. A Space Force would surely have need to contest the ability to project power into it from terrestrial domains and must have the martial means—weapons—necessary to contest access there. To do nothing *from* space to oppose an attack *into* space from a ground-based antisatellite weapon (as Russia, China, and India have all demonstrated) is in this construction ludicrous.

It follows that command of a domain is only truly possible from military assets operating *in* the domain. We can envision an enemy being able to scour the skies from the ground or sea, for example, but that does not guarantee it can operate there if I also have that ability. In this case the air domain would be mutually contested, but not commanded. When command of the domain—either generally or locally, indefinite or temporary—is achieved, only then can effects *from* the domain be delivered. If you cannot get *into* the domain, you cannot *use* the domain. In times of war or conflict, a military force must be prepared to fight its way into a domain to use it. This is the essence of the "boots on the ground" argument in a land campaign. A similar argument is critical for the sea and air domain; let us call it the "oars in the water" or the "wings in the air" principle, respectively. If space power is as vital to America's defense as General Kwast asserts, it cannot be maintained much less achieved if it is lost unless and until the US Space Force is allowed to develop the best means to fight *in* and *from* space.

This brings to fore issues of operational command that have been trotted out in opposition to a separate Space Force, again obviated by the experience of airpower. Going back to De Seversky, "the Army may retain minor aviation adjuncts—certain aerial transport facilities and local reconnaissance auxiliaries, for instance. But in operations where Army and Air Force are components of a larger team, there is no more reason for the infantry to own and control the aviation than there is for aviation to own and control the infantry and tank divisions. Whether the Army or the Air Force will command a given operation will depend on the nature of the job to be done. The success of the operation will depend upon how well each of these components has been developed in its own sphere, and how well they have been trained together. We can no more hope for maximum exploitation of [*space*] possibilities—in readiness for such operations—if development of the [*Space Force*] is subordinate to [*air*] minds than we could expect the best development of ground forces by subordinating them to naval minds" (259–60).

The nature of the job to be done is the central issue of the preceding paragraph. Airmen are acutely aware that their mantra "An Airman in charge of the air campaign" enhances coordination and is vital to the success of any large-scale joint operation. Further, "the guiding mind of a High Command or Commander in Chief must be taken for granted at all times. That applies to the Army and Navy [*and Air Force*] no less than to a separate [*Space Force*]. No military service in wartime acts on its own whims. Their coordination, however, does not require a merger of their internal organizations. Each derives its strength from its specialized per-

sonnel, its full sense of responsibility, its ability to extract the last ounce of effectiveness without interference from an alien mentality" (260).

"The task of the guiding intelligence is, in fact, enormously facilitated when it can deal with a self-sufficient [*space*] arm. There are then no leading strings held tight by the older services to hamper coordination. [*Space*] power would be available for use without bureaucratic impediments. The very principle of effective coordination calls for true independence and equality on the part of the collaborating elements. There can be no coordination between [an] Air Force and a [*Space Force*] that is subordinate to the [*Air Force*]; in that case there can only be orders based on [*aviation*] ideas and blind obedience by [*space professionals*]" (261).

"[*Space*], it cannot be too often repeated, is a separate element, distinct from land and sea [*and air*]—an element with its own [*spatial*] relations, its own laws and problems. It is a continuous and uninterrupted element enveloping the entire globe; strategically speaking every political division and every differentiation between [*space*]-over-land, [*space*]-over-water, [*and space-over-air*] is artificial and meaningless. [It] calls for a continuous and undivided [*Space*] Command and fully uniform [*space*] equipment, co-operating where necessary with the Army, Navy, [*and Air Force*] expecting their co-operation where needed" (263).

"On examination it becomes clear that [these are] the official views of the existing military Departments. These military minds are confusing two categories of unity. No one denies that [the] Army, Navy, [*and Air Force*] need unity of command in their respective spheres. This includes authority over their various auxiliaries and authority over units of other services working with them for specific tasks. If the sole function of [*space forces*] were to serve as the adjunct to the surface forces, the [views] would be entirely correct. Actually, it is only true so far as it goes—and unfortunately it leaves out of consideration entirely *the most vital and decisive* [*spacefaring*] *which operates alone* in its own environment" (emphasis added) (263–64). Just as with land, sea, and airpower, there are critical missions envisioned for space that require *it* to be the supported force. The entrenched "military minds do not grasp this idea themselves—that coordinated action with other services is a secondary function of [*space*] power. The primary function of [*military space*] power is to destroy the hostile [*space*] forces, to strike an enemy directly across long distances—in brief to take and hold [*space*]. That is as distinct an undertaking as the conquest and control of the seas by a navy used to be" (264). If all [*space capabilities*] were possessed "by the ground, sea, [*and air*] forces [as mere] adjuncts, [we] should have no genuine [*space*] power at all" (264).

In an eerily parallel analogy, De Seversky was significantly concerned with the reorganization of the land and air forces that had recently occurred, much as Kwast is today skeptical of the less than fully independent Space Force. De Seversky states, "The War Department in the late spring of 1941 presented the country with a minor reform dressed up to look like a major reorganization. It announced the 'unification of its air activities in a new unit to be known as The Army Air Forces.' In informing Congress of the plan, the Secretary of War emphasized the 'autonomy' of the new setup and its 'unity of command'" (266).

He continues: "This verbal garnishing was most unfortunate. Insofar as it created the illusion the Army had made a 'compromise' on the issue of self-administered air power, the public was being misled. Actually the reorganization merely revamped one of the Army's subordinate sections, while keeping it as subordinate as ever. The 'unification' announced did not even touch the basic issues raised by advocates of an Air Department. No one can object to necessary internal reform of the Army's aeronautical structure. But there is every reason to object strenuously to an attempt to palm off a departmental reform, such as frequently occurs inside any service, as a substitute for a separate Air Force" (266).

"True, our General Staff now provides equal representation for air officers; the fifty-fifty arrangement ought to result in better air-ground coordination," De Seversky adds. "The air members, having had Army training, understand ground problems and can be most useful in formulating co-operative efforts. The ground members, however, having had no . . . aviation training, can make no real contribution to purely air strategy. On the contrary, they will act as a brake on real aeronautical thinking and a constant source of interference. The Air Forces will be even more firmly tied into the ground strategy than before" (278–79).

The similarity to the Space Force is striking. "The common denominator of all the objections to an independent [*Space*] Force is that the United States has built [*space*] forces of a sort under the aegis of the [other] services and can therefore improve them without limit under the same monitorship. . . . The answer is that *at present we have no* [*space*] *power at all* [emphasis in original]. We have a miscellany of [*space assets*], good, bad, and indifferent, but no [*space*] power in the sense we have defined air power in these pages [because] those [assets] will not reflect a unified [*space*] strategy to be used by a unified [*space*] command" (279).

The full emancipation of air and space forces is only the latest iteration of a long debate. That is, "the inquisitive might go back to the debates on a separate and independent Navy Department in the last decade of the

eighteenth century, before the Navy was given autonomous status in 1798. It is an enlightening experience. You will find that the War Department used substantially the same arguments for holding onto the naval forces that are now being used against emancipation of [*space*] power; and you will find, too, that the spokesmen for genuine sea power used about the same reasoning in support of their thesis that is being advanced today by advocates of a separate [*Space*] Department" (282).

De Seversky reflects, "Reading the Congressional Record of April 25, 1798, I thought I was listening to typical . . . objections to the freeing of air [*and space*] power. Congressman Livingston, for instance, grew ironical about the silly business of a separate Navy. 'To carry this idea to its full extent,' he explained, 'it would not only be necessary to have separate departments, but also a great variety of subdivisions; they must have . . . commissioners of gun barrels and of ramrods.' How often have we heard the same type of irony from naval men asking if we also wanted a separate department of submarines or naval artillery" (282)!

"On the other side of the debate were men like Representative [Harrison Gray] Otis [Federalist-MA] who insisted that 'the services of the War and Naval Departments were . . . perfectly distinct,' adding, 'as well might a Merchant be sent to do the business of a Lawyer; a Lawyer that of a Physician; a Carpenter that of a Bricklayer; or a Bricklayer that of a Carpenter,' as expect Army men to lead the Navy. And Representative [Samuel] Sewall [Federalist-MA] declared that 'it was well known that an officer might be well acquainted with the business of the army, without knowing anything about the business of the navy.' Navy men opposing the separate air [*or space*] power might recall the fact that the bill authorizing the formation of a separate Navy passed the house by the narrow margin of 47 to 41" (283).

The echoes of that historic deliberation reverberate in Kwast's argument. Further justifications for separation—based not on a hatred of the parent or extant services but a genuine desire for the full potential of the new domain to be realized—follow logically. "The technology of war advances with such rapid strides that it quickly outstrips the mental pace of humdrum old-style leaders. Today tactical vision must be coupled with engineering vision. Military leaders must be able to foresee both the new equipment and the tactical implications of that equipment. . . . Even men trained in aviation have difficulty in visualizing the air power of tomorrow; how hapless, therefore, are those trained in totally different fields, who psychologically are incapable of considering space as a separate tactical sphere! That is why a realistic program calls for unequivocal separation of new

weapons from old assumptions, especially in relation to [*space*] power, the newest, the fastest-growing, and most revolutionary military force" (284).

"A new military art, such as inherent in [*spacefaring*], must not be restrained by subservience to . . . old military conceptions. To do so is to lose . . . by default. I am not reflecting on the intelligence or the patriotism or even the professional competence of the Army, Navy, and [*Air Force*] strategists. I am simply taking note that they think as infantry and naval [*and air*] leaders and can never release their minds for [*space*] warfare. [They] will accommodate [*space power*] to their strategy, instead of accommodating their strategy to [*space power*]. Only 100 per cent [*space professionals*] can conceive and carry through a 100 per cent [*space*] strategy" (284–85).

Further evidence of entrenched and obsolete thinking is what happens to so-called reformers. They are labeled radicals and zealots. They are vilified, and their careers are threatened. De Seversky specifically recalls "the crucifixion of General Mitchell," who spoke out for full development of airpower and was court-martialed for it (285). "Even if never expressed in words or orders, those preferences percolate down through the ranks. They make themselves effective in subtle pressures of obedience to authority, loyalty to superiors, and honest hunger for promotion. To the extent that a [*space power*] idea infringes on the traditional sphere of an older service, or throws doubt on the efficacy of that service in a given situation, it is slowed up and frequently doomed to premature death. The channels through which it must move are adjusted to the needs and inertias of an older tradition. Military services, like everything else that is alive, [instinctually have] a will to survive which fights, consciously or unconsciously, against every reality that limits their functions" (285–86).

"The most talented [*space*] strategist or designer cannot contribute all that he has when he is condemned to work within the framework of limited and essentially false basic plans of strategy. It is not merely a matter of 'guts' in speaking out. No one can deny the importance of discipline in the armed forces. Once he has made his recommendations to his superiors, a [*space*] officer feels it is his soldierly duty to say no more—even if he is heartbroken by the manner in which his deeply felt views are filed and forgotten. . . . The subject [*of space power*] is ringed with a taboo. . . . The [*space*] industry, too, is effectively silenced. It does business with the same generals and admirals. Moreover, it is sworn to secrecy in a manner that precludes practical criticism of [*space*] policy" (286).

De Seversky quotes a then-recent Navy memorandum to complete his point: "No officer in the Army or Navy who is even slightly familiar with his profession fails to realize that aviation is a major and essential com-

ponent of both services. Moreover, the airmen who are most vocal in urging now the formation of an additional independent air force are men who are not in the Army, not in the Navy, and are either unfamiliar with or not keenly interested in the broader aspects of national strategy, national defense, and the science of modern warfare." De Seversky responds that "it does not take a mathematical genius to figure out why the most vocal airmen are not in the Army and not in the Navy! The personal tragedies of those who dared to be vocal are only minor items in the larger tragedy of an American air force condemned to domination by the older services" (285–86).

De Seversky goes on to cite a contemporary book of the time on airpower by the chief of Army aviation, Gen Henry H. Arnold, as "a convenient case in point (of airman's inherent subservience, I suppose)" (289). He notes Arnold's admission that "the air forces of all countries will eventually be separated from land and sea forces" and, further, that in America "this long step should be taken, if it is taken at all, only after careful planning and mature thought, and with no zest for radical reform. There should be a stage of gradual evolution as against other knife-cutting of binding ties." De Seversky adds that Arnold "almost in the same breath . . . contradicts himself by writing: '*We shall be fortunate if our time for that reorganization of an independent air force comes in the relative calm of peace or at worst in the preparatory and not in the fighting stage* (emphasis added)'" (289). A continuing argument against an independent Space Force is that it will not be ready until after it has armed combat experience in war, as in the case of the Air Force. I am reminded of the story of the man who would not go into the water until he knew how to swim. To wait until war comes to be properly organized, trained, and equipped for war fully is not an argument made by serious military strategists.

De Seversky notes about Arnold that "this sort of double talk, if it is not the product of honest confusion, may be a device for conveying the truth to those who understand, without offending the powers that be. The general surely knew when he wrote [this] that we were close to the reality of war and that there was no time for the 'gradual evolution' he proposed. He must be quite adequately aware that twenty years of delay—from the time when General Mitchell first demanded independence of aviation, have brought an evolution of approximately zero. Evolutionary processes, too, can be arrested by artificial restraints" (289). Space power has suffered even more impediments.

By 1958 the Air Research and Development Command (ARDC) was soliciting ideas and developing plans that "outlined a rationale for an

'Earth Military Orbital Space Force'" in the "1965–1980 time frame."[5] In April 1959, Adm Arleigh Burke, chief of naval operations, with the concurrence of Army chief of staff Gen Maxwell Taylor, proposed creation of a single unified military space command. However, Gen Thomas White, Air Force chief of staff, "opposed the proposal because . . . it violated the practice of treating space systems on a functional basis and integrating weapons within unified commands. He argued that space systems represent only a better means of performing existing missions and should be assigned to the appropriate unified or specified command."[6] De Seversky had deftly rebutted these arguments against independent airpower, as cited previously in this article.

In late 1981, frustrated by the soaring costs and lack of progress that appeared to have allowed the Soviet Union to leap ahead of American military space capabilities, Congress passed House Resolution 5130 requiring the US Air Force to report to Congress on the feasibility of establishing an independent space command.[7] Entrenched interests in the DOD strongly opposed the move on the grounds it was not needed, would duplicate bureaucracies, and would cost too much despite the findings of a January 1982 General Accounting Office report that a separate space command coordinating all military space activities would instead result in overall cost *savings*. Two years later, the first US Space Command was inaugurated with the goal of transitioning to an independent Space Force by the end of the century.[8]

And still, military space progress languished. In 1999 Congress authorized a commission headed by former secretary of defense Donald Rumsfeld to investigate. The commission's warnings were chilling. Its 2001 report stated flatly that the US was unprepared for "a Space Pearl Harbor," a devastating attack in space that would render its military incapable of projecting power effectively overseas.[9] Among the report's recommendations was to begin preparations for the establishment of an independent Space Force, beginning with a presidential advisory group within three years and full transition within 10 years.[10] Two years later, US Space Command was decommissioned.

History matters. Major de Seversky will get the last word in this essay—modified for today in terms that echo resoundingly in General Kwast's clarion call: "At bottom the objectives to true and independent [*space power*] derive from a flaw in vision. The critics are unable to see the potentialities of [*space power*] beyond the horizons of its present equipment and its present tactics. They do not take in the full majestic sweep of the inevitable progress of [*spacecraft*]; they base their thinking on [*space-*

faring] as they know it today. . . . Not only developments around the corner, but immediate possibilities which trained [*space*] minds know to be simple and sober realities, orthodox strategists dismiss as fantastic and far-fetched" (290).

"The eventual emancipation of [*space power*] from . . . its 'subordinate status' is inevitable. With that emancipation will come a blossoming of [*spacefaring*] talent, a heightened morale among [*space*] personnel, [and] a long-delayed chance for American [*space power*] to come into its own. . . . Every week of delay is a costly and futile restraint on that unfoldment, and a postponement of victory" (291). SSQ

Everett C. Dolman
Air Command and Staff College

Notes

1. Stephen Kwast, "Where the Space Force Must Go," *Politico*, 17 January 2020, https://www.politico.com/.

2. Alexander P. de Seversky, *Victory Through Air Power*, Book-of-the-Month Club ed. (New York: Simon and Schuster, 1942).

3. Page numbers for references to De Seversky's *Victory Through Air Power* are indicated in parentheses throughout this article.

4. See CQ Staff, "Air Force Opposes Creation of Space Corps," Roll Call, 22 June 2017, https://www.rollcall.com/; Sandra Erwin, "As Acting Air Force Secretary, Donovan Takes Lead Role Advocating for Independent Space Force," *Space News*, 17 July 2019, https://spacenews.com/; and Sandra Erwin, "Air Force Changes Message on Space Force amid Criticism It Stifled Debate," *Space News*, 2 September 2019, https://spacenews.com/.

5. Brent Ziarnick, *Tough Tommy's Space Force: General Thomas S. Power and the Air Force Space Program*, Drew Paper no. 34 (Maxwell AFB, AL: Air University Press, 2019), 55, 57, https://media.defense.gov/.

6. David Spires, *Beyond Horizons: A Half Century of Air Force Space Leadership*, rev. ed. (Maxwell AFB, AL: Air University Press, 1998), 76, https://www.airuniversity.af.edu/.

7. Thomas Karas, *The New High Ground* (New York: Simon & Schuster, 1983), 17–20; and Spires, *Beyond Horizons,* 76.

8. Spires, *Beyond Horizons*, 217–21, 230–34.

9. *Report of the Commission to Assess United States National Security, Space Management and Organization* (Washington, DC: The Commission, 2001), xiii–xv. https://aerospace.csis.org/.

10. *Report of the Commission to Assess United States National Security*, 80, 83–84, 94.

FEATURE ARTICLE

The Great Divide in US Deterrence Thought

Keith B. Payne

Abstract

This article assesses different sides of the nuclear deterrence debate by examining opposing narratives supporting the relative simplicity versus complexity of maintaining deterrence.* The "easy deterrence" narrative posits that the essential requirements for stable mutual deterrence are not difficult to understand or meet and that the functioning of mutual deterrence can be considered largely predictable and reliable. In sharp contrast to this easy deterrence narrative, the alternative narrative envisions deterrence as difficult to establish and sustain and potentially requiring greater nuclear capabilities, contingency planning, and for some, strategic defensive capabilities. This article emphasizes the speculative nature of predictions about deterrence but suggests that the "difficult deterrence" narrative, with its recommended spectrum of deterrence threat options and focus on credibility, seems the most prudent in the contemporary threat environment. While the easy deterrence narrative has considerable confidence in deterrence because all rational or sensible leaders are expected to be cautious when confronted with a threat of societal destruction, the difficult deterrence narrative contends that such an expectation may be a "fatal error." Difficult deterrence offers neither a definitive solution to the threat of nuclear use nor ease. It offers no cooperative global transformation and disarmament or confidence that deterrence will work easily and predictably across time. Instead, the difficult deterrence narrative confronts a dilemma that the easy deterrence concept avoids.

*****

For seven decades, the question at the heart of the US nuclear policy debate has been how best to deter—specifically, the ease or difficulty of acquiring and maintaining effective deterrence capabilities. Contending easy and difficult deterrence narratives start from several

*This article is adapted from the author's book *Shadows on the Wall: Deterrence and Disarmament* published by the National Institute Press in March 2020.

common realist assumptions but posit contrary expectations about opponent calculations, decision-making, and behavior. The different expectations separating these narratives establish largely incompatible standards for deterrence and lead logically and directly to contrary definitions of deterrence requirements and the conditions for stability. The differences separating the easy and difficult deterrence narratives are so stark that proponents of the former often mistakenly contend that the latter is preparation for nuclear "war fighting," not a deterrence strategy. The constant push and pull between these two alternative conceptions of deterrence and their respective requirements helps explain the decades-long US public nuclear policy debate. Observers wanting to understand how seemingly equally credentialed experts can make diametrically opposed claims about deterrence requirements must first understand the two competing narratives underlying those opposing claims. They must look behind the curtain.

Proponents of each narrative often criticize the other for offering risky force recommendations. They apply different and largely incompatible metrics to answer the question, How much is enough for deterrence? and unsurprisingly reach contrary conclusions. For example, depending on which deterrence narrative is the basis for judging, maintaining the US triad of strategic nuclear forces is either essential for continued deterrence or an unnecessary, destabilizing waste of resources. The integrity of the claim depends on the veracity of the narrative. However, neither side can rightly claim to "know" what will be required for the deterrence of unknown opponents in unknown circumstances over unknown stakes—it must be a matter of some speculation given the woefully incomplete information we have about the future. This article assesses both sides of the debate and suggests that the most prudent choice in the contemporary threat environment is the difficult deterrence narrative's recommended spectrum of deterrence threat options and its focus on the credibility of those options. This suggestion, however, is made with full recognition of the unavoidable speculation involved. When it comes to projecting the future functioning of deterrence and its requirements against yet-unknown opponents in yet-unknown conditions, we are all amateurs looking at shadows on the wall.

The Easy Deterrence Narrative

During the Cold War, renowned academics—including Kenneth Waltz, Thomas Schelling, Bernard Brodie, and Robert Jervis—advanced basic points of an enduring narrative about mutual nuclear deterrence

commonly referred to as a "stable balance of terror"—"stable" meaning that neither side would have irresistible incentives to employ nuclear weapons or engage in provocations that would likely escalate to nuclear use. The simplified (and popularized) easy deterrence narrative is a pastiche that draws from these scholars' innovative analyses. It incorporates various points appearing in their works but does not reliably reflect their variation, subtlety, or nuance. The discussion here *does not suggest* that Waltz, Schelling, Brodie, or Jervis necessarily advanced or even accepted all facets of the easy deterrence narrative, but that it incorporates various points found in their works.

The easy deterrence narrative posits that the essential requirements for stable mutual deterrence are not difficult to understand or meet and correspondingly that the functioning of mutual deterrence can be considered largely predictable.[1] Those deterrence requirements are rational or "sensible" leadership decision-making (or "a little common sense") and properly controlled and safeguarded capabilities for strategic nuclear retaliation even after suffering an attack.[2] Schelling concludes that with such survivable second-strike capabilities, "a powerfully stable mutual deterrence results."[3] This narrative of relatively easy deterrence survived the Cold War and continues to be a powerful theme in public discussions of deterrence and its requirements.

The confidence of the easy deterrence narrative follows in part from three traditional axioms of realist thought. First, the international system is anarchic with mistrust among its members. Second, in this system, great powers generally prioritize the pursuit of national survival in their decision-making. Third, great powers are rational in that pursuit.[4] Easy deterrence confidence follows from both a macro view of the international system and its anarchic structure and from the expectation that individual units (countries) will behave in predictable ways when facing a severe nuclear threat. Leaders able to reason will seek to avoid those actions that could precipitate an opponent's devastating nuclear response. As Waltz states, "Deterrent policies derive from structural theory, which emphasizes that the units of an international political system must tend to their own security as best they can. . . . A little reasoning leads to the conclusions that to fight nuclear wars is impossible and that to launch an offensive that might prompt nuclear retaliation is obvious folly. To reach those conclusions, complicated calculations are not required. . . . [Deterrence] depends on fear. To create fear, nuclear weapons are the best possible means."[5] He adds, "In a nuclear world, to act in blatantly offensive ways is madness."[6] In short, the priority goal of protecting national sur-

vival and the fear naturally created by nuclear weapons combine to compel caution and provide deterrence.[7]

In 1983, the Harvard Nuclear Study Group observed that since the establishment of nuclear deterrence, the international system had avoided great power war despite many pressures that previously would likely have triggered conflict. The study group attributed this stability to the nuclear "balance of terror." Why this unprecedented stability? It is because "nuclear weapons have created what we call the *crystal ball effect*. . . . This crystal ball effect helps give the nuclear world at least some measure of stability. Statesmen in the atomic age can envision the destruction of a full-scale nuclear war and it makes them determined to avoid it" (emphasis in original).[8]

Using the same "crystal ball" metaphor, Waltz points to the *unique* effectiveness of nuclear weapons for deterrence: "With conventional weapons, the crystal ball is clouded. With nuclear weapons, it is perfectly clear."[9] That is, "in a conventional world, deterrent threats are ineffective because the damage threatened is distant, limited, and problematic. Nuclear weapons make military miscalculations difficult and politically pertinent prediction easy."[10] A properly structured balance of terror is an overwhelmingly powerful deterrent because leaders must be cautious when national survival is at risk.

President John Kennedy's national security advisor, McGeorge Bundy, observed that US and Soviet Cold War decision-making reflected this powerful peacekeeping effect of nuclear deterrence: "The stalemate that keeps nuclear peace between the superpowers is so deep and strong that it is not affected by the relative ruthlessness of the two societies or their different experience of twentieth-century war. What each can do to the other, whoever goes first, is more than enough to stay every hand that does not belong to a madman. . . . The imperative of avoiding nuclear war imposes great caution on both governments."[11]

Waltz elaborates on the same expectation: "Differences among nuclear countries abound, but for keeping the peace what difference have they made? . . . In a nuclear world, any state—whether ruled by a Stalin, a Mao Zedong, a Saddam Hussein, or a Kim Jong Il—will be deterred by the knowledge that aggressive actions may lead to its own destruction. . . . Who cares about the [differing] 'cognitive' abilities of leaders when nobody but an idiot can fail to comprehend [nuclear weapons'] destructive force."[12] Waltz also suggests that "in a nuclear world any state will be deterred by another state's second-strike forces. One need not become preoccupied with characteristics of the state that is to be deterred or scrutinize its leaders."[13] Why is this so? The answer, according to Bundy, is that "in

the real world of real political leaders—whether here or in the Soviet Union—a decision that would bring even one hydrogen bomb on one city of one's own country would be recognized in advance as a catastrophic blunder; ten bombs on ten cities would be a disaster beyond history; and a hundred bombs on a hundred cities are unthinkable."[14]

For leaders who are not "mad," war between nuclear powers in a proper balance of terror simply *must* be avoided—no goal short of national survival could justify an action that would seriously risk "unthinkable" national destruction. Thus, "if decision makers are 'sensible,' peace is the most likely outcome."[15] This expectation regarding leadership decision-making inspires confidence in the predictable functioning of nuclear deterrence among states. Indeed, Bundy coined the term "existential deterrence"—meaning that a balance of terror creates the conditions needed for effective mutual deterrence without regard to many other factors—including the character of the sides involved or their relative nuclear capabilities.[16] Waltz concurs, stating, "The [deterrence] effects of nuclear weapons derive not from any particular design for their employment in war but simply from their presence."[17]

This point of the easy deterrence narrative leads to considerable confidence in a balance of nuclear terror to produce the reliable functioning of mutual deterrence. There are numerous popular, expert, and official Cold War and post–Cold War illustrations of this confidence in the predictable effect of nuclear deterrence. For example,

- "In US-Soviet relations, the current nuclear postures have substantially solved the problem of deterring deliberate nuclear attack. Under present conditions, no rational leader could conclude that his or her nation would be better off with a nuclear war than without one."[18]
- "In the light of the certain prospect of retaliation there has been literally no chance at all that any sane political authority, in either the United States or the Soviet Union, would consciously choose to start a nuclear war. This proposition is true for the past, the present, and the foreseeable future. For sane men on both sides, the balance of terror is overwhelmingly persuasive."[19]

In short, this narrative deems deterrence to be reliably effective in preventing large-scale attacks. In the absence of a "madman," the crystal ball effect unique to nuclear weapons is expected to provide even the most aggressive leadership with clarity regarding the intolerable catastrophe that would attend nuclear conflict, thus preventing nuclear war and con-

flicts considered likely to escalate to nuclear war. A balance of terror provides existential deterrence.

Easy Deterrence Components

As noted, axioms of traditional realist thought underlie the easy deterrence narrative's expectation of the predictable, effective functioning of nuclear deterrence. However, three additional extra-realist propositions about leadership decision-making and the functioning of deterrence contribute to the conclusion that deterrence is largely predictable and "easy": societal threats, graduated options, and uncertainty.

Societal threats. The easy deterrence narrative typically contends that the nuclear "second-strike" capabilities needed for deterrence *are the survivable forces required to threaten punishment on an opponent's societal assets*. Nuclear capability *beyond* that needed to threaten societal destruction does not add to deterrence; it is unnecessary "overkill."[20] The expectation that a punitive threat of societal destruction can serve as the basic measure of adequacy for reliable deterrence was particularly prominent in Cold War academic treatments of the subject and in official US declarations regarding deterrence. Societal assets could include cities, industry, power, transportation, and population. However, from early in the Cold War, threats to population and cities became a type of shorthand for measuring deterrence strategies: "A force sufficient to kill the enemy's population and destroy his wealth is an adequate deterrent,"[21] and "Deterrence comes from having enough to destroy the other's cities; this capability is an absolute, not a relative, one."[22] The expectation of effective deterrence from societal threats (cities) has continued to be a theme in public commentary in the post–Cold War era: "Deterrence today would remain stable even if retaliation against only ten cities were assured."[23]

This general expectation that punitive threats against an opponent's society can deter reliably is not based on the unreasonable presumption that all opponents are enlightened and dutiful civil servants. Instead, "what government would risk sudden losses of such proportion or indeed of much lesser proportion? Rulers want to have a country that they can continue to rule," and this would be put at risk by nuclear deterrent threats to their societies.[24] Deterrence works reliably and predictably at nuclear force levels that are easy to acquire and maintain because opponents with a modicum of common sense will place decisive value on the preservation of their nation's societal assets—which generally are relatively few, undefended, and highly vulnerable to modest numbers of nuclear weapons.[25]

Confidence in deterrence based on punitive societal threats was not limited to academic discussions. In the 1960s, Secretary of Defense Robert McNamara's public and then-classified pronouncements regarding nuclear deterrence specified the US threat levels to Soviet society (population and industry) considered adequate for reliable nuclear deterrence. The precise numbers of his "assured destruction" metric for US deterrence capabilities shifted somewhat over time, but in a declassified 1964 draft presidential memorandum, Secretary McNamara defined US deterrence requirements as the US ability to destroy "25 percent of [the Soviet] population (55 million people) and more than two-thirds of [Soviet] industrial capacity."[26] McNamara's "assured destruction" metric represented the "flat of the curve" for the number of US nuclear weapons needed to threaten the specified levels of destruction on the Soviet population and industry. Beyond a specific number of weapons, the additional level of societal destruction possible with each additional weapon rapidly diminished. The computation of the declining marginal value of additional US nuclear weapons against Soviet societal targets determined the percentiles declared as deterrence standards and the capabilities necessary to meet those standards.[27]

By the mid-1960s, Secretary McNamara expressed great confidence in the reliability of this type of deterrence threat: "Such a capability would, with a high degree of confidence, ensure that we could deter under all foreseeable conditions, a calculated, deliberate nuclear attack on the United States."[28] He "had come to believe that the US deterrent capability, the nation's strategic offensive forces, not the damage-limiting strategic defensive forces, protected American society."[29]

There is no single commonly agreed number of nuclear weapons deemed adequate for deterrence based on punitive threats to an opponent's societal assets. However, many scholars suggest a range from "hundreds" to far fewer.[30] This deterrence narrative and metric has since become a prominent, enduring theme in public commentary about US nuclear deterrence requirements. The following are some examples:

- "Most professional analysts of the subject believe that the prospect of about one hundred thermonuclear warheads exploding over urban areas is more than enough to deter either side from starting a nuclear war. . . . I personally believe that very much smaller numbers are sufficient to deter war; I have used numbers like one hundred only because it is customary to do so in such arguments."[31]
- "No current or conceivable threat to the United States requires it to maintain more than a few hundred survivable nuclear weapons. The

delivery of fewer than a hundred warheads could destroy the society and economy of any country, and tens of detonations could kill more people than have ever been killed in any previous war."[32]

- "A reasonably small force of several hundred weapons would allow that state to strike back over 100 times before it had to negotiate. No state on the planet could withstand that sort of punishment, and no sane leader would run that sort of risk."[33]
- "Ten to one hundred survivable warheads should be more than enough to deter any rational leader from ordering an attack on the cities of the United States or its allies."[34]
- "Fewer than 100 warheads is sufficient to inflict a wholly unacceptable level of damage on a continental-sized economy, and suggests that—even for the most enthusiastic proponent of nuclear deterrence—maintaining an arsenal at higher than that level is unnecessary."[35]

In short, this narrative contends that deterrence is easy to achieve and sustain because a punitive nuclear threat of societal destruction requires relatively few weapons and is made transparent by the crystal ball effect. A modest survivable second-strike capability will serve to compel all but "mad" leaders to "draw back from the brink."[36]

Graduated options for deterrence. An important additional element of this easy deterrence narrative is the contention that US capabilities for limited or graduated nuclear threats can help to deter limited attacks or prevent escalation in the event an initial failure of deterrence leads to an opponent's limited attack. Schelling, Waltz, and Jervis essentially endorsed the threat of graduated (limited) response options as part of a US deterrence strategy to demonstrate US will and thereby deter limited attacks and help control escalation if deterrence initially fails.[37] They identified the capability for graduated demonstration-of-will threats as particularly important for extending US nuclear deterrence coverage to allies.

This rationale for graduated nuclear options is not based on the potential military effects of limited nuclear employment but on certain expectations. One is that brandishing limited nuclear threats can help deter an opponent's limited nuclear provocation by having proportional response options. Another is that having graduated threat options can help deter an opponent's escalation if deterrence initially fails by demonstrating US resolve to escalate further if necessary—but doing so well below the ultimate deterrence threat of societal destruction. That "ultimate threat" could be held in reserve to help ensure that the opponent would continue to have

an overwhelming incentive *not* to further escalate a conflict for fear of unleashing its own nuclear destruction.

This rationale for limited nuclear threats also does not presume that a "limited" nuclear war could be fought safely or with any certainty that a limited war would remain so in any meaningful sense. It also is not an endorsement of nuclear "war fighting." Rather, the existence of graduated options is intended to contribute to both the deterrence of limited attacks and the reestablishment of deterrence ("intra-war deterrence") to bound escalation following a limited attack. Brandishing limited nuclear response options and thereby demonstrating resolve for these deterrence purposes has been an apparent feature of declared US deterrence policy since the mid-1970s.[38] As Secretary of Defense James Schlesinger observed at the time, "To the extent that we have selective response options—smaller and more precisely focused than in the past—we should be able to deter such challenges. Nevertheless, if deterrence fails, we may be able to bring all but the largest nuclear conflicts to a rapid conclusion before cities are struck. Damage may thus be limited and further escalation avoided."[39] The capacity to issue graduated nuclear threats necessitates the manifest planning and control of forces that enable graduated threats to be known to opponents for deterrence purposes.[40]

Uncertainty deters. The easy deterrence narrative typically includes another expectation about the functioning of deterrence that facilitates the conclusion that deterrence is easy and reliable. This second expectation is that opponents' *uncertainty* about whether, when, and how the United States actually would execute its nuclear deterrent threat can provide *adequate credibility* for effective deterrence. An opponent's uncertainty about the US deterrent threat can deter attack because the consequences of US retaliation would be so catastrophic if executed: "Uncertainty of response, not certainty, is required for deterrence because, if retaliation occurs, one risks losing so much."[41] Thomas Schelling posited that deterrence depends *not* on an opponent's certainty that nuclear catastrophe would follow its highly aggressive action but on the *"chance"* that catastrophe would follow. Schelling famously called this "the [deterrence] threat that leaves something to chance."[42]

Early in the Cold War, Schelling explained why uncertain nuclear threats can deter in a balance of terror: "Any situation that scares one side will scare both sides with the danger of a war that neither wants, and both will have to pick their way carefully through the crisis, never sure that the other knows how to avoid stumbling over the brink."[43] The possibility or

chance of great loss provides adequate deterrence because "a low probability of carrying a highly destructive attack home is sufficient for deterrence."[44]

This easy deterrence expectation that the chance of threat execution can provide effective deterrence is particularly important to the integrity of America's extended nuclear deterrence umbrella to allies. It answers the concern that America's threat of employing nuclear weapons in defense of allies should be seen as an unconvincing bluff, not an effective deterrent, *because doing so could easily result in the destruction of the United States*.[45] However, extended deterrence *can* be sufficiently credible to work—even if US execution of the threat would be self-destructive—because the opponent *could never be sure that the United States would not execute its deterrent threat* by way of an irrational act or the fog of war, and that chance is adequate to deter.[46] Consequently, the seemingly innocuous proposition that uncertain nuclear threats can deter reliably plays a central role in the easy deterrence narrative, particularly for extended deterrence.

This contention that deterrence can be based on uncertainty regarding the US nuclear threat contributes significantly to defining how much is enough for deterrence in terms that are easily achieved and sustained. How so? Effective deterrence does *not* require that the country attempting to deter *convince* opponents that it would, in fact, deliberately, rationally execute its nuclear deterrent threat if provoked. The country seeking to deter can forego those offensive or defensive strategic capabilities that might otherwise be thought necessary to *fully convince* opponents of the certainty of its deterrence threat—that it surely would be executed as threatened. Such additional requirements are unnecessary for effective deterrence because threat credibility with that sense of certainty is unnecessary. In short, not only are the required second-strike nuclear forces relatively modest and easily sustained, but additional strategic capabilities *are unnecessary to convince opponents of the logical credibility of the threat*.

Easy Deterrence: Reasons for Optimism

This narrative leads to considerable optimism about the ease and predictability of deterrence. However, Waltz, Brodie, and Jervis acknowledge that deterrence is not "ironclad" or foolproof. It could be upended by leaders who are not "sensible" in that "a small-number system can always be disrupted by the actions of a Hitler and the reactions of a Chamberlain."[47] Jervis elaborates on why this is so: "Miscalculations are possible, even in situations that seem very clear in retrospect, and states are sometimes willing to take what others think are exorbitant risks to try to reach highly valued goals."[48]

Despite this caveat regarding deterrence that "nothing in this world is certain," considerable "comfort" and "cautious optimism" are warranted with regard to the reliability of deterrence—not because leaders are reliably virtuous but because the dangers involved in reckless behavior are so obvious (the crystal ball effect).[49] According to Waltz, "With nuclear weapons, it's been proven without exception that whoever gets nuclear weapons behaves with caution and moderation. Every country—whether they are countries we trust and think of as being highly responsible, like Britain, or countries we distrust greatly for very good reasons, like China during the Cultural Revolution—behaves with such caution."[50] Brodie adds, "The leaders of no country will wish to risk the total destruction of their country. . . . We should not complain too much because the guarantee is not ironclad."[51]

In addition, even if an individual leader is reckless, "those who direct the activities of great states are by no means free agents."[52] There may be a variety of pressures contributing to sensible behavior that are beyond the character of a particular leader or leadership, including the possible refusal of those who carry out orders to engage in reckless actions. As 2007 Nobel Laureate professor Roger Myerson says in this regard, "Our most dangerous adversaries are not lone madmen, however, but are leaders with political support from many people who have normal hopes and fears. Psychopathic militarists like Hitler become a threat to our civilization only when ordinary rational people become motivated to support them as leaders."[53] For nuclear deterrence in a proper balance of terror to fail, "one would have to believe that a whole set of leaders might suddenly go mad."[54]

Also, as noted, the absence of war among great powers since the end of World War II suggests the predictable reliability of nuclear deterrence: "Never since the Treaty of Westphalia in 1648 . . . have great powers enjoyed a longer period of peace than we have known since the Second World War. One can scarcely believe that the presence of nuclear weapons does not greatly help to explain this happy condition."[55] This extraordinary history, according to Waltz, argues *against* the expectation that nuclear deterrence is likely to fail. Indeed, he finds it "odd" that "a happy nuclear past leads many to expect an unhappy nuclear future."[56]

Finally, this narrative includes the expectation that if deterrence fails at some point, graduated nuclear options and intra-war deterrence can help prevent escalation. While there are no guarantees, "even if deterrence should fail, the prospects for rapid de-escalation are good."[57]

Easy Deterrence Force Planning

The force recommendations that follow from the easy deterrence narrative focus on the need for survivable nuclear capabilities that *are able* to hold an opponent's societal assets at risk but *are not able to threaten the opponent's own deterrence forces*. What is the reason for this balance? Stable deterrence follows from the ability to threaten the opponent's societal assets, but a capability to threaten the opponent's own deterrence forces is not consistent with a stable *mutual* balance of terror. Forces able to strike an opponent's offensive deterrence forces on the ground and air defenses or ballistic missile defenses intended to protect cities are likely ineffective and unnecessary for deterrence. And they would threaten to upset stability by calling into question the opponent's deterrence capability. Such forces could cause an opponent to doubt the effectiveness of its deterrent and thus drive the opponent to move in haste to gain the possible advantage of striking first rather than waiting to absorb an initial undeterred blow.[58]

Offensive or defensive forces that might threaten the pre- or post-launch survivability of an opponent's deterrence forces could in this way create what Schelling called "the reciprocal fear of surprise attack."[59] Avoiding such "destabilizing" forces promotes reliable deterrence because "the *likelihood* of war is determined by how great a reward attaches to jumping the gun, how strong the incentive to hedge against war itself by starting it, [and] how great the penalty on giving peace the benefit of the doubt in a crisis" (emphasis in original).[60]

Confidence in the reliable working of deterrence also contributes to the conclusion that attempting to physically protect society against strategic nuclear attack not only is unnecessary for deterrence and potentially destabilizing but also is of limited potential value because a stable balance of terror provides reliable protection via deterrence—and does so *in the absence of such defenses*. Waltz summarizes this point with the rhetorical question, "Why should anyone want to replace stable deterrence with unstable defense?"[61] He further states, "In a nuclear world defensive systems are predictably destabilizing. It would be folly to move from a condition of stable deterrence to one of unstable defense."[62]

Correspondingly, the easy deterrence narrative's general guidelines for stable deterrence *favor* the survivable, modest offensive capabilities able to threaten societal assets, which could include "targets that are crucial to a nation's modern economy, for example, electrical, oil, and energy nodes [and] transportation hubs."[63] However, the same guidelines *argue against* (1) offensive nuclear systems able to threaten the pre-launch survivability of an opponent's own deterrence forces and (2) strategic defensive capabilities

designed to physically protect one's own society. The concern that strategic defenses could destabilize deterrence has been a prominent theme in US strategic thought and policy for decades and continues to be present.[64]

In summary, based initially on several points of traditional realist thought and the three elaborations described above, early in the Cold War a compelling and even comforting deterrence narrative emerged that posits the ease and efficacy of a balance of terror to prevent nuclear or large-scale conventional war. US policy never followed this deterrence narrative in its entirety, and official public characterizations of US deterrence requirements became increasingly distant beginning in the mid-1970s. However, the easy deterrence narrative's relatively modest and narrow force requirements for a stable balance of terror and related arguments against "destabilizing" strategic offensive and defensive capabilities were, and continue to be, prominent themes in the US public debate about deterrence and force requirements.

The Difficult Deterrence Narrative

During the Cold War, a largely separate set of prominent academics and scholars contributed to a markedly different nuclear deterrence narrative. This alternative narrative certainly concurs that deterrence is valuable, even necessary, for the United States in an anarchic international system. It includes some variation and has evolved over decades. However, in sharp contrast to the easy deterrence narrative, this alternative narrative envisages deterrence as *difficult* to establish and sustain and as potentially demanding considerably greater nuclear capabilities, contingency planning, and for some, strategic defensive capabilities. Correspondingly, its general force guidelines are quite different from those of the easy deterrence narrative.

This alternative deterrence narrative emerged early in the Cold War—at least in part in response to the inadequacies some scholars saw in the sanguine easy deterrence narrative. They included most prominently Herman Kahn and Albert Wohlstetter and, more recently, Colin Gray. For example, Wohlstetter begins his famous 1958 paper, *The Delicate Balance of Terror*, with the following:

> I should like to examine the stability of the thermonuclear balance which, it is generally supposed, would make aggression irrational or even insane. The balance, I believe, is in fact precarious, and this fact has critical implications for policy. Deterrence in the 1960's will be neither inevitable nor impossible but the product of sustained intelligent effort, attainable only by continuing hard choice. . . .

> . . . While feasible, it will be much harder to achieve in the 1960's than is generally believed. One of the most disturbing features of current opinion is the underestimation of this difficulty.[65]

This classic early work captures much of the basic criticism of the easy deterrence narrative and points toward a competing deterrence narrative derived from the works of Kahn, Wohlstetter, Gray, and others. In contrast to the easy deterrence narrative, it sees the pursuit of deterrence as an ongoing and difficult challenge, with no fixed approach and no corresponding finite and fixed set of nuclear capabilities that can predictably provide the desired deterrent effects.

Bernard Brodie's commentary on Wohlstetter's contention of a "delicate balance of terror" illustrates a fundamental disagreement between these two competing deterrence narratives: "I could never accept the implications of [Wohlstetter's] title—that the balance of terror between the Soviet Union and the United States ever has been or ever could be 'delicate.' My reasons have to do mostly with human inhibitions against taking monumental risks or doing things which are universally detested."[66] Brodie's comments demonstrate that while both approaches to deterrence share some points of realist origin, they posit (1) strikingly different expectations about possible leadership decision-making and behavior; (2) different expectations about the functioning of deterrence; and (3) correspondingly, different answers to the question of how much is enough for deterrence.

While the easy deterrence narrative has considerable confidence in deterrence because rational or sensible leaders can generally be expected to be cautious when confronted with a threat of societal destruction, the difficult deterrence narrative contends that such an expectation may be a "fatal error." This is because "not all actors in international politics calculate utility in making decisions in the same way. Differences in values, culture, attitudes toward risk-taking, and so on vary greatly."[67] Rational or sensible leadership decision-making can include unique decision-making factors that drive leaders' perceptions and calculations of value, cost, and risk in surprising, unpredictable directions. An examination of multiple international crises leads to the assessment that "the personality of an individual determines the reaction to information and events. A leader's nationality, passion, idealism, cynicism, pragmatism, dogmatism, stupidity, intelligence, imagination, flexibility, stubbornness, and so on, along with mental disorders such as depression, anxiety, and paranoia, shape reactions and decisions during a crisis."[68]

Colin Gray points to the variability in leadership perceptions and calculations in his full rejection of the easy deterrence narrative. Such uncertain-

ties present deterrence planners (and defense planners in general) with an incomplete basis for confident prediction of opponent behavior "no matter how cunning their methodology or polished their crystal ball."[69] As a result, "there is massive uncertainty over 'what deters' (who? on what issue? when?)."[70]

In short, the functioning of deterrence "is heavily context dependent."[71] The confident expectation of an opponent's sensible caution when confronting a severe societal deterrence threat may be upset by a variety of factors that may not be obvious in advance to an outside observer. For example, an opponent's hierarchy of values may posit an expectation of *intolerable cost* associated with *not* acting; an opponent may be willing to take great risks in unwavering pursuit of a cherished goal or be confident that it need not conciliate because the deterrer itself will yield; an opponent may be unwilling or unable to recognize great risk; and unexpected technical, operational, or organizational factors may prevent the hoped-for deterrence outcome from prudent leadership decision-making.[72]

A 2014 study by the National Research Council of the National Academy of Sciences emphasizes that accounting for such factors is critical in considerations of deterrence.[73] Indeed, advancements in cognitive science have demonstrated the significant degree to which decision-making often is governed not only by prudent cost-benefit calculations but also by many possible less predictable factors. These include people's tendencies to act on the basis of their emotional and cognitive biases, such as seeking confirmation of their opinions; being overly optimistic; focusing on their wins versus losses, or survivorship; and making decisions based on expectations of normalcy versus planning for new or potential events.[74]

Numerous analyses of historical case studies indicate that these types of factors can affect leadership decision-making and thus the functioning of deterrence. [75] In one such study, King's College professor Richard Ned Lebow notes that "even the most elaborate efforts to demonstrate prowess and resolve [for deterrence] may prove insufficient to discourage a challenge when policy makers are attracted to a policy of brinkmanship as a necessary means of preserving vital strategic and domestic political interests." He further suggests that "these cases and others point to the importance of motivation as the key to brinkmanship challenges. To the extent that leaders perceive the need to act, they become insensitive to the interests and commitments of others that stand in the way of the success of their policy."[76]

The difficult deterrence narrative does not share the expectation that the crystal ball effect will reliably mitigate all such inherently human

decision-making factors and thereby reliably lead to opponents' caution and prudence in response to even severe punitive deterrence threats. Because the definition of what constitutes "sensible" leadership thinking and behavior can vary greatly, generalized expectations of how easily deterrence *should* function vis-à-vis sensible leaders and the finite forces that *should* reliably deter them may be misleading at best *apart from the specifics of a case*.[77] This conclusion contrasts sharply with the easy deterrence expectation that sensible leaderships can be expected to respond similarly to a severe deterrence threat, that is, with caution.

Variation in Decision-Making

A reliably effective deterrent threat must overcome the possibly wide variation in opponents' perceptions, calculations, and decision-making. For the difficult deterrence narrative the implications of this seemingly academic point are enormous. For example, different leaderships may place highest priority on different values, national or personal—potentially dashing confident expectations of how deterrence *should* function if leaders are sensible. While Schelling, Waltz, Jervis, and Brodie clearly recognize the potential of variability in decision-making and the fact that deterrence is not "ironclad," it is the difficult deterrence narrative that is influenced greatly by the potentially distorting effects of this variation on the predictable functioning of deterrence and on its requirements.

The difficult deterrence narrative points to various historical case studies in which opponents did not reason or behave as expected by the easy deterrence narrative as evidence that deterrence may not function as it projects. While many historical examples are available, two brief cases illustrate this point. The first occurred in 1962, when Nikita Khrushchev moved nuclear weapons to Cuba despite his expectations that, as a consequence, "they can attack us and we shall respond" and that "this may end in a big war."[78] During the same crisis and in an expression of socialist ideological zeal, the Cuban leadership urged Soviet leaders to launch a nuclear attack against the United States despite its recognition that the consequences would be a horrific war and the destruction of Cuba.[79]

In the second example, occurring in October 1973, Egypt and Syria launched a massive armored attack against Israel to restore national honor despite the reported possibility of Israeli nuclear retaliation.[80] US leaders were surprised by this large-scale attack on Israel and the Yom Kippur War because, according to Secretary of State Henry Kissinger, "Our definition of rationality did not take seriously the notion of [Egypt and Syria] starting an unwinnable war to restore self-respect. There was no defense

against our own preconceptions."[81] The Egyptian and Syrian leaderships' decision to risk another war with Israel followed from their extreme motivation to change the status quo and restore honor.[82] Their motivation to act overshadowed their caution, a dynamic that was outside Kissinger's "definition of rationality" and easy deterrence projections.

The difficult deterrence narrative is driven by the point that this variability in leadership decision-making means that *deterrence is not easy*, in part because deterrence planning must be done "with reference to the unique details of the case in hand."[83] It must, for example, reflect an understanding of the potentially unique values or assets held most dear by different opponents because those values or assets may vary and yet must be put at risk for deterrence purposes. In contemporary deterrence policy jargon, this narrative has evolved to include the corresponding point that deterrence strategies must be "tailored" to take into account the unique context and characteristics of the opponent in question.[84]

The easy deterrence narrative typically contends that a punitive US threat to an opponent's societal infrastructure reliably constitutes a near-universally applicable deterrence threat. But having an arsenal suited to threaten "easy" societal targets may be inadequate for deterrence if the material assets or intangible values opponents hold most dear are other than *societal assets*. To focus only on one general form of punitive deterrence threat per the easy deterrence narrative risks having a strategy that does not apply to the opponent in question when necessary. Consequently, this difficult deterrence narrative contends that deterrence requirements must include nuclear forces capable of threatening a spectrum of plausible opponent values and assets, potentially including an opponent's diverse and hardened military targets. Kahn insisted during the Cold War that an "adequate" US nuclear deterrent "demands more and better offensive forces" than the simple capability to threaten societal assets.[85] This call for greater and more diverse nuclear capabilities does not reflect a "nuclear war-fighting" goal vice deterrence as easy deterrence critics often contend. Rather, it reflects a more demanding definition of deterrence requirements. Indeed, the logic of the difficult deterrence narrative is captured by Kahn's observation that planning for deterrence should not be limited to expectations of

> a complacent and cautious enemy. Even a frown might do that [deter]. Our attitude should be the same as an engineer's when he puts up a structure designed to last twenty years or so. He does not ask, "Will it stand up on a pleasant June day?" He asks how it performs under stress, under hurricane, earthquake, snow load, fire, flood, thieves, fools and

> vandals. . . . Deterrence is at least as important as a building, and we should have the same attitude toward our deterrent systems. We may not be able to predict the loads it will have to carry, but we are certain there will be loads of unexpected or implausible severity.[86]

Physically Defending against Nuclear Attack

The differences separating the easy and difficult deterrence narratives transcend the latter's greater requirements for strategic nuclear capabilities. For *a select cadre of those contributing to the difficult deterrence narrative*—including Herman Kahn, Donald Brennan, and Colin Gray—it also included the requirement for some capabilities to physically defend the United States from nuclear attack. Their argument for strategic defensive capabilities follows again from the significance for deterrence of the variation in leadership decision-making—in this case revolving around *the need for deterrence threat credibility*.

Herman Kahn was particularly critical of the easy deterrence contention that uncertainty or chance can provide sufficient threat credibility for reliable deterrence.[87] He emphasized the connection between physically defending the United States and having a *sufficiently credible* extended nuclear deterrent for allies. Kahn insisted that a "not incredible" US deterrence threat is needed to extend deterrence coverage to allies reliably, and that achieving *that level of threat credibility is dependent on the US capability to protect American society to some extent*. Why so? Kahn argued, in a true "balance of terror," the US extended deterrent was likely to *be incredible* because the opponent's capability for a nuclear reply against US society could be seen as precluding US willingness to employ nuclear weapons on behalf of an ally. He believed that such an act could, in effect, be suicidal for the United States and thus not adequately credible for extended deterrence, stating that "it will be irrational [for the United States] to attack and thus insure a Soviet retaliation unless we have made preparations to counter this retaliation."[88] His basic point was that deterrence credibility is based not only on the US threat that can be posed against an opponent, but on the punishment the opponent could inflict in return on the United States if Washington were to carry out its deterrent threat.[89] If an opponent's punishment of the United States is likely to be seen as intolerable, even if the US deterrent threat is recognized as severe by the opponent, *that threat may not be sufficiently credible to deter in a crisis.*

Neither Kahn nor Gray suggests that for deterrence purposes such defensive preparations must be "perfect" (likely a hopeless goal), but rather sufficiently effective to lead the opponent to conclude that it cannot

dismiss the US extended nuclear deterrent given US homeland vulnerability. This point of the difficult deterrence narrative contends that US extended deterrence responsibilities establish a requirement for some US homeland defensive capabilities that the easy deterrence narrative typically deems "destabilizing."[90]

Kahn and Gray also advanced a separate rationale for strategic defense capabilities that follows logically from the difficult deterrence narrative's theme that deterrence is subject to potential failure. Kahn emphasized that because deterrence can fail despite best efforts to deter, the United States requires some defensive capabilities to reduce the level of possible catastrophe in the event of war. This is a matter of national prudence. As Kahn noted, "War can still occur and it is better to survive the war than not. Therefore one needs to have systems that can reduce the damage done in a war."[91] He emphasized that "so long as" strategic defensive capabilities are "technologically and economically possible," they are needed for both extended deterrence credibility and "because it is prudent to take out insurance against a war's occurring unintentionally."[92] Kahn's colleague, Donald Brennan, elaborated the case for strategic defense of the United States during the Cold War, arguing that it was "bizarre" that US policy should prioritize offensive threats to Russians over defensive capabilities for Americans.[93] Kahn, Brennan, and Gray were mindful of the technical and financial challenges confronting these strategic defense goals. They contended, nevertheless, that the level of defenses needed for extended deterrence *credibility* was likely available and that a meaningful level of strategic defense for society could be available with reasonable investment, time, and policy attention.

US policy has moved incrementally and on a bipartisan basis over the past several decades toward acceptance of strategic ballistic missile defense capabilities for the United States *against limited* missile threats and theater missile defenses for allies and friends.[94] This move in policy in favor of defensive capabilities for the United States against limited missile threats appears to have been motivated by the emerging and unprecedented nuclear threats from "rogue" states. Remarked President George W. Bush, "In such a world, Cold War deterrence is no longer enough. To maintain peace, to protect our own citizens and our own allies and friends . . . we need a new framework that allows us to build missile defenses to counter the different threats of today's world."[95] This shift also reflects an apparent general belief that US missile defense capabilities are technically feasible/affordable against a "rogue" state's *limited* strategic missile threat—as op-

posed to the contemporary prospects for defending against the large-scale missile threats posed by Russia and China.[96]

However, the acceptance of the need for physical protection of the United States remains within some familiar easy deterrence–oriented limits. The Obama administration's unclassified 2013 *Report on Nuclear Employment Strategy of the United States* says that "the United States seeks to improve strategic stability by demonstrating that it is not our intent to negate Russia's strategic nuclear deterrent."[97] The Trump administration's 2019 *Missile Defense Review* says, "The United States *relies on nuclear deterrence* to prevent potential Russian or Chinese nuclear attacks employing their large and technically sophisticated intercontinental missile systems" (emphasis added).[98] US security against the large-scale nuclear threats posed by these great powers appears to be based *on deterrence*, *not* on capabilities for physical protection against such a nuclear attack. Contemporary US policy in this regard corresponds to Kahn's call for strategic defenses for the United States against limited nuclear missile threats, but *not* against large-scale Russian or Chinese nuclear threats.

Implications for the Evolving Nuclear Policy Debate

Nuclear policy debates tend to focus on some specific nuclear system, such as an ICBM or cruise missile. Advocates and opponents offer contrasting claims that the system in question surely is needed for deterrence or certainly is excessive and "destabilizing." These conflicting claims generally are extensions of the difficult or easy deterrence narratives' different expectations about the functioning of deterrence and speculation about the future contexts and types of opponents against which deterrence is expected to operate. They often reflect one deterrence narrative or the other as the basis for judgement because these narratives' differing expectations of context and character of the opponent determine how "stability" is defined and how deterrence requirements are calculated. However, rarely do the discussions focus on the narratives behind the competing arguments that a nuclear system is essential or excessive for deterrence—it is much easier simply to assert that "it's destabilizing" or "it's essential" as if there is a known, objective basis for making such statements. There is not; there are competing narratives based on differing speculative expectations about the future.

If the easy deterrence narrative's answers to the key questions about context and opponent are considered the more valid, then the associated easy deterrence force posture may be considered the most reasonable. If the difficult deterrence narrative's answers to these questions are considered

the more valid, then its more demanding associated force posture requirements may be deemed the most reasonable. Each narrative is the most reasonable *if judged from the prism of its own respective assumptions and logical framework*. However, each is woefully wrongheaded if judged by the other's: the easy deterrence force requirements are wholly inadequate if judged by the standards of the difficult deterrence narrative; the difficult deterrence requirements, in turn, are excessive and destabilizing if judged by the easy deterrence metrics. Both such judgements, however, are largely speculative by definition because they concern the future functioning of deterrence and its requirements—which will be shaped by currently unknown details of context and opponent.

In short, definitive contemporary claims about what will or will not be required for deterrence or what is excessive should be recognized for what they are—based on the different speculative assumptions about the future context and character of opponents underlying the competing narratives. This speculation cannot be resolved with great confidence by better methodologies or sharper analyses. There simply are too many inherent unknowns regarding the many possible factors that can affect the future functioning of nuclear deterrence and its requirements. As Kahn emphasized, there is a fortunate absence of empirical data regarding the outbreak of nuclear war on which to base definitive conclusions about the functioning of nuclear deterrence.[99] Gray provides the implications of this point: "It is all but self-evident that there can be no objectively correct answer. None of the candidate answers are testable, save by the verdict of future events."[100]

Which narrative ultimately is the more accurate depends then on which will prove to have captured the character of future contexts and opponents more precisely: Will the potentially diverse opposing leaderships' goals, perceptions, values, and modes of decision-making render opponent behavior unpredictable, even seemingly irrational to American observers—thereby making the functioning of deterrence particularly challenging—as is anticipated by the difficult deterrence narrative? Or, as anticipated by the easy deterrence narrative, will the contexts and character of opponents render punitive societal threats and the uncertainty of their execution adequate for deterrence to function predictably and reliably? If so, the easy deterrence force recommendations should correspondingly be adequate and the potentially "destabilizing" effects of physical defenses may indeed be of greater concern than the value of whatever level of added deterrence credibility and protection they might provide. As noted, however, the answers to these questions about the future cannot be known with precision

and confidence in the present given the irreducible uncertainties pertinent to the functioning of deterrence. Looking back from the twenty-second century, it may be clear via the unraveling of history that one of these narratives or the other offered a more accurate basis for deterrence policy in the mid to late twenty-first century, but that simply cannot be known with confidence in the present.

This lack of certainty does not fit well with the political demands for policy planners to identify with great confidence the specific effects of moving in one direction or another, but it may be the best that honestly is possible. Nevertheless, as Kahn observed, in the absence of the cooperative global transformation enabling nuclear disarmament, preparation for deterrence must go forward. Abdicating in frustration because it is impossible to predict with certainty the composition of "stability," the precise requirements for deterrence, or a policy direction that *ensures* the functioning of deterrence would be to consciously leave all to chance and luck—a notoriously bad strategy. Instead, policy guidance for deterrence must be as informed as possible with full recognition that the unavoidable uncertainties about the future preclude credible claims that one narrative or the other is "objectively correct." With this significant caveat about deterrence and nuclear policy, here is the key question: Given what may be anticipated about future contexts and opponents, is it possible to suggest whether the easy or difficult deterrence narrative offers an approach to deterrence policy *that is more prudent*? As Hans Morgenthau emphasized, in matters of national security, prudence should be the priority consideration given the stakes involved.[101]

Where Is Prudence?

Which deterrence narrative ultimately is the more prudent depends on which seems to be more suitable to deter war in the future given the limited information available in the present. No "objectively correct answer" is possible, but informed commentary is. For example, it is possible to observe that evidence from history and contemporary studies of cognition suggest strongly that opponents are likely to have a diversity of goals, perceptions, values, and modes of decision-making—some known to outsiders, others unknown. Leaderships have frequently pursued surprising goals and risked national security in ways that observers, including those in the United States, considered highly unlikely and even irrational at the time.[102] The easy deterrence narrative may be correct in its expectation that uncertain punitive nuclear threats will enforce caution in all rational or sensible opponents and thus have the needed deterrent effect;

it is impossible to claim otherwise with certainty. But the driving concern of the difficult deterrence narrative is reasonable: at least some future opponents' decision-making and behavior may be contrary to easy deterrence expectations—*as has been the case in the past*—and affect the functioning of deterrence in unexpected ways—again, as it has in the past. This expectation is buttressed by the expanding number and diversity of nuclear threats to the United States—including from revisionist, expansionist states and states with leaderships that are unfamiliar and/or highly eccentric by familiar Western norms. This dynamic appears to increase the probability that the decision-making of diverse opponents will be varied and shape the functioning of deterrence in surprising directions.

Given the potential stakes at risk, it would seem that the burden of proof is on the easy deterrence narrative to explain why those responsible for US national security should now be confident that contexts and opponents will predictably and reliability fit the comforting profile it posits of opponents who will, when necessary, be deterred by uncertain US societal threats, that is, easy deterrence. But this explanation must acknowledge that the character, behavior, and apparent calculations of some pertinent past leaderships have been well outside that particular profile for a variety of reasons. Why can it now be expected with confidence that the potential variability in opponents' future decision-making will *not* lead them to surprising—even *apparently* irrational—behavior? What new factors in state behavior now point with high confidence to opponents whose decision-making and behavior can be relied upon to prove so predictably sensible and cautious that deterrence can be expected to work easily? Easy deterrence proponents are welcome to present the evidence and logic behind confidence in this expectation. The implications for US deterrence policy and considerations of proliferation are profound if it is deemed most plausible.[103]

There is, however, a challenge in doing so. Assurances of predictably prudent opponent behavior seem open to serious question given the reality of eccentric, occasionally reckless behavior of some US opponents—now including those with growing or potential nuclear capabilities, such as North Korea and Iran—and also given the significant role nuclear weapons appear to play in Russia's and China's respective efforts to recover or expand their places "in the sun." Even during the Cold War, Herman Kahn acknowledged that it would be "reckless" for an opponent to strike the United States with nuclear weapons, but "even more reckless" for the United States to rely on an opponent's "extreme caution and responsibility" for security à la easy deterrence.[104]

The difficult deterrence–recommended diverse and flexible nuclear threat options and planning *cannot* ensure the functioning of deterrence in every possible contingency—as its contributors fully acknowledge. Nothing can "ensure" deterrence because opponents ultimately decide to be deterred or not. But a broader and more flexible range of threat options may help expand the parameters for deterrence to apply to opponents who require more than an uncertain threat of societal destruction to be deterred. Having a spectrum of deterrence threat options and focusing on threat credibility seem only prudent in the contemporary threat environment given the diversity of opponents and their nuclear threats, the potential variability of their decision-making, and the range of possible deterrence goals.

In addition, it seems particularly imprudent to place so much confidence in the reliable functioning of deterrence that little or no provision is made for physically defending against even the *limited* nuclear attacks that might be mounted or threatened by a North Korea, Iran, or other new nuclear power in the future. Here again, the difficult deterrence narrative seems the more prudent; easy deterrence offers no provision for the reduction of damage if deterrence and intra-war deterrence fail—its focus is on societal threats and argues *against* strategic defenses. In short, Waltz's easy deterrence–oriented rhetorical question of why anyone should want to replace stable deterrence with unstable defense now seems easily answered.[105] In the contemporary security environment, some strategic defense capabilities may be wanted not to replace deterrence, but to help protect society in those limited attack scenarios in which the reliable functioning of deterrence is suspect and strategic defenses offer some potential for meaningfully limiting the consequences of deterrence failure. To be sure, there is room to debate the prospective threats that fit this category, but some almost certainly do.

The prudence of difficult deterrence also includes a continuing role for diplomacy—Kahn emphasized the possible value of negotiations and agreements that are in the mutual security interests of the United States and opponents.[106] The realist's caveat to this point, however, is to recognize that most, perhaps all, opponents are likely to pursue diplomacy, including arms control negotiations, to advance their own self-interested strategic goals—not as a selfless act for the greater global good or to advance an easy American concept of nuclear deterrence stability. Expectations otherwise are likely to be frustrated.

Conclusion

The easy deterrence narrative is encouraging, even comforting, when compared to the alternative narrative that deterrence is difficult, requires tailoring to specific opponents, and takes constant effort to sustain—and even with that, the prospects for war still cannot be dismissed. The difficult deterrence narrative hardly offers a definitive, satisfying, or in any way comforting prescription, and it sits atop the foreboding realist perspective that the international system is anarchic; it seems designed to lack appeal.

Nevertheless, the Obama administration's commitment to rebuild the US strategic nuclear triad of bombers, ICBMs, and sea-launched missiles—after decades of relative US inactivity—appears to reflect some basic themes of the difficult deterrence narrative and correspondingly the need for US nuclear modernization.[107] Proponents of sustaining the triad largely follow the difficult deterrence narrative's definition of requirements and emphasize the need for the *flexibility and credibility* provided by the triad. In contrast, critics of the modernization program charged that the Obama administration had "lost focus and momentum" and was pursuing "excessive strategic capabilities."[108] They opined that "it is past time for the Obama administration to take a hard look at where the US nuclear arsenal is heading. . . . It is time to change course."[109] Yet the Obama administration sustained its triad modernization plans, and the same underlying difficult deterrence themes appear to be reflected in the Trump administration's continuation of the US rebuilding program.

In short, contemporary positions for or against the nuclear triad modernization program tend to follow one narrative or the other as the basis for their respective arguments. Largely following the easy deterrence narrative, critics make confident claims about force requirements for deterrence and in doing so tend to see no deterrence need for, and considerable potential instability in, comprehensive plans to modernize the strategic triad. In line with the difficult deterrence narrative, proponents conclude otherwise.[110] Today's competing arguments are largely unintelligible in the absence of an understanding of the divergent easy and difficult narratives because the contending arguments about what is or is not "stabilizing" or required for deterrence are extensions of their different assumptions and projections. As the triad modernization debate picks up, a narrative roadmap truly is needed to understand the meaning behind the competing arguments.

The easy deterrence narrative is comforting and convenient in many ways while the difficult deterrence narrative can only be described as discomforting, even jarring. Herman Kahn acknowledged that his difficult deterrence prescription lacked popular appeal: "This is a difficult, unpleas-

ant, and emotional subject: the points raised are often irritating or dismaying, and many readers transfer their irritation and dismay to the author."[111] Nevertheless, Kahn insisted that facts and logic did not allow him an alternative, more attractive course. When criticized by a congressman during congressional testimony for "putting this cold war logic of nuclear war and overkill for two peoples in such remorseless terms," Kahn replied, "Do you prefer a warm human error, a nice emotional mistake?"[112]

There is little chance that the difficult deterrence narrative will be described as comforting or appealing. Difficult deterrence offers neither a definitive solution to the threat of nuclear use nor ease. It offers no cooperative global transformation and disarmament or confidence that deterrence will work easily and predictably across time. Instead, the difficult deterrence narrative confronts a dilemma that the easy deterrence narrative avoids: the contention that deterrence is necessary because cooperative global disarmament is unlikely, but also difficult and fallible because leadership decision-making is variable and unpredictable. This is a profound dilemma. In response, difficult deterrence offers the following for the foreseeable future: (1) tailoring deterrence to be as effective as possible; (2) diplomacy to ease friction when possible; and (3) if deterrence fails, mitigating to the extent possible the humanitarian consequences with intra-war deterrence and feasible defensive preparations. This is a troubling prescription in many ways. It includes no promised ease or certain happy ending, but it is critical to understand if its basic points are the most plausible. If so, this narrative that appears least appealing is, nevertheless, also the most prudent. SSQ

Keith B. Payne

The author is a cofounder of the National Institute for Public Policy and professor emeritus, Missouri State University. He has served as a senior advisor to the Office of the Secretary of Defense and as deputy assistant secretary of defense and as a member of the congressional commission on US strategic posture. He is an award-winning author of numerous books and articles. Dr. Payne received a bachelor of arts degree (honors) in political science from UC Berkeley and a doctorate (with distinction) from the University of Southern California. The author would like to thank Colin Gray, Stephen Cimbala, Matthew Costlow, Michaela Dodge, Eric Edelman, Kurt Guthe, Robert Joseph, Thomas Mahnken, Richard Mies, David Trachtenberg, and the anonymous reviewers for their helpful comments on the manuscript. The views expressed in this article are the author's alone and do not represent any institution with which he is or has been affiliated.

Notes

1. For example, Waltz indicates that the "second-strike" strategic nuclear capability needed for this stable nuclear "balance of terror" is easily acquired and maintained: "It does not take much to deter. . . . In the nuclear business deterrence is cheap and easy." Waltz in Scott D. Sagan and Kenneth N. Waltz, *The Spread of Nuclear Weapons: A Debate Renewed; With New Sections on India and Pakistan, Terrorism, and Missile Defense* (New York: Norton

and Co., 2003), 142, 152. Waltz also notes that "the invulnerability of a sufficient number of warheads is easy to achieve. . . . With nuclear weapons, [deterrence] stability and peace rest on easy calculation of what one country can do to another. Kenneth N. Waltz, "Nuclear Myths and Political Realities," *The American Political Science Review* 84, no. 3 (September 1990): 732, 734. Further, Jervis states, "The healthy fear of devastation, which cannot be exorcised short of the attainment of a first-strike capability, makes deterrence relatively easy." Robert Jervis, "Why Nuclear Superiority Doesn't Matter," *Political Science Quarterly* 94, no. 4 (Winter 1979/1980): 617–18, https://www.jstor.org/.

2. Kenneth Waltz contends that rational calculations are *not* necessary for nuclear deterrence to function, "only a little common sense." See Waltz in Sagan and Waltz, *Spread of Nuclear Weapons,* 154.

3. Thomas C. Schelling, *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1960), 251.

4. These are realist "bedrock assumptions." See John J. Mearsheimer, *The Tragedy of Great Power Politics* (New York: W. W. Norton & Co., 2001), 30–31.

5. Waltz in Sagan and Waltz, *Spread of Nuclear Weapons,* 154.

6. Kenneth Waltz, "The Spread of Nuclear Weapons: More May be Better," *Adelphi Papers*, no. 171 (London: International Institute for Strategic Studies, 1981), https://www.mtholyoke.edu/.

7. James Wood Forsyth Jr., "The Common Sense of Small Nuclear Arsenals," *Strategic Studies Quarterly* 6, no. 2 (Summer 2012): 94, https://www.airuniversity.af.edu/.

8. Albert Carnesale et al., *Living with Nuclear Weapons* (New York: Bantam Books, 1983), 44. See also Gordon Graig and Alexander George, *Force and Statecraft: Diplomatic Problems of Our Time* (New York: Oxford University Press, 1995), 225.

9. Waltz in Sagan and Waltz, *Spread of Nuclear Weapons,* 114.

10. Waltz, "More May Be Better."

11. McGeorge Bundy, *Danger and Survival: Choices about the Bomb in the First Fifty Years* (New York: Random House, 1988), 592.

12. Waltz in Sagan and Waltz, *Spread of Nuclear Weapons,* 117, 132.

13. Waltz, "Nuclear Myths and Political Realities," 737–38.

14. McGeorge Bundy, "To Cap the Volcano," *Foreign Affairs* 48, no. 1 (October 1969): 10, https://www.jstor.org/.

15. Robert Jervis, "The Political Effects of Nuclear Weapons: A Comment," *International Security* 13, no. 2 (Fall 1988): 81.

16. McGeorge Bundy, "Bishops and the Bomb," *New York Review of Books* 30, no. 10 (16 June 1983): 3–4. See also Bernard Brodie, *War and Politics* (New York: Macmillan Publishing, 1973), 412.

17. Waltz, "Nuclear Myths and Political Realities," 738.

18. Graham Allison, Albert Carnesale, and Joseph Nye, "Conclusion," in *Fateful Visions: Avoiding Nuclear Catastrophe*, eds. Joseph Nye, Graham Allison, and Albert Carnesale (Cambridge, MA: Ballinger Publishing, 1988), 216.

19. Bundy, "To Cap the Volcano," 9.

20. See, for example, Herbert York, *Race to Oblivion: A Participant's View of the Arms Race* (New York: Simon and Schuster, 1970), 42, 46–48.

21. Ralph Lapp, *Kill and Overkill: The Strategy of Annihilation* (New York: Basic Books, 1962), 140.

22. Jervis, "Why Nuclear Superiority Doesn't Matter," 618. Schelling further asserts, "We should find virtue in the technological discoveries that enhance the anti-population potency of our retaliatory weapons." Schelling, *Strategy of Conflict*, 239, 233, 236–37. See also Thomas C. Schelling, "Reciprocal Measures for Arms Stabilization," in *Arms Control, Disarmament and National Security*, ed. Donald G. Brennan (New York: George Braziller, 1961), 167; York, *Race to Oblivion*, 42, 46–48; McGeorge Bundy, William J. Crowe, Jr., and Sidney D. Drell, *Reducing Nuclear Danger: The Road Away from The Brink* (New York: Council on Foreign Relations Press, 1993), 95; Bernard Brodie, "The Anatomy of Deterrence," *World Politics* 11, no. 2 (January 1959): 177; and Jerome H. Kahan, *Security in the Nuclear Age: Developing U.S. Strategic Arms Policy* (Washington, D.C.: The Brookings Institution, 1975), 330.

23. Bruce Blair et al., "Smaller and Safer: A New Plan for Nuclear Postures," *Foreign Affairs* 89, no. 5 (September/October 2010): 10, https://www.jstor.org/.

24. Waltz, "More May Be Better."

25. As noted in Steven Pifer and Michael E. O'Hanlon, *The Opportunity: Next Steps in Reducing Nuclear Arms* (Washington, D.C.: Brookings Institution Press, 2012), 20–21.

26. Robert McNamara to Lyndon Johnson, draft memorandum, subject: "Recommended FY 1966–FY 1970 Programs for Strategic Offensive Forces, Continental Air and Missile Defense Forces, and Civil Defense," 3 December 1964, Department of Defense/Office of the Secretary of Defense, Freedom of Information Act, 4. Hereafter, McNamara to Johnson, draft memorandum, 1964.

27. See Alain C. Enthoven and K. Wayne Smith, *How Much Is Enough? Shaping the Defense Program, 1961–1969* (New York: Harper and Row, 1971), 67, 207–8; and McNamara to Johnson, draft memorandum, 17.

28. McNamara to Johnson, draft memorandum, 4.

29. Edward J. Drea, *McNamara, Clifford, and the Burdens of Vietnam, 1965–1969,* vol. 6, Secretaries of Defense Historical Series (Washington, D.C.: Historical Office, Office of the Secretary of Defense, 2011), 347, https://history.defense.gov/.

30. See a listing of such recommendations in Keith B. Payne and James R. Schlesinger, *Minimum Deterrence: Examining the Evidence* (Fairfax, VA: National Institute Press, 2013), 4–5, https://www.nipp.org/.

31. York, *Race to Oblivion,* 167–68.

32. Union of Concerned Scientists, *The Obama Administration's New Nuclear Policy: An Assessment of the "Nuclear Posture Review"* (Washington, D.C.: Union of Concerned Scientists, 8 April 2010).

33. James Wood Forsyth Jr., "Nuclear Weapons and Political Behavior," *Strategic Studies Quarterly* 11, no. 3 (Fall 2017): 122, https://www.airuniversity.af.edu/.

34. Steve Fetter, "Nuclear Strategy and Targeting Doctrine," in *The Nuclear Turning Point: A Blueprint for Deep Cuts and De-alerting of Nuclear Weapons,* ed. Harold A. Feiveson (Washington, D.C.: Brookings Institution Press, 1999), 57.

35. Gareth Evans and Yoriko Kawaguchi, *Eliminating Nuclear Threats: A Practical Agenda for Global Policymakers*, Report of the International Commission on Nuclear Non-proliferation and Disarmament (Canberra: International Commission on Nuclear Non-proliferation and Disarmament, 2009), 194, https://ycsg.yale.edu/.

36. Waltz in Sagan and Waltz, *Spread of Nuclear Weapons,* 109.

37. Thomas C. Schelling, *Arms and Influence* (New Haven, CT: Yale University Press, 1966), 102–12; Thomas Schelling, "Comment," in *Limited Strategic War*, eds. Klaus

Knorr and Thornton Read (New York: Frederick A. Praeger, 1962), 248–50; Thomas Schelling, "Surprise Attack and Disarmament," in *NATO and American Security*, ed. Klaus Knorr (Princeton, NJ: Princeton University Press, 1959), 207; Waltz, "Nuclear Myths and Political Realities," 733–34; and Jervis, "Why Nuclear Superiority Doesn't Matter," 618, 628–29.

38. In 1974, Secretary of Defense James Schlesinger observed that all previous planned US options involved "literally thousands of weapons" and emphasized the need for limited nuclear response options. He announced publicly that the United States would introduce limited nuclear threat options to provide greater credibility for the deterrence of limited threats. He said this was made necessary because increased Soviet nuclear capabilities had rendered the credibility of large-scale US response options to limited attacks "close to zero." James Schlesinger, *US/USSR Strategic Policies*, Testimony in US Senate, Committee on Foreign Relations, 93rd Cong., 2nd Sess., 4 March 1974, 9, and 7, 12–13, 55. See also James Schlesinger, *Annual Defense Department Report, FY 1976 and FY 197T* (Washington, D.C.: USGPO, 5 February 1975), II-3–II-4, https://history.defense.gov/; and James Schlesinger, *Annual Defense Department Report, FY 1975* (Washington, D.C.: US Government Printing Office, 4 March 1974), 32, 42, https://history.defense.gov/.

39. Schlesinger, *Annual Defense Department Report, FY 1975,* 38.

40. Secretary of Defense Schlesinger emphasized that capability for limited nuclear response options required no increase in the number of nuclear weapons, but did require the planning and control necessary for limited strike options. See Schlesinger, *US/USS.R. Strategic Policies*, 17; and Schlesinger, *Annual Defense Department Report, FY 1975*, 44.

41. Waltz in Sagan and Waltz, *Spread of Nuclear Weapons,* 24.

42. See in particular Schelling, *Strategy of Conflict*, chap. 8.

43. Schelling, *Arms and Influence,* 99.

44. Waltz, "More May Be Better." Robert Jervis makes the same claim: "Even a very small probability of escalation is sufficient to deter serious encroachments." Jervis, "Why Nuclear Superiority Doesn't Matter," 619.

45. In 1979 Henry Kissinger remarked publicly that "our European allies should not keep asking us to multiply strategic assurances that we cannot possibly mean, or if we do mean, we should not want to execute, because if we execute, we risk the destruction of civilization." Henry Kissinger, "The Future of NATO," in *NATO, The Next Thirty Years*, ed. Kenneth Myers (Boulder, CO: Westview Press, 1981), 8.

46. Thomas Schelling elaborated this proposition that uncertainty deters in some detail. See the discussion in Keith Payne, *The Great American Gamble* (Fairfax, VA: National Institute Press, 2008), 31–35.

47. Kenneth Waltz, *Theory of International Politics* (Boston: McGraw Hill, 1979), 175.

48. Jervis, "Why Nuclear Superiority Doesn't Matter," 633. See also Brodie, *War and Politics*, 403, 430.

49. Waltz in Sagan and Waltz, *Spread of Nuclear Weapons*, 126.

50. Waltz in Scott Sagan, Kenneth Waltz, and Richard Betts, "A Nuclear Iran: Promoting Stability or Courting Disaster?" *Journal of International Affairs* 60, no. 2 (Spring/Summer 2007): 137, https://www.jstor.org/. (Transcript of a debate between Waltz and Sagan, Kellogg Conference Center, Columbia University School of International and Public Affairs, 8 February 2007.)

51. Bernard Brodie, *War and Politics* (New York: Macmillan, 1973), 403, 430.

52. Waltz, *Theory of International Politics, 176.*

53. Roger B. Myerson, *Force and Restraint in Strategic Deterrence: A Game-Theorist's Perspective* (Carlisle, PA: US Army War College, Strategic Studies Institute, November 2007), 22.

54. Waltz, "Nuclear Myths and Political Realities," 737.

55. Waltz, 744.

56. Waltz, "More May Be Better."

57. Waltz; see also Schelling, *Strategy of Conflict*, 253.

58. Schelling, *Strategy of Conflict*, 233.

59. Schelling, 207–29.

60. Schelling, *Arms and Influence*, 285.

61. Waltz, *Spread of Nuclear Weapons*, 153.

62. Waltz, "Nuclear Myths and Political Realities," 743.

63. Hans M. Kristensen, Robert S. Norris, and Ivan Oelrich, *From Counterforce to Minimal Deterrence: A New Nuclear Policy on the Path toward Eliminating Nuclear Weapons*, Occasional Paper no. 7 (Washington, D.C.: Federation of American Scientists and the National Resources Defense Council, April 2009), 31–32, https://fas.org/.

64. Secretary of Defense Harold Brown made this point with remarkable clarity in his fiscal year 1980 annual Department of Defense report: "*In the interests of [deterrence] stability, we avoid the capability of eliminating the other side's deterrent,* insofar as we might be able to do so. In short, we must be quite willing—as we have been for some time—to accept the principle of mutual deterrence, and design our defense posture in light of that principle" (emphasis added). Harold Brown, *Department of Defense Annual Report, Fiscal Year 1980* (Washington, D.C.: US Government Printing Office, 1979), 61, https://history.defense.gov/. More recently, see Department of Defense, *Ballistic Missile Defense Review Report* (Washington, D.C.: Secretary of Defense, February 2010), 12–13, https://apps.dtic.mil/; and Eric Gomez, "It Can Get You into Trouble, but It Can't Get You Out," in *America's Nuclear Crossroads: A Forward-Looking Anthology*, eds. Caroline Dorminey and Eric Gomez (Washington, D.C.: CATO Institute, 2019), 17–28.

65. Albert Wohlstetter, *The Delicate Balance of Terror*, doc. P-1472 (Santa Monica, CA: RAND Corporation, 1958), https://www.rand.org/.

66. Bernard Brodie, "The Development of Nuclear Strategy," *International Security* 2, no. 4 (Spring 1978): 69, https://www.jstor.org/.

67. Gordon A. Craig and Alexander L. George, *Force and Statecraft: Diplomatic Problems of Our Time*, 3rd ed. (New York: Oxford University Press, 1995), 188.

68. Jonathan Roberts, *Decision-Making during International Crises* (New York: St. Martin's Press, 1988), 162–63.

69. Colin S. Gray, *Strategy and Defence Planning: Meeting the Challenge of Uncertainty* (London: Oxford University Press, 2014), 1.

70. Colin Gray, *Nuclear Strategy and Strategic Planning* (Philadelphia: Foreign Policy Research Institute, 1984), 47.

71. Craig and George, *Force and Statecraft*, 192.

72. For a thorough discussion of this final organizational factor interfering with the predictable functioning of deterrence, see Scott Sagan in Sagan and Waltz, *Spread of Nuclear Weapons*, 46–87.

73. National Research Council, *U.S. Air Force Strategic Deterrence Analytic Capabilities: An Assessment of Tools, Methods, and Approaches for the 21st Century Security Environment* (Washington, D.C.: National Academies Press, 2014), 35–39, 93–95, https://doi.org/.

74. For a recent concise summary of this point and some of its implications for deterrence, see Iain King, "What Do Cognitive Biases Mean for Deterrence?," *The Strategy Bridge*, 12 February 2019, https://thestrategybridge.org/. See also Thomas Scheber, "Evolutionary Psychology: Cognitive Function, and Deterrence," in *Understanding Deterrence*, ed. Keith B. Payne (New York: Routledge, 2013), 65–92.

75. See Keith Payne, *Deterrence in the Second Nuclear Age* (Lexington: University Press of Kentucky, 1996), 79–119; and Keith Payne, *The Fallacies of Cold War Deterrence and a New Direction* (Lexington: University Press of Kentucky, 2001), 1–77.

76. Richard Ned Lebow, "The Deterrence Deadlock: Is There a Way Out?" in Robert Jervis, Richard Ned Lebow, and Janice Gross Stein, *Psychology and Deterrence* (Baltimore: Johns Hopkins University Press, 1985), 183.

77. Lawrence Freedman, *Deterrence* (Malden, MA: Polity Press, 2004), 117.

78. Quoted in Aleksander Fursenko and Timothy Naftali, *One Hell of a Gamble: Khrushchev, Castro, and Kennedy, 1958–1964* (New York: W. W. Norton and Company, 1997), 171, 241.

79. Viktor Semykin, interview for "The Missiles of October: What the World Didn't Know," *ABC News*, Journal Graphics transcript no. ABC-40, 17 October 1992, 21.

80. Avner Cohen, *Israel and the Bomb* (New York: Columbia University Press, 1998), 342. Prof. John Mearsheimer notes, "Egypt and Syria knew that Israel had nuclear weapons in 1973, but nevertheless they launched massive land offensives against Israel." Mearsheimer, *Tragedy of Great Power Politics*, 132.

81. Henry Kissinger, *Years of Upheaval* (Boston: Little, Brown & Co., 1982), 465.

82. Craig and George, *Force and Statecraft*, 181, 191.

83. Colin S. Gray, *Maintaining Effective Deterrence* (Carlisle, PA: Strategic Studies Institute, US Army War College, August 2003), 15, https://apps.dtic.mil/.

84. Office of the Secretary of Defense, *2018 Nuclear Posture Review* (Washington, D.C.: Office of Assistant Secretary of Defense, 2018), 25–27, https://media.defense.gov/. For an early post–Cold War (1996) discussion of the need to "tailor" deterrence, see Payne, *Deterrence in the Second Nuclear Age*, 128–29, 156–57.

85. Herman Kahn, "United States Central War Policy," in *Beyond the Cold War*, ed. Robert A. Goldwin (Chicago: Rand McNally, 1966), 51. See also Herman Kahn, *On Thermonuclear War* (Princeton, NJ: Princeton University Press, 1960), 557.

86. Kahn, *On Thermonuclear War*, 137–38.

87. Kahn, "United States Central War Policy," 46, 51.

88. Kahn, *On Thermonuclear War*, 133. See also Michael Howard, "The Forgotten Dimensions of Strategy," *Foreign Affairs* 57, no. 5 (Summer 1979): 983, https://www.jstor.org/.

89. Kahn, *On Thermonuclear War*, 32.

90. This linkage between some US defenses for society and extended deterrence credibility continues in contemporary policy discussions. See, for example, Jina Kim and John K. Warden, "Limiting North Korea's Coercive Nuclear Leverage," *Survival* 62, no. 1 (February/March 2020): 31, https://doi.org/.

91. Herman Kahn, "Some Comments on Controlled War," in Knorr and Read, *Limited Strategic War,* 64.

92. Herman Kahn, *The Nature and Feasibility of War and Deterrence*, P-1888-RC (Santa Monica, CA: RAND Corporation, 20 January 1960), 39–40, https://www.rand.org/. Colin Gray presented this point concisely: "Nuclear war is possible, and the US government owes it to generations of Americans—past, present, and future—to make

prudent defense preparations to limit damage to domestic American values to the extent feasible in the event of nuclear war." Gray, *Nuclear Strategy and Strategic Planning*, 8.

93. Donald G. Brennan, "The Case for Population Defense," in *Why ABM: Policy Issues in the Missile Defense Controversy*, eds. Johan J. Holst and William Schneider Jr. (New York: Pergamon Press, 1969), 116.

94. Office of the Secretary of Defense, *2019 Missile Defense Review* (Washington, D.C.: Department of Defense, 2019), 25–37, https://media.defense.gov/. The Obama administration's 2010 "Ballistic Missile Defense Review" lists as first priority defending "the homeland against the threat of limited ballistic missile attack." Department of Defense, Office of Public Affairs, "2010 Ballistic Missile Defense Review (BMDR) Fact Sheet," 3 March 2010, 2, https://www.mda.mil/.

95. The White House, President George W. Bush, "Remarks by the President to Students and Faculty at National Defense University," news release, 1 May 2002, https://georgewbush-whitehouse.archives.gov/. See also Department of Defense, *Nuclear Posture Review: Report to the Congress in Response to Sections 1041 (as Amended) and 1042 of the Floyd D. Spence National Defense Authorization Act for Fiscal Year 2001, PL 106-398*, December 2001, iii, 7, https://www.esd.whs.mil/.

96. Numerous US government statements observe that contemporary US strategic missile defense capabilities are effective against the prospect of limited missile attacks against the United States. Gen John Hyten, vice-chairman of the Joint Chiefs of Staff, said recently, "I don't say 100 percent very often. I have 100 percent confidence in those capabilities against North Korea." See Lee Haye-ah, "U.S. Military Leader Expresses Full Confidence in Ability to Defend against New N.K. Missiles," Yonhap News Agency, 18 January 2020, https://en.yna.co.kr/. Also, the 2010 BMDR states, "The United States is currently protected against limited ICBM attacks." Department of Defense, "2010 Ballistic Missile Defense Review (BMDR) Fact Sheet," 2. See also Office of the Secretary of Defense, *2019 Missile Defense Review*, XII; the statement by Adm Harry Harris in Missy Ryan, "North Korean Threat Could Require New Missile Defenses for Western US," *The Washington Post*, 14 February 2018, https://www.washingtonpost.com/; and the statement by Lt Gen Trey Obering in Rowan Scarborough, "U.S. Puts Faith in Missile Defense," *Washington Times*, 24 June 2006, https://www.washingtontimes.com/. As early as 1991, the Arms Control Association (frequently critical of missile defense) observed, "There is little doubt that it is technically possible to protect the United States against a handful of missiles launched by accident, a mad commander, or a Third World country." Arms Control Association, Background Paper, *New Star Wars Plan: Unnecessary Destruction of the ABM Treaty*, February 1991, 1.

97. Department of Defense, *Report on Nuclear Employment Strategy of the United States Specified in Section 491 of 10 U.S.C.* (Washington, DC: Department of Defense, June 2013), 3, https://apps.dtic.mil/.

98. Office of the Secretary of Defense, *2019 Missile Defense Review*, 8.

99. See Payne, *Deterrence in the Second Nuclear Age*, 7n7.

100. Gray, *Strategy and Defence Planning,* 2.

101. Hans J. Morgenthau, *Politics among Nations: The Struggle for Power and Peace* (New York: Alfred A. Knopf, 1962), 10.

102. Described in Keith Payne, *The Fallacies of Cold War Deterrence and a New Direction* (Lexington: University Press of Kentucky, 2001), 1–15.

103. Waltz expressed a logical but countercultural conclusion of his expectations, that is, that nuclear proliferation could entail positive results if it expands the regions made reliably stable by mutual deterrence. Waltz, "Spread of Nuclear Weapons," 168.

104. Kahn, *Nature and Feasibility of War,* 30.

105. Waltz in Sagan and Waltz, *Spread of Nuclear Weapons,* 153.

106. Kahn, *Nature and Feasibility of War*, 43.

107. See Department of Defense, "Remarks by Secretary Carter to Troops at Minot Air Force Base, North Dakota," transcript, 26 September 2016, https://www.defense.gov/; and Brad Roberts, *The Case for U.S. Nuclear Weapons in the 21st Century* (Stanford, CA: Stanford University Press, 2016), 35, 99, 103–4, 192–94, 260–62, 268–71. The Obama administration's rejection of a "no first use" declaratory policy also appears to correspond to the difficult deterrence narrative. See Amy Woolf, *U.S. Nuclear Weapons Policy: Considering "No First Use*," Congressional Research Service (Washington DC: CRS, 1 March 2019), https://fas.org/.

108. Daryl Kimball, "Mr. President, 'Yes, We Can,'" *Arms Control Today*, Arms Control Association, 6 July 2015, http://www.armscontrol.org/.

109. Tom Collina and Will Saetren, "Time to Cut America's Nuclear Triad," *The National Interest*, 17 December 2015, http://nationalinterest.org/.

110. See, for example, the statement by Adm Charles Richard in C. Todd Lopez, "Stratcom Commander: Failing to Replace Nuclear Triad Akin to Disarmament," DOD News, 28 February 2020, https://www.defense.gov/.

111. Herman Kahn, "The Arms Race and Some of Its Hazards," in Brennan, *Arms Control, Disarmament and National Security*, 89–90.

112. Statement of Herman Kahn, Director, Hudson Institute, before the US House of Representatives, Committee on Foreign Affairs, Subcommittee on National Security Policy and Scientific Developments, Hearings, 91st Cong., 1st sess., in House, *Strategy and Science: Toward a National Security Policy for the 1970's* (Washington, D.C.: US Government Printing Office, 1969), 154, e-book, https://books.google.com/.

PERSPECTIVE

Surviving the Quantum Cryptocalypse

Jon R. Lindsay

Abstract

The quantum threat to cybersecurity is an example of a self-denying prophesy: the more credible the threat narrative, the more concerted the effort to counter it. Quantum computing poses a security threat because digital encryption currently depends on the computational difficulty of certain mathematical problems such as factoring large numbers that would be exponentially easier to solve with a quantum computer. Although experimental machines are not yet powerful enough to undermine public encryption, they do demonstrate that quantum computers are able, under some circumstances, to outperform the fastest classical supercomputers. Indeed, the quantum threat is so credible that the scientific community has been working on cryptographic countermeasures that will soon be certified for public use. Research is also well underway on new quantum networks that can enhance cryptographic security. The size of the quantum window of vulnerability depends on relative rates of engineering progress in quantum computing and quantum-safe alternatives, as well as political considerations about how long secrets need to be protected. There are reasons to be cautiously optimistic that countermeasures are maturing faster than the threat. Nevertheless, the quantum threat should be taken seriously, which is precisely why it might never materialize.*

*****

The security of almost every digital application on classified and unclassified networks relies on a small number of cryptographic protocols. The security of key protocols such as Rivest-Shamir-Adleman (RSA) relies on the computational intractability of certain mathematical problems, such as factoring large numbers. Quantum computers might be able to solve these problems exponentially faster. Quantum information science is a fast-developing field at the intersection of

*I am grateful to Ben Garfinkel, David Meyer, and several anonymous reviewers for comments on the technical portions of this paper. Any mistakes are my responsibility.

quantum physics and computer science. It uses counterintuitive concepts from quantum physics that make it possible to perform calculations that are impossible for even the fastest classical supercomputers. In principle, a large-scale, fully functional, universal quantum computer could factor very large numbers in a matter of hours.

The maturation of quantum computing would thus pose a categorical threat to the confidentiality, integrity, and availability of the entire cyber domain.[1] An intelligence adversary with the right kind of machine could potentially break RSA, decrypt classified data, and forge digital signatures. All networks and applications on those networks, public and private, using vulnerable cryptography would be put at risk. Because military operations in all physical environments—land, sea, air, space—rely on many of the same information technologies and networks that power the global economy, a systematic vulnerability in the cyber domain would become a systematic vulnerability in all domains. Classified information could be collected, altered, or deleted. Personal, financial, legal, logistic, and operational data could be manipulated to influence tactical and strategic operations. Malware could be installed at will to enable espionage or disrupt critical infrastructure. Disinformation could be disseminated from the secure accounts of senior officials, heightening the credibility of foreign deception efforts. The authentication codes protecting sensitive equipment and weapons stockpiles could be falsified, facilitating illicit proliferation. Given the ubiquitous importance of cyberspace, the systematic compromise of cybersecurity would be a strategic problem of the first order.

The threat of a spooky quantum vulnerability is easy to exaggerate, which makes it tempting to downplay the threat.[2] Indeed, history is littered with expectations of technological transformation that never came to pass.[3] Threats in theory are often limited by challenges in practice, so the realization of the quantum threat will likely depend on institutional capacity as much as scientific potential.[4] Nevertheless, the quantum threat cannot be dismissed out of hand since the scientific state of the art is advancing rapidly.[5] Recent breakthroughs in the lab have demonstrated that it is possible for experimental quantum machines to perform some calculations faster than classical supercomputers, even as the ability to break RSA is still a long way off. The window of vulnerability to quantum computing has not yet opened, but it is increasingly plausible that it could open in the future. Indeed, it is precisely because scientific progress in quantum computing has made the threat so credible that the cryptographic community has redoubled efforts to field countermeasures.

Prophesies that are both believable and undesirable tend to become self-denying. The more that progress in quantum computing portends a "cryptocalypse," the more likely that scientists and policy makers will take steps to keep this from happening.[6] Cryptographers have already identified alternatives to RSA that rely on different mathematical problems believed to be intractable for both classical and quantum computers. The US National Institute of Standards and Technology (NIST) is currently evaluating, and will soon certify, new standards that can be incorporated into cyber systems. Furthermore, quantum mechanics can also be leveraged to create totally new types of secure data networks. Operational prototypes exist in China, Europe, and North America. Chinese scientific progress in quantum information science has been especially motivating for the US government, and both China and the US have dramatically increased their investment in this area in recent years. One implication of this investment is the liklihood that "quantum safe" offsets will be available and implemented long before anyone is able to field a threatening quantum computer.

In this article I explain why the quantum threat may be a self-denying prophesy. First, I provide a quick overview of the quantum threat to public encryption. Next, I discuss the potential impact of quantum computing on the balance between cyber offense and defense. Third, I review progress in the development of countermeasures to the quantum threat, then offer three scenarios based on different assumptions about engineering progress in quantum technology. Finally, I conclude with some cautious optimism about the prospects for quantum defense over offense.

The Quantum Threat to Public Encryption

The quantum threat emerges at the nexus of cybersecurity, cryptology, and quantum computing. The security of cyberspace depends on the computational difficulty of certain mathematical functions, which turn out to be vulnerable to certain quantum algorithms. The threat of quantum cryptanalysis (code breaking) has also inspired the development of various forms of quantum-safe cryptography (code making) such as classical post-quantum cryptography (PQC) and quantum key distribution (QKD).[7] Table 1 parses out these different technologies. This section focuses on the offensive (cryptanalytic) threat posed by quantum computing, while the defensive (cryptographic) remedies of PQC and QKD are discussed in a later section.

Table 1. Classical and quantum information technologies compared

Cryptologic applications	Classical information technology	Quantum information technology
General applications that rely on cryptography for security	Intelligence, communication, administration, command and control, automation, governance, diplomacy, law enforcement, science, engineering, manufacturing, finance, commerce, advertising, entertainment	Scientific modeling and simulation, quantum sensing and measurement, data storage and search, machine learning and artificial intelligence
Classical cryptography vulnerable to quantum cryptanalysis	Rivest-Shamir-Adleman (RSA) Diffie-Helman (DH) Elliptic Curve Cryptography (ECC) Advanced Encryption Standard (AES) Secure Hash Algorithm (SHA)	Shor's algorithm provides an exponential speedup vs. RSA, DH, and ECC. Grover's algorithm provides a polynomial speedup vs. AES and SHA.
Quantum-safe cryptography	Post-Quantum Cryptography (PQC)	Quantum Key Distribution (QKD)

Pundits often assert that quantum computing is "equivalent to opening a combination lock by trying every possible number and sequence simultaneously" or that it can easily solve hard problems like "the traveling salesman problem."[8] Such descriptions are either wrong or extremely misleading. Quantum computing leverages the counterintuitive phenomena of quantum physics to solve mathematical problems. Whereas a digital bit must be one *or* zero, a quantum bit (qubit) can be a "superposition" of one *and* zero. Multiple qubits can be "entangled" to represent more information than can be represented with separate qubits. It is important to appreciate that quantum computers offer performance improvements only for mathematical problems for which a suitable quantum algorithm has been discovered. Furthermore, physical implementations of quantum computers must be able to run quantum algorithms at scale (i.e., with thousands or millions of qubits) while detecting and correcting errors. Difficult outstanding engineering challenges abound. The online appendix summarizes a few key technical concepts to explain how quantum computing works in principle and why it is difficult to implement in practice.[9]

In principle, quantum computing imperils the security of popular cryptographic protocols like RSA. RSA is an example of an asymmetric protocol, which uses different keys for encryption and decryption.[10] Asymmetric encryption, invented independently by British intelligence and American academics in the 1970s, is invaluable for secure internet communication.[11] It is distinguished from symmetric encryption, which uses the same key for both operations; prominent examples include the famous Enigma machine and modern block ciphers like Advanced Encryption Standard (AES). Distributing the same key throughout a large dispersed

organization has always presented a serious security challenge; for example, the Allies were able to break into Enigma networks when they captured German ships carrying common key material.[12] In asymmetric encryption, by contrast, the so-called public key can be openly revealed to allow other people to send encrypted messages that only the recipient can decrypt by using a secret private key. The private key can also be used to create digital signatures that anyone can verify with the public key.

A critical requirement of asymmetric encryption is that it must be extremely difficult to guess the private key from the public key. Modern RSA works because the public key is based on a very large number (i.e., two to the power of 2048) while the private key is based on its prime factors. With ordinary classical computers, it is easy to multiply two large prime numbers together, but it is exponentially harder to factor the result. A typical desktop computer would need more than six quadrillion years to crack 2048-bit RSA.[13] However, in 1994 Peter Shor discovered a quantum algorithm that can theoretically factor prime numbers (and calculate discrete logarithms) exponentially faster than the fastest known classical methods.[14] If one assumes the existence of a powerful quantum computer, therefore, Shor's algorithm could in principle enable successful cryptanalytic attacks in a matter of hours, an astounding improvement compared to the countless lifetimes required by the fastest classical supercomputers today.[15]

RSA is widely used in implementing public key infrastructure (PKI), which links real-world individuals and organizations to cryptographic keys to facilitate secure communication and digital authentication.[16] Military PKI systems, for example, employ a common access card (CAC) with an embedded chip that stores the keys enabling an authorized user to log on to classified and unclassified networks. PKI underwrites the security of military communications, financial transactions, and intellectual property and the privacy of civil society around the world. Digital signatures produced with RSA certify the authenticity of digital messages and facilitate the installation of software from trusted vendors. Breaking RSA would make it possible to decrypt secure data and install arbitrary code on protected networks.

RSA is not the only protocol that matters in modern cryptosystems. Quantum computing provides only a modest advantage against symmetric ciphers like AES or Secure Hash Algorithms (SHA) using other methods such as Grover's algorithm. Unfortunately, PKI necessarily relies on asymmetric ciphers like RSA, Diffie-Helman (DH), and Elliptic Curve Cryptography (ECC), all of which can be defeated with Shor's algorithm.[17] RSA is the linchpin of most modern implementations of PKI,

and there are no quick fixes short of replacing vulnerable asymmetric protocols with something else. The development of a functional quantum computer able to break RSA, and thereby compromise PKI, would imperil the privacy and authenticity of the entire cyber domain.

Shor's algorithm has been known since 1994, but for many years it seemed like little more than a theoretical curiosity. For all practical purposes, it appeared infeasible to build an actual quantum machine powerful enough to run Shor's algorithm with reliable error correction. Completely eliminating all PKI dependence on vulnerable protocols, moreover, would have required a massive update of government and private sector cryptosystems, or the construction of an entirely new quantum communications infrastructure based on immature technology. These would have been major undertakings, to say the least. The theoretical threat posed by Shor's algorithm thus did not seem like a practical urgency.

This perception changed in the 2010s as academic and corporate labs demonstrated working prototypes. A solid-state machine in 2012 was able to "run a three-qubit compiled version of Shor's algorithm to factor the number 15, and successfully find the prime factors 48% of the time."[18] Since then, quantum computers have factored numbers much larger than 15, but still nothing as large as a 2048-bit RSA key. The most dramatic experimental breakthrough to date occurred in September 2019, when a 53-qubit machine known as Sycamore achieved a milestone known as "quantum supremacy." Sycamore, built by Google and physicists at the University of California, Santa Barbara, ran a quantum algorithm faster than could be simulated by the world's fastest classical supercomputer (the IBM Summit at the Oak Ridge National Laboratory).[19]

To crack RSA with the most efficient known method, a quantum computer must be able to keep 20 million qubits in coherence (i.e., maintaining superposition and entanglement without losing quantum information) for several hours, which is what is required.[20] There is still a long way to go before this will be possible. Prototype machines have been able to maintain fewer than 100 qubits in coherence for short amounts of time. In 2017, IBM maintained 50 qubits in coherence for 90 microseconds.[21] In 2019, Google's Sycamore maintained 53 qubits in coherence for three minutes, a dramatic improvement to be sure but a long way from cracking RSA. Sycamore might be likened to the Wright Flyer: a gross contraption compared to what might come later, yet nonetheless a harbinger of a new technological era.[22] The many unknowns and major engineering challenges ahead make it difficult to hazard a guess about whether a large-

scale quantum computer remains 10 or 100 years away, but anything less seems overly optimistic.

The Offense-Defense Balance in Cyberspace

The ability to break RSA would in principle provide a capable intelligence adversary with a formidable offensive advantage. Yet quantum information science (in particular PQC or QKD) also has the potential to restore the advantage to defense, again in principle. If defensive offsets are not developed in time, however, a dangerous window of vulnerability to quantum attack could open. Windows are important in international relations because political actors are tempted to jump through them.[23] An actor with an uncontested capability to perform quantum cryptanalysis would be tempted to use it to gain intelligence advantages, which might then be parlayed into military or economic advantages.

The race between offensive measures and defensive countermeasures is as old as war itself. Offensive advantage, moreover, is never just an immutable characteristic of weapon systems. The offense- defense balance in any era depends on organizational and geostrategic context, not simply technology.[24] Yet scientific principles and engineering feasibility constrain the strategic and operational art of the possible.[25] Technical trends establish the boundary conditions for any potential window in which offense has the advantage. This window can and does change as actors take the initiative to build new weapons and find new ways to use them.

For example, between the world wars technological trends shaped the offensive potential of bombers and the defensive potential of radar. The Royal Air Force (RAF) worked out an air defense scheme after World War I that relied on acoustic mirrors along the Channel Coast able to detect an aircraft 10 or more miles away.[26] Yet as aircraft speeds increased, acoustic mirrors could no longer provide sufficient warning of incoming bombers in time to launch fighters to intercept them. Technological innovation made the "Channel gap" a pressing strategic problem for the RAF, which was not resolved until the emergence of radar a few years before World War II. Importantly, the exploitation of the technological potential for both strategic bombing and air defense required complementary organizational innovation, an area in which Britain performed well while Germany did not.[27] While offensive advantage can be fleeting, it can still be a very real and consequential factor for strategic competition in the window of time before defensive innovation prevails. The question

is how long it takes for any given threat, or countermeasures to it, to become practically feasible.

Quantum computing has the potential to alter the offense-defense balance in cyberspace, but this is not a simple proposition. The cyber domain is often described as intrinsically offense dominant, but in fact the balance is mutable.[28] The hacker does not always get through, in part because cybersecurity has appreciably improved in recent years.[29] There are many reasons for this development, to include the emergence of a multibillion-dollar information security industry, the increased use of active network monitoring and counterintelligence methods such as threat hunting, and the rise of specialized government agencies focused on cybersecurity and military units such as US Cyber Command. These improvements do not imply that we can simply ignore serious cyber threats, however, as recent episodes like the 2016 Russian influence campaign and 2017 NotPetya attacks make clear. On the contrary, it is precisely because we *do* have to worry about serious cyber threats that we have become better at detecting and defending against them. If cyberspace is a contested domain, it is also contestable.[30] Offense does not categorically hold the advantage.[31]

The contest between offense and defense in cyberspace is dynamic and conducted at many levels. Hidden vulnerabilities and clandestine exploits are the coin of the realm for offensive cyber operations. Attackers have incentives to keep their exploits secret because revelation can prompt the defender to patch or reconfigure systems. Many vulnerabilities in software systems tend to be transitory because they can be quickly patched or mitigated once revealed, yet vulnerabilities at the hardware or protocol layers can take longer to remediate. It can take a while to develop and acquire viable substitutes, and even once available, network dependencies can raise the costs of testing and switching to the new components.[32]

Unlike with many cyber vulnerabilities, unfortunately, mere knowledge of the quantum threat to RSA is not enough to close it. Shor's algorithm has been known for a quarter century, as noted, but not yet mitigated. There is no simple patch available because entirely new cryptosystems are needed. The quantum threat is a striking instance of what cybersecurity professionals call a "class break," a vulnerability that categorically affects an entire class of technology versus just particular targets.[33] Shor's algorithm is about the biggest class break imaginable.

According to one prominent physicist, "If a quantum computer is ever built, much of conventional cryptography will fall apart."[34] As the general council of the National Security Agency (NSA) explains, "The strategic advantage here would be for one country to surreptitiously acquire such a

capability and maintain it for perhaps several years or more. Other countries would not realize that everything from their weapons systems to financial transactions would be vulnerable during that period; and that would include not only current activity but also the historic, encrypted communications collected and retained by the winner in anticipation of this very capability."[35] The former president of a major research university argues that Chinese progress in quantum technology "presents the United States with its new 'Sputnik moment.'. . . Whoever gets this technology first will also be able to cripple traditional defenses and power grids and manipulate the global economy."[36]

Chinese developments thus provide a sense of urgency in these matters. China has named quantum informatics a key plank in its "13th Five-Year Plan" for technology and innovation, and it is building the world's largest quantum laboratory.[37] Even though China has historically struggled to catch up in science, in quantum information technology it has been the first to achieve several important milestones.[38] China launched the first satellite for quantum science, demonstrating the ability to leverage the entanglement of particles—described by Einstein as "spooky action at a distance"—from orbit, an unprecedented distance. China has also built a large-scale experimental quantum network between Beijing and Shanghai. China hopes not only to improve its general economic competitiveness by investing in quantum technology but also to shore up its perceived vulnerability to US cyber operations—highlighted by the Snowden leaks—by developing more secure quantum networks. Chinese strategists have started writing about "quantum hegemony," and the United States is taking note.[39]

It is important to appreciate that quantum networking is a related but distinct category of technology from quantum computing. Both technologies draw on quantum mechanics, but the similarities end there. China's recent achievements in satellite-enabled quantum experiments and its Beijing-Shanghai link are all in the realm of quantum communications rather than computation. Chinese progress in quantum computing has been less impressive, and here North America remains the leader. Quantum computing offers advantages to the offense (cryptanalysis) while quantum communications offers advantages to the defense (cryptography). However, these cryptologic advantages do not map directly onto military advantages. Cryptographic security (defense) is needed to cover plans and preparations for a military offensive, and cryptanalytic achievements (offense) can provide intelligence that helps to strengthen military defenses against surprise attack. Furthermore, both types of quantum

technologies are systemic variables, whereas the offense-defense balance in any given case usually depends more on dyadic factors such as the organizational capacity of rivals.[40]

It is far from clear how well either China or the United States will be able to operationalize quantum technology, even as there are reasons to suspect that the US military and intelligence community may have important relative advantages in this respect.[41] What is clearer is that geopolitical competition has become a major catalyst for both countries to invest in quantum information science. Active political rivalry on the scientific frontier makes the cyber offense-defense balance more important, even as it tends to make it more ambiguous.

Defending against the Quantum Threat

Scientific breakthroughs can give rise to new threats to national security, and scientific research can also produce countermeasures to them.[42] Yet this counteraction does not happen by itself. To realize any effective countermeasure, actors must invest resources and political will. Actors may show little interest in preventative action when a threat is diffuse, far away, or hard to understand. Yet as time horizons shorten and threats begin to seem more palpable, the imperative for preventative action becomes more urgent.[43] The incentives to invest in applied scientific research will also tend to increase when a geopolitical rival invests in the same threatening technology. The quantum threat has long seemed diffuse and uncertain. Yet real achievements by a real competitor like China are helping to dramatize the urgency of the problem. Balancing in politics and balancing in science can become one and the same.[44]

Quantum-safe cryptography, as I use the term here, includes both PQC and QKD. These innovations are inspired, in part, by the threat posed by Shor's algorithm and experimental progress in quantum computers. If offsets can be fielded soon, the quantum threat window may not ever open in the first place.

PQC works by using mathematical problems difficult for both classical and quantum computers to solve (i.e., PQC is not vulnerable to Shor's algorithm). Candidate problems include finding the shortest vector in a lattice, decoding error-correction codes, and solving systems of multivariate equations over finite fields.[45] PQC runs on classical computers, providing security against classical and quantum attacks. Because quantum computers have very specialized applications, classical computers will almost certainly remain the best choice for many applications. Even quantum systems

will still incorporate some classical components. Therefore, PQC will be needed to ensure the security of classical computers in the future.

In the United States, the NIST "has initiated a process to develop and standardize one or more additional public-key cryptographic algorithms . . . that are capable of protecting sensitive government information well into the foreseeable future, including after the advent of quantum computers."[46] The NIST has received, and is evaluating, nearly 70 submissions from two dozen countries.[47] The NSA, meanwhile, has signaled that it "will initiate a transition to quantum resistant algorithms in the not too distant future," cautioning against adopting strong protocols like ECC and instead waiting for PQC.[48] While the NIST should approve PQC alternatives within the next few years, the full transition could still take a decade more. Previous transitions (e.g., to AES) took much longer than anticipated due to economic and organizational constraints. In the ideal case, new PQC protocols would simply be swapped in for current cryptographic primitives to minimize the need to reengineer all the other systems that depend on them. More likely, however, "PQC standardization . . . will need a new wineskin to hold the new wine."[49] So long as classical computing power continues to increase, the additional computational overhead of PQC will probably not pose a general barrier to implementation. However, the greater resource-intensiveness of PQC could pose a problem for more constrained and bandwidth-limited military applications (such as ship-to-shore networks). This problem might be mitigated by judiciously limiting the use of computationally intensive primitives within the overall cryptographic system, just as slower RSA is used to open a session conducted with faster AES today.

The alternative to PQC is QKD. Quantum mechanics can be leveraged to create new kinds of communication networks that use a totally different approach to cryptography. QKD exploits the Heisenberg Uncertainty Principle to detect the presence of an eavesdropper. Since the act of measuring quantum data can change them, an eavesdropper in the channel would increase detectable error rates. QKD thus makes it possible to securely distribute unique keys between geographically separated parties (which was the original justification for inventing asymmetric encryption like RSA).[50] The practical feasibility of QKD over large distances, including between satellites in orbit and ground stations, has been demonstrated in numerous experiments.[51] Research is underway to develop quantum routers and networks that can preserve entangled states while scaling up to greater numbers of users, higher bandwidths, and longer distances, along with reliable quantum repeater and memory devices that do not

destroy quantum state.[52] These challenges are perhaps less formidable than those associated with general-purpose quantum computing, but they are still difficult. Yet there are also promising signs of progress.[53]

QKD is hardly a silver bullet. The same mechanism that prevents the eavesdropper from copying the data (i.e., the act of tapping the quantum circuit causes an increase in random errors) also enables the adversary to impose a service denial attack on the quantum channel. An attempt to copy data every time it is transmitted has the potential to force every connection to reset. QKD also does not protect data integrity against side channel attacks on the engineering implementation of the system or social engineering attacks on the gullibility of human operators. Elaborations such as "measurement-device-independent QKD" can close some loopholes, but they still assume that the preparation of photons for transmission will be unobserved and that communicators will also have an authenticated classical channel.[54] This does not preclude some types of man-in-the-middle attacks.

Any transition to quantum communication networks (with QKD) will also be difficult. Quantum networks rely on very different principles than does the installed base of classical digital networks around the world. If switching to PQC will be hard, QKD could be even harder. Adoption of PQC, insofar as security motivates consideration of quantum networking, will probably be more feasible for most organizations and states. As cryptographer Tom Berson wryly notes, QKD is a "new, difficult, expensive way to achieve an outcome which we have, for decades, been achieving easily and cheaply."[55] For most practical network applications, PQC to shore up classical networks will be available more quickly, feasibly, and reliably without attempting to transition to a wholly new quantum network architecture protected by QKD. Quantum networking may yet become attractive for novel applications other than cryptography that have no classical equivalent, such as certifying deletion or sharing out quantum computational resources.[56]

Assessing the Quantum Window of Vulnerability

It is difficult, even irresponsible, to make specific predictions about progress at the scientific frontier, but it is possible to gain some clarity about the relative bounds of the problem. In particular, it is possible to say something about the size of the technological window of vulnerability based on relative estimates about the maturation of offensive and defensive innovation. Nontechnical considerations also affect the size of the window. Foremost among these is the length of time that secrets need to be kept.

The latent value of secrecy will vary depending on the encrypted data's content and policy priorities. Some secrets are extremely perishable, such as the current location of mobile military assets in war or a negotiating position in a deal that will be concluded in the next few days. By contrast, weapon designs and other capabilities that require significant investment may need longer protection if revelation would enable an adversary to develop countermeasures. Politically sensitive covert action might be kept secret for a long time if revelation would be embarrassing to the government or allies or concerns activities of exceptionally long duration.[57] Intelligence sources and methods are particularly sensitive. Historical data can enable the adversary to better understand an adversary's doctrine or even identify long-running operations. For example, the US Army intercepted a batch of KGB communications about agent operations in the West in the 1940s and was able to decrypt some of them due to improper reuse of one-time pads by KGB agents.[58] The Army decrypted only a small fraction of these messages (known as the Venona files) before the Soviets discovered the compromise and switched to a different system. Nonetheless, the ongoing decryption and analysis of the Venona trove enabled the Allies to uncover the Cambridge Five spy ring (including Guy Burgess and Kim Philby) as well as operations against the Manhattan Project (including Julius and Ethel Rosenberg). Venona continued to illuminate KGB methods and facilitate Western counterintelligence throughout the Cold War.[59]

Figure 1 summarizes three different scenarios based on three successively longer estimates of the time it will take for an attacker to field a fully functional, large-scale quantum computer that can crack RSA.[60] The threat window is bounded on the attacker's side by the rapid development, slightly delayed development, or extremely delayed development of quantum computers, denoted by $t_{qc\text{-}rapid}$, $t_{qc\text{-}delayed}$, and $t_{qc\text{-}extreme\text{-}delay}$. These might be considered as 5, 20, or 50 years from now, respectively, but any specific estimates would be misleading. My focus here rather is on the relative size of the window. The window is bounded on the defender's side by the amount of time it will take the defense to transition to quantum-safe cryptosystems secured by PQC or QKD (denoted $t_{q\text{-}safe}$) and the amount of time that organizations want to keep their secrets from an adversary (up to t_{secret}). The point $t_{q\text{-}safe}$ is the earliest possible point that quantum-safe encryption is technically feasible, even as any organizational implementation will take some additional time. Whether or not a target can implement PQC or QKD properly is a critical factor in any given case, but my focus here is on technological boundary conditions.

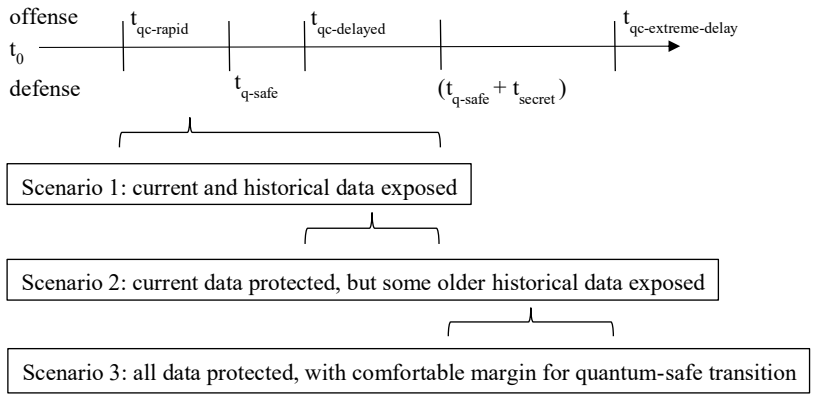


Figure 1. Windows of vulnerability to quantum decryption

Scenario 1 ($t_{qc\text{-}rapid}$) is the best case for offense; scenario 3 ($t_{qc\text{-}extreme\text{-}delay}$) is the best case for defense; and scenario 2 ($t_{qc\text{-}delayed}$) is a mixed case. The first scenario assumes a breakthrough in quantum computing in the next few years, occurring either in public or in secret, that enables an intelligence agency to begin bulk decryption of data secured with contemporary PKI. No quantum-safe offsets are available at the time of this breakthrough (i.e., $t_{q\text{-}rapid} < t_{q\text{-}safe}$ for whatever reason). At that point, most financial transactions, military communications, private personal information, and other data will be exposed. It would still be necessary for the attacker to be able to access, assess, analyze, and disseminate sensitive data, which are all nontrivial organizational performances. If these (difficult) conditions are met, however, then the quantum-enabled attacker could read confidential data, forge digital signatures, and install arbitrary code. Even perishable, time-sensitive, current data would be exposed in the time between a quantum computing breakthrough and the introduction and adoption of viable quantum-safe cryptosystems (i.e., the interval from $t_{q\text{-}rapid}$ to $t_{q\text{-}safe}$). Access to time-sensitive data might even enable an adversary to manipulate markets or disrupt operations. Such an ability could provide intelligence and influence in the short term and erode trust in the global economy in the long term.

Scenario 1 is the worst case for the defender because the quantum computing breakthrough occurs prior to the implementation of quantum-safe cryptography. Even after quantum-safe cryptography is deployed at $t_{q\text{-}safe}$, any data encrypted and stored prior to that date, using old encryption protocols, will still be vulnerable. Any data encrypted prior to $t_{q\text{-}safe}$ in an unsafe protocol may retain some strategic or tactical value for as long as t_{secret} and will thus remain vulnerable to quantum decryption up until $t_{q\text{-}safe}$

+ t_{secret}. After that point, all historical secrets will have lost their intelligence value for understanding military operations or political policy.

In the other two scenarios, defensive innovators are first past the post, allowing current data to be protected from quantum decryption. These cases differ depending on whether any historical data is also exposed. The second-best (or second-worse) case for the defender is scenario 2, where a quantum computing breakthrough is delayed until just after quantum safe implementation (i.e., $t_{q\text{-}safe} < t_{qc\text{-}delayed}$). Scenario 2 is problematic because some old data that were encrypted in the old format will become exposed after the quantum breakthrough, and these will still have some intelligence value to the adversary. All historical data encrypted and stored prior to $t_{q\text{-}safe}$ will become readable to the adversary in the interval from $t_{qc\text{-}delayed}$ to $t_{q\text{-}safe} + t_{secret}$. A proactive intelligence adversary might even begin harvesting encrypted data before the quantum computing breakthrough in anticipation of decrypting them afterwards.

The best case for the defender is scenario 3, where a breakthrough is delayed until long after the quantum-safe transition. In this case, there is nothing valuable left to decrypt after $t_{qc\text{-}extreme\text{-}delay}$. If progress in quantum computing is so delayed, or quantum-safe offsets are available so soon, then no valuable data are exposed. Perhaps the engineering obstacles of entangling millions of fully functional coherent qubits will prove too formidable. For whatever reason, quantum-safe offsets are in place far in advance of the emergence of a powerful quantum computer. When that day finally comes, all data that retain any political or economic utility have long since been encrypted in quantum-safe formats. Any ancient data remaining on servers, still encrypted in unsafe formats, will have long since gone stale (i.e., $t_{q\text{-}safe} + t_{secret} < t_{qc\text{-}extreme\text{-}delay}$). The adversary will thus find no value even in decrypting the old data that it has stockpiled in anticipation of acquiring a quantum computer.

Scenario 3 provides a cushion for the transition to PQC or QKD that is missing in the other two scenarios. This margin (i.e., the interval between $t_{q\text{-}safe} + t_{secret}$ and $t_{qc\text{-}extreme\text{-}delay}$) is important because rolling out the PQC standards that are eventually certified by the NIST is sure to be a long and difficult process. The longer a quantum breakthrough is delayed, or the sooner the quantum-safe offset is available, the more time organizations will have to upgrade their cryptosystems. Those organizations that highly prioritize cybersecurity may be able to upgrade to PQC relatively quickly, once it is available. Many others will delay because of the difficulty of ensuring backward compatibility with their legacy installed base of software. If a quantum computer becomes available during the period of

incomplete transition to PQC, then systems that do not use PQC, or data exchanged with systems that do not use it, will remain vulnerable. In effect this would amount to a localized reversion to scenarios 1 or 2 for some organizations, despite the global availability of PQC per scenario 3. Rather than a discrete point in time, $t_{q\text{-safe}}$ should really be thought of as a fuzzy band that will vary by organization and industry.

In the final analysis, I assess scenario 1 (early quantum computing breakthrough) to be *least* likely while scenario 3 (the triumph of quantum-safe defense) is far *more* likely. Scenario 2 (some historical data exposed to quantum cryptanalysis) deserves to be taken seriously, both because there might be a surprising breakthrough in the midrange and because the quantum-safe transition will be uneven.

How to Stop Worrying and Love the Cryptocalypse

The prospect of a devastating quantum threat to cybersecurity is an example of a self-denying prophesy. The magnitude and credibility of the threat inspires the search for countermeasures to mitigate it. The more convincing the doomsayer's prophesy, the harder its potential victims work to postpone catastrophe.[61] Quantum computing has the potential to create a dramatic "class break" in the computational infrastructure of modern military and economic power. This threat should be taken seriously thanks to recent engineering progress in quantum computing. Indeed, scientists and states are taking it *so* seriously that the most dangerous eventuality is unlikely to come to pass. The US government is taking the quantum threat—and opportunity—particularly seriously because China is betting big on quantum technology.

Self-denying prophesies are common in military history. British prime minister Stanley Baldwin famously said in 1932 that "the bomber will always get through." In 1940, of course, German bombers did not always get through. British fears of strategic bombing, heightened by the RAF's own rhetoric, encouraged the RAF in the interwar years to build the astonishingly successful air defense system that won the Battle of Britain. Likewise, in the eternal race between code makers and code breakers, the looming threat of quantum decryption is already encouraging innovation in quantum-safe encryption. This does not mean that future systems will provide perfect operational security, any more than the RAF's integrated air defense system could intercept every bomber. Baldwin would have been considerably less motivating, however, had he cautioned that the bomber only sometimes gets through, depending on a complex interaction of social and technical factors.

Predicting the interaction of scientific progress, international politics, and secret intelligence is especially difficult. Resolution of the many uncertainties and empirical speculations mentioned in this article will take further assessment of technical progress; and indeed, further technical progress. How much confidence can we have that the quantum threat window will not open? My estimates are informed by current trends, but a future breakthrough is always possible. A well-resourced intelligence agency like the NSA might develop a working quantum computer in secret before the completion of PQC implementation. Documents leaked by Edward Snowden suggest that the NSA has included funding for research into "a cryptologically useful quantum computer" as part of an $80 million research program on "Penetrating Hard Targets."[62] If the NSA were to succeed, is it realistic to believe that its quantum coup could be kept secret? In the 1940s, Bletchley Park secretly developed its Bombe and Colossus machines to break the Enigma and Lorenz cryptosystems, respectively. Britain kept its triumphs secret for decades in order to keep on exploiting Warsaw Pact countries using similar cryptosystems.[63] However, this feat is unlikely to be replicated in the age of quantum computing. The conditions of absolute operational security at Bletchley Park differ starkly from today's world of pervasive leaks and penetrating intelligence. Bletchley Park had a virtual monopoly on the computer scientists of its day (including the brilliant Alan Turing), but the locus of innovation in computer science has long since passed out of government hands. Major firms like Google and IBM are racing to be the first to develop quantum computers for lucrative commercial and scientific applications beyond the national security domain (such as drug discovery and scientific modeling), and there is a cottage industry of reporting on quantum progress in the technical trade press. There is so much investment pouring into commercial and academic quantum science that cryptographers will have plenty of warning well before the quantum threat becomes imminent, an eventuality that remains many years if not many decades away. According to quantum computing expert Scott Aaronson, "It seems improbable that the NSA could be that far ahead of the open world without anybody knowing it."[64]

The PQC transition, by contrast, is already underway and should be well advanced within the next decade. One might reasonably expect PQC to mature sooner and ultimately be more widely implemented than QKD, if only because PQC protocols are designed to be analogous with current cryptographic protocols. Quantum networking technology is perhaps more mature than quantum computing, but, nevertheless, the implementation problems in large-scale quantum communications are legion. It will likely

be PQC rather than QKD—classical rather than quantum protocols—that will provide widespread protection against the threat of quantum cryptanalysis. The widespread implementation of PQC is going to be especially difficult for military systems with widespread dependencies on legacy cryptosystems (and RSA). A thorough survey of military systems will be crucial to ensuring that critical functions and data are prioritized for protection. This transition will inevitably have to be phased, with local upgrades installed and tested in less critical areas to gain confidence in the fixes. This process is sure to be long and complicated, but progress may be expedited if senior leadership gives cybersecurity the priority it deserves.

Quantum cryptanalysis may still be decades away, but some secrets might retain their value for many decades. There are likely things of interest about the early Cold War that remain hidden in the secret archives of intelligence agencies. Given the longevity of some secrets, there is no room for complacency about the quantum threat. Indeed, the entire argument here relies on practitioners *not* being complacent. It is the very plausibility and danger of the threat that mobilize scientific and institutional action. The prospect of quantum decryption sometime in the next few decades is sufficiently likely, and the risks of relying on vulnerable protocols like RSA for cryptographic security are sufficiently great, that effort to develop and implement quantum-safe networks should be a high priority.

Current US government interest in quantum information science is encouraging in this regard. As of this writing, the Trump administration's fiscal year 2021 budget request features generous funding for "industries of the future" like "artificial intelligence (AI), quantum information sciences (QIS), 5G/advanced communications, biotechnology, and advanced manufacturing." Even as the administration slights scientific research in other areas, including biosecurity, the budget includes "$210 million for the National Science Foundation (NSF) for QIS research, doubling the FY 2020 Budget for QIS," and "$237 million for DOE's [Department of Energy's] Office of Science to support QIS research. This will bolster quantum information efforts at the national laboratories and in academia and industry."[65] Nearly half a billion is earmarked for quantum technology, including $25 million to build a quantum internet connecting 17 national labs.[66] While Congress is unlikely to pass the 2021 budget intact, it is suggestive of the administration's priorities. Moreover, funding for quantum science is likely to be spared the squabbles that embroil more controversial budget items. Despite the extreme polarization in contemporary American politics, there is bipartisan support for increasing investment in quantum science. As the DOE under secretary for science points out, "The

dollars we have put into quantum information science have increased by about fivefold over the last three years."[67] This funding is motivated in no small part by the concern that China could leapfrog ahead of the United States. Investment in quantum information technology has thus become an important component of what the 2018 National Defense Strategy describes as "the re-emergence of long-term, strategic competition between nations."[68]

If the prospect of quantum-safe security via QKD is not enough of a motivation for investing in quantum networking, there are other positive reasons to invest. Quantum networks may enable some applications that are simply infeasible with classical networks. These include encryption schemes allowing users to certify the deletion or retention of data, detect tampering, and create unique time windows for decryption.[69] Quantum computing also holds great promise for scientific modeling and drug discovery.

This article has only explored the technical bounds of the possible, but many other social factors affect the window of vulnerability. Organizational institutions, human behavior, industrial policy, and strategic interaction can squander technological advantages. They can also compensate for technological weaknesses. Even if quantum-safe networks are not available before quantum computers (scenario 1), protecting some secrets will still be possible. Target organizations will still find ways to hide their most valuable secrets by using physically isolated networks or abstaining from digital encoding altogether. Conversely, even in a world of secure quantum-safe networks (scenario 3), it will be still possible to collect secrets by attacking the insecure human endpoints of the network. Strong cryptography, classical or quantum, does not automatically translate into strong information security. Gullible humans, flawed security policy, and sociotechnical complexity can inadvertently expose data protected by quantum-safe systems.[70]

Endemic friction in the sociotechnical implementation of cryptology is something of an insurance policy for *both* offense and defense in any of the three scenarios. The actual performance of either quantum decryption or quantum-safe encryption is unlikely to live up to its full potential. Even if I am too pessimistic about the scientific prospects of quantum computing relative to quantum-safe alternatives, quantum computers will still have to operate in human organizations that offer little reason for optimism. The practical implication is clear. Organizations cannot rely solely upon technology for cryptologic advantage. Information assurance begins and ends with a workforce that understands and cares about the confidentiality, integrity, and availability of relevant data. More complex information

technologies require an even higher level of technical acumen and awareness from personnel, and an even stronger commitment on the part of leadership to maintaining a robust cybersecurity posture. Offensive cyber advantage, conversely, depends on knowing how to exploit the behavior of organizations that fail to maintain their guard.

No technical advantage can be sustained forever, if indeed it can be realized in the first place. In the case of quantum computing, the credible fear that a geopolitical adversary might realize a major intelligence advantage has already mobilized considerable effort for prevention. It is important to sustain this effort. Quantum computing may yet have other important military applications, but we should make sure that an exponential improvement in cryptanalysis will not be one of them. The 2^n horsemen of the cryptocalypse should be just believable enough to make themselves irrelevant. SSQ

Jon R. Lindsay

Jon R. Lindsay is an assistant professor at the Munk School of Global Affairs and Public Policy and Department of Political Science at the University of Toronto. He served in the US Navy in the aviation, intelligence, and special warfare communities. He is the author of *Information Technology and Military Power* (Cornell, 2020) and volumes on deterrence (Oxford, 2019) and cybersecurity (Oxford, 2015).

Notes

1. John Mulholland, Michele Mosca, and Johannes Braun, "The Day the Cryptography Dies," *IEEE Security & Privacy* 15, no. 4 (July/August 2017): 14–21, DOI: 10.1109/MSP.2017.3151325.

2. Frank L. Smith, "Quantum Technology Hype and National Security," *Security Dialogue*, April 2020, https://doi.org/10.1177/0967010620904922.

3. Lawrence Freedman, *The Future of War: A History* (New York: PublicAffairs, 2017).

4. Jon R. Lindsay, "Demystifying the Quantum Threat: Infrastructure, Implementation, and Intelligence Advantage," *Security Studies* 29, no. 2 (2020): 335–61, DOI: 10.1080/09636412.2020.1722853.

5. For a thorough technical assessment, see National Academies of Sciences, Engineering, and Medicine, *Quantum Computing: Progress and Prospects* (Washington, DC: The National Academies Press, 2019), https://doi.org/10.17226/25196.

6. Victoria Craw, "Quantum Computing Set to Revolutionise Everything from National Security to Drug Design and Financial Investments," *News.Com.Au*, 29 January 2018, http://www.news.com.au/technology/. See also "Risk Analysis of Quantum Computer Attacks on Digital Signatures," accessed 21 April 2020, https://www.quantumcryptopocalypse.com/.

7. There are other interesting applications of quantum information science that are beyond the scope of this article, ranging from new types of remote sensing to improvements in scientific modeling and machine learning. For a survey of potential applications, see Michael J. Biercuk and Richard Fontaine, "The Leap into Quantum Technology: A

Primer for National Security Professionals," *War on the Rocks* (blog), 17 November 2017, https://warontherocks.com/; and National Academies of Sciences, Engineering, and Medicine, *Quantum Computing*.

8. Vivek Wadhwa, "Quantum Computers May Be More of an Imminent Threat than AI," *Washington Post*, 5 February 2018, https://www.washingtonpost.com/.

9. The appendix for this article is available online at https://www.airuniversity.af.edu/.

10. R. L. Rivest, A. Shamir, and L. Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems," *Communications ACM* 21, no. 2 (February 1978): 120–26, https://dl.acm.org/doi/10.1145/359340.359342.

11. Simon Singh, *The Code Book: The Science of Secrecy from Ancient Egypt to Quantum Cryptography* (New York: Random House, 1999), 243–92.

12. David Kahn, *Seizing the Enigma: The Race to Break the German U-Boat Codes, 1933–1945*, rev. ed. (Annapolis, MD: Naval Institute Press, 2012), 149–60, 255–66.

13. DigiCert, "Check Our Numbers: The Math behind Estimations to Break a 2048-bit Certificate," accessed 17 May 2019, https://web.archive.org/.

14. Peter W. Shor, "Algorithms for Quantum Computation: Discrete Logarithims and Factoring," in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science* (Los Alamitos, CA: IEEE Computer Society, 1994), 124–34, http://citeseerx.ist.psu.edu/.

15. Craig Gidney and Martin Ekera, "How to Factor 2048 Bit RSA Integers in 8 Hours Using 20 Million Noisy Qubits," arXiv, 6 December 2019, https://arxiv.org/.

16. See Niels Ferguson, Bruce Schneier, and Tadayoshi Kohno, *Cryptography Engineering: Design Principles and Practical Applications* (Indianapolis, IN: John Wiley & Sons, 2010).

17. M. Roetteler and K. M. Svore, "Quantum Computing: Codebreaking and Beyond," *IEEE Security and Privacy* 16, no. 5 (September 2018): 22–36, DOI: 10.1109/MSP.2018.3761710.

18. Erik Lucero et al., "Computing Prime Factors with a Josephson Phase Qubit Quantum Processor," *Nature Physics* 8, no. 10 (October 2012): 719–23, https://www.nature.com/.

19. Frank Arute et al., "Quantum Supremacy Using a Programmable Superconducting Processor," *Nature* 574, no. 7779 (October 2019): 505–10, https://doi.org/10.1038/nphys2385.

20. Gidney and Ekera, "How to Factor 2048 Bit RSA Integers."

21. Samuel K. Moore, "IBM Edges Closer to Quantum Supremacy with 50-Qubit Processor," *IEEE Spectrum: Technology, Engineering, and Science News*, 15 November 2017, https://spectrum.ieee.org/tech-talk/; and Martin Giles and Will Knight, "Google Thinks It's Close to 'Quantum Supremacy.' Here's What That Really Means," MIT *Technology Review*, 9 March 2018, https://www.technologyreview.com/.

22. Scott Aaronson, "Why Google's Quantum Supremacy Milestone Matters," *New York Times*, 30 October 2019, https://www.nytimes.com/.

23. Marc Trachtenberg, "A 'Wasting Asset': American Strategy and the Shifting Nuclear Balance, 1949–1954," *International Security* 13, no. 3 (1988): 5–49, DOI: 10.2307/2538735; and Stephen Van Evera, *Causes of War: Power and the Roots of Conflict* (Ithaca: Cornell University Press, 1999), chap. 4.

24. Charles L. Glaser and Chaim Kaufmann, "What Is the Offense-Defense Balance and Can We Measure It?," *International Security* 22, no. 4 (Spring 1998): 44–82, https://

web.stanford.edu/; and Stephen Biddle, "Rebuilding the Foundations of Offense-Defense Theory," *The Journal of Politics* 63, no. 3 (2001): 741–74, https://www.jstor.org/.

25. For example, orbital physics constrains the employment of antisatellite weapons today as much as in the Cold War despite tremendous advances in spacecraft. Thus, a technical primer over three decades old remains relevant: Ashton B. Carter, "Satellites and Anti-Satellites: The Limits of the Possible," *International Security* 10, no. 4 (Spring 1986): 46–98, https://www.jstor.org/.

26. John Ferris, "Fighter Defence before Fighter Command: The Rise of Strategic Air Defence in Great Britain, 1917–1934," *Journal of Military History* 63, no. 4 (1999): 845–84, https://www.jstor.org/.

27. David Zimmerman, *Britain's Shield: Radar and the Defeat of the Luftwaffe* (Stroud, UK: Sutton Publishing, 2001).

28. Erik Gartzke and Jon R. Lindsay, "Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace," *Security Studies* 24, no. 2 (2015): 316–48, http://dx.doi.org/10.1080/09636412.2015.1038188; and Rebecca Slayton, "What Is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment," *International Security* 41, no. 3 (2017): 72–109.

29. Richard A. Clarke and Robert K. Knake, *The Fifth Domain: Defending Our Country, Our Companies, and Ourselves in the Age of Cyber Threats* (New York: Penguin, 2019).

30. According to the 2018 *Command Vision for US Cyber Command*, "Cyberspace is an active and contested operational space in which superiority is always at risk." United States Cyber Command, *Achieve and Maintain Cyberspace Superiority: Command Vision for US Cyber Command* (Ft. Meade, MD: USCYBERCOM, June 2018), 6, https://www.cybercom.mil/.

31. Erica D. Borghard and Shawn W. Lonergan, "Cyber Operations as Imperfect Tools of Escalation," *Strategic Studies Quarterly* 13, no. 3 (2019): 122–45; and Max Smeets, "The Strategic Promise of Offensive Cyber Operations," *Strategic Studies Quarterly* 12, no. 3 (2018): 90–113, https://www.airuniversity.af.edu/.

32. Max Smeets, "A Matter of Time: On the Transitory Nature of Cyberweapons," *Journal of Strategic Studies* 41, no. 1–2 (2017): 6–32, https://doi.org/10.1080/01402390.2017.1288107.

33. Bruce Schneier, "Class Breaks," *Schneier on Security* (blog), 3 January 2017, https://www.schneier.com/.

34. Gilles Brassard quoted in Hoi-Kwong Lo and Norbert Lütkenhaus, "Quantum Cryptography: From Theory to Practice," *Physics in Canada* 63, no. 4 (2007): 191–96, https://arxiv.org/.

35. Glenn S. Gerstell, "I Work for N.S.A. We Cannot Afford to Lose the Digital Revolution," Opinion, *New York Times*, 10 September 2019, https://www.nytimes.com/.

36. C. L. Max Nikias, "This Is the Most Important Tech Contest since the Space Race, and America Is Losing," *Washington Post*, 11 May 2018, https://www.washingtonpost.com/.

37. Central Committee of the Communist Party of China, "The 13th Five-Year Plan for Economic and Social Development of the People's Republic of China (2016–2020)," December 2016, https://en.ndrc.gov.cn/; and Eamon Barrett, "Google and NASA Have Claimed Quantum Supremacy, but China Is Not Far Behind the U.S.," *Fortune*, 30 October 2019, https://fortune.com/2019/.

38. Dan Breznitz and Michael Murphree, *Run of the Red Queen: Government, Innovation, Globalization, and Economic Growth in China* (New Haven, CT: Yale University Press, 2011).

39. Elsa B. Kania and John Costello, *Quantum Hegemony? China's Ambitions and the Challenge to U.S. Innovation Leadership* (Washington, DC: Center for a New American Security, 2018), https://www.cnas.org/; and Taylor Owen and Robert Gorwa, "Quantum Leap: China's Satellite and the New Arms Race," *Foreign Affairs*, 7 September 2016, https://www.foreignaffairs.com/.

40. Biddle, "Rebuilding the Foundations of Offense-Defense Theory."

41. Greg Austin, *Cybersecurity in China: The Next Wave* (Zurich: Springer, 2018).

42. Robert L. Paarlberg, "Knowledge as Power: Science, Military Dominance, and U.S. Security," *International Security* 29, no. 1 (2004): 122–51, https://www.jstor.org/.

43. David M. Edelstein, *Over the Horizon: Time, Uncertainty, and the Rise of Great Powers* (Ithaca, NY: Cornell University Press, 2017).

44. Scholars in the realist tradition of international relations generally expect political actors to balance against threats. A classic statement is found in Kenneth N. Waltz, *Theory of International Politics* (Reading, MA: Addison-Wesley, 1979). Another version of balancing theory argues that threatening intentions must be present to provoke balancing, not just threatening capability; see Stephen M. Walt, *The Origins of Alliance* (Ithaca, NY: Cornell University Press, 1990). Some counter that intentions are so fickle that states tend to balance primarily against military power; see John J. Mearsheimer, *The Tragedy of Great Power Politics* (New York: Norton, 2001). Others highlight historical anomalies, for example, Paul Schroeder, "Historical Reality vs. Neo-Realist Theory," *International Security* 19, no. 1 (1994): 108–48, https://doi.org/10.2307/2539150. My goal is not to adjudicate this debate but simply to suggest that the intuitive concept of political balancing can be extended to the realm of scientific threats. Indeed, geopolitics is often the catalyst for the emergence of a scientific threat in the first place. While the nature of military and scientific threats may differ, the political dynamics of balancing responses—worrying about uncertainty, dithering under long-time horizons, free riding on others, mobilizing action against manifest threats, for example—have a familiar rhyme and reason.

45. Lily Chen et al., *Report on Post-Quantum Cryptography*, NISTIR 8105 (Gaithersburg, MD: National Institute of Standards and Technology, April 2016), http://dx.doi.org/10.6028/NIST.IR.8105; K. Lauter, "Postquantum Opportunities: Lattices, Homomorphic Encryption, and Supersingular Isogeny Graphs," *IEEE Security Privacy* 15, no. 4 (2017): 22–27, DOI: 10.1109/MSP.2017.3151338; and Daniel J. Bernstein et al., "Post-Quantum RSA," Cryptology ePrint Archive, 19 April 2017, https://eprint.iacr.org/.

46. Computer Security Resource Center, "Announcing Request for Nominations for Public-Key Post-Quantum Cryptographic Algorithms," National Institute of Standards and Technology, 20 December 2016, https://csrc.nist.gov/.

47. Dustin Moody, "Let's Get Ready to Rumble: The NIST PQC 'Competition,'" paper presented at PQCrypto 2018, Fort Lauderdale, FL, 11 April 2018, https://csrc.nist.gov/.

48. Bruce Schneier, "NSA Plans for a Post-Quantum World," *Lawfare* (blog), 21 August 2015, https://www.lawfareblog.com/. The NSA notes that "for those partners and vendors that have not yet made the transition to Suite B elliptic curve algorithms, we recommend not making a significant expenditure to do so at this point but instead to prepare for the upcoming quantum resistant algorithm transition." Quoted in Schneier.

49. L. Chen, "Cryptography Standards in Quantum Time: New Wine in an Old Wineskin?," *IEEE Security and Privacy* 15, no. 4 (2017): 51–57, https://doi.org/10.1109/MSP.2017.3151339.

50. According to Eleni Diamanti et al., "Encryption keys generated by QKD can be used in a symmetric cipher scheme, such as Advanced Encryption Standard, which is quantum resistant, for enhanced security, or they can be combined with the one-time-pad encryption scheme for unconditional security." See Eleni Diamanti et al., "Practical Challenges in Quantum Key Distribution," *Npj | Quantum Information* 2 (8 November 2016): 3, https://doi.org/10.1038/npjqi.2016.25.

51. P. D. Townsend, "Secure Key Distribution System Based on Quantum Cryptography," *Electronics Letters* 30, no. 10 (May 1994): 809–11, https://ieeexplore.ieee.org/.

52. Diamanti et al., "Practical Challenges in Quantum Key Distribution"; Christoph Simon, "Towards a Global Quantum Network," *Nature Photonics* 11, no. 11 (November 2017): 678–80, https://doi.org/10.1038/s41566-017-0032-0; Stephanie Wehner, David Elkouss, and Ronald Hanson, "Quantum Internet: A Vision for the Road Ahead," *Science* 362, no. 6412 (19 October 2018), https://doi.org/10.1126/science.aam9288; and Acín et al., "Quantum Technologies Roadmap."

53. According to Diamanti et al., "In China, the deployment of a 2,000 km QKD network between Shanghai and Beijing is underway; in Europe, after the [Secure Communication based on Quantum Cryptography] network demonstration in 2008, the UK is now creating a quantum network facilitating device and system trials, and the integration of quantum and conventional communications; in Japan, QKD technologies will be tested to secure transmission of sensitive genome data; and the US has also started installing its own QKD network." See Diamanti et al., "Practical Challenges in Quantum Key Distribution," 1–2.

54. Diamanti et al.

55. Tom Berson (@nd2t), "Quantum Key Distribution," Twitter, 23 July 2019, 4:30 p.m., https://twitter.com/.

56. Simon, "Towards a Global Quantum Network"; and Wehner, Elkouss, and Hanson, "Quantum Internet."

57. Information Security Oversight Office, "The President: Executive Order 13526; Classified National Security Information," 29 December 2009, National Archives, https://www.archives.gov/.

58. Michael Warner and Robert Louis Benson, "Venona and Beyond: Thoughts on Work Undone," *Intelligence and National Security* 12, no. 3 (July 1997): 1–13; and Richard Aldrich, *GCHQ: The Uncensored Story of Britain's Most Secret Intelligence Agency* (London: HarperCollins, 2010), 72–88.

59. Michael Warner, *The Rise and Fall of Intelligence: An International Security History* (Washington, DC: Georgetown University Press, 2014), 153.

60. This framework draws on Michele Mosca, "Cybersecurity in an Era with Quantum Computers: Will We Be Ready?," paper presented at the 5th Conference on Quantum Cryptography, Tokyo, 2015, Cryptology ePrint Archive, Report 2015/1075, https://eprint.iacr.org/.

61. Jean-Pierre Dupuy, *A Short Treatise on the Metaphysics of Tsunamis*, trans. Malcolm B. DeBevoise (East Lansing: Michigan State University Press, 2015). Dupuy points out that such disasters are usually represented as being beyond control—acts of god—yet are actually made more or less likely by the behavior of their victims.

62. Steven Rich and Barton Gellman, "NSA Seeks to Build Quantum Computer That Could Crack Most Types of Encryption," *Washington Post*, 2 January 2014, https://www.washingtonpost.com/.

63. Stephen Budiansky, "Colossus, Codebreaking, and the Digital Age," in *Colossus: The Secrets of Bletchley Park's Codebreaking Computers*, ed. B. Jack Copeland (New York: Oxford University Press, 2006), 52–63.

64. Rich and Gellman, "NSA Seeks to Build Quantum Computer."

65. White House, "Advancing United States Leadership in the Industries of the Future," FY 21 fact sheet, 2 February 2020, https://www.whitehouse.gov/.

66. Cade Metz, "White House Earmarks New Money for A.I. and Quantum Computing," *New York Times*, 10 February 2020, https://www.nytimes.com/.

67. Metz.

68. Department of Defense, *Summary of the 2018 National Defense Strategy of the United States of America: Sharpening the American Military's Competitive Edge* (Washington, DC: Department of Defense, 2018), 2, https://dod.defense.gov/.

69. Anne Broadbent and Christian Schaffner, "Quantum Cryptography beyond Quantum Key Distribution," *Designs, Codes and Cryptography* 78, no. 1 (January 2016): 351–82, https://doi.org/10.1007/s10623-015-0157-4.

70. For further discussion of the social context of quantum cryptology, see Lindsay, "Demystifying the Quantum Threat."

PERSPECTIVE

Hypersonic Technology: An Evolution in Nuclear Weapons?

Lt Col Nathan B. Terry, USAF
Paige Price Cone

Abstract

Hypersonic delivery systems are a grave concern because they are potentially fast and maneuverable enough to evade existing defensive systems. As the US military considers upgrading its nuclear arsenal, hypersonic delivery systems are one possible option. Increased research on hypersonic technologies over the past two decades demonstrates there is technical feasibility for hypersonic conventional weapons. The case for nuclear-armed hypersonic weapons (NAHW) is more complicated. This article considers NAHWs from the point of view of deterrence thinking and suggests a NAHW is consistent with current US thinking about deterrence with respect to existing ballistic missiles, cruise missiles, and missile defense systems. However, we conclude that there are few advantages to hypersonic nuclear delivery systems relative to existing nuclear weapon delivery systems.*

*****

The Department of Defense under secretary for research and engineering, Michael Griffin, recently declared hypersonic technology to be his top technology priority.[1] The former commander of US Strategic Command, Gen John Hyten, says the US currently does not have "any defense that could deny the employment" of hypersonic weapons.[2] These statements demonstrate how hypersonic weapons present unique strategic opportunities and challenges. For example, hypersonic weapons promise to defeat existing missile defense systems—something limited salvos of intercontinental ballistic missiles (ICBM) may be unable to do.[3] Currently, the US, Russia, and China are actively working to develop advanced hypersonic weapon systems, and other countries are interested as well.[4] Public reports describe US hypersonic development in

*The authors would like to thank Dr. James Platte, Dr. Carrie Lee, and Dr. Bob Greendyke for helpful discussions.

terms of conventional systems capable of providing a prompt, long-range strike capability.[5]

There has been significant discussion of conventional hypersonic weapons. Much of this research considers whether conventionally armed hypersonic weapons might prove destabilizing.[6] However, there has been little specific examination of whether hypersonic delivery systems for nuclear weapons may prove destabilizing. We argue that US nuclear-armed hypersonic weapons (NAHW) will not be destabilizing in terms of nuclear deterrence.

The analysis starts by considering how evolutionary technological changes developed concurrently with deterrence thinking and how previous scholars evaluated technology's impact on deterrence thinking. Then it assesses how a NAHW might affect deterrence thinking. We compare future NAHWs against three existing nuclear-related technological systems: ICBMs, cruise missiles, and missile defense. The analysis examines the historical development of each element to show that US hypersonic technology is evolutionary relative to these elements. It also considers whether the elements will combine synergistically. Next, our analysis appraises the potential implications of two sides having NAHWs, again in the context of the key parameters of existing systems. Ultimately, we conclude that hypersonic development is evolutionary; therefore, NAHWs will not be destabilizing relative to existing nuclear weapons delivery technology or offer great advantage. While a historical analysis of hypersonic component technology shows rapid advancement, no NAHW has been openly fielded.

Evolutions and Technology

Hypersonic vehicles, commonly characterized as highly maneuverable systems traveling at speeds of at least Mach 5, comprise two classes of hypersonic systems: hypersonic cruise missiles and hypersonic boost glide vehicles.[7] The word *hypersonic* generally refers to these two systems. ICBMs and submarine-launched ballistic missiles (SLBM) travel faster than Mach 5 but are not maneuverable, so they are not considered hypersonic weapons for purposes of this article. Here the focus is on the implications of high-speed, maneuverable nuclear weapon systems. Maneuverability allows NAHWs to potentially evade missile defense systems. This makes them potentially useful against adversaries with effective defenses against ICBMs, SLBMs, or nuclear-armed cruise missiles.

Hypersonics is not actually a single technology. Rather, it is a class of related technologies that must be combined together to form an opera-

tionally useful system. The Defense Advanced Research Projects Agency (DARPA) reports that successful hypersonic systems require the effective combination of a number of technologies, including high-speed supersonic combustion ramjet (scramjet) engines, high-temperature materials capable of managing the high heat loads associated with hypersonic flights, advanced manufacturing techniques, and advanced vehicle configurations.[8] An analysis of the journal publications in each area, shown in figure 1,[9] demonstrates that hypersonic technologies have been changing rapidly over the last few decades.[10]

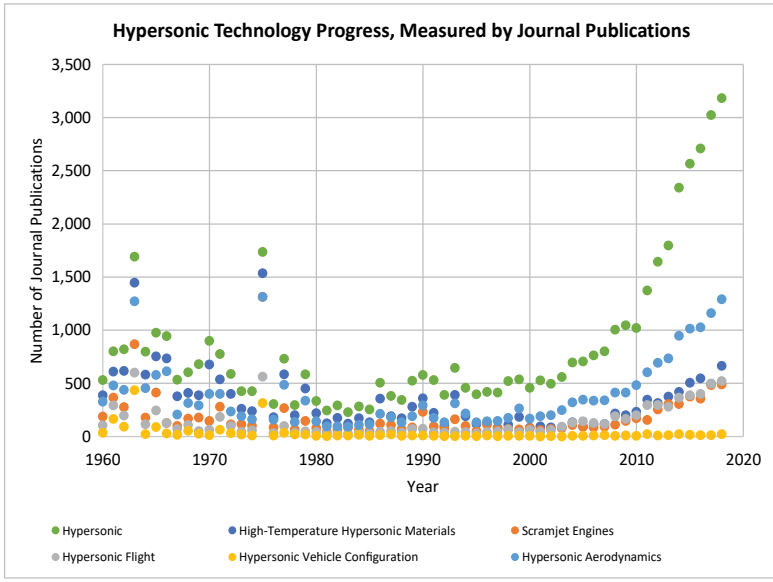


Figure 1. Number of journals published yearly reporting progress in hypersonic and hypersonic component technology

The rapid changes to hypersonic technologies do not necessarily lead to radical alterations in our thinking about deterrence. Many technological changes are evolutionary improvements in technology that only catalyze evolutions in strategic thinking. However, in the case of nuclear weapons, the capability improvement was so radical that the new technology revolutionized how nations thought about war. When a new technology like hypersonic delivery systems is developed, how should it be evaluated in terms of deterrence thinking? Will it turn out to be a breakthrough that significantly changes deterrence, or will it rather be an important but incremental change in existing technology?

Bernard Brodie proposed nuclear deterrence in 1946 in response to the tremendous power of nuclear weapons. To him, nuclear bombs represented a 700-fold increase of the destructive power provided by bombers, en-

abling a single aircraft on a single mission to potentially destroy a city.[11] While B-29s were vulnerable to various defenses—including fighter aircraft, antiaircraft weapons, or even preemptive ground attacks—the potential for destruction represented by nuclear-armed bombers was so great that Brodie argued the main purpose of nuclear weapons was prevention, not fighting.[12]

This concept was a revolutionary development in the thinking about war. Large armies had been used for centuries both to prevent attacks and to carry them out. As Thomas Schelling explained, prior to the advent of nuclear weapons, only the loser was punished—and then, only after it lost.[13] Nuclear weapons could destroy so much and so quickly, he argued, that annihilation could come to either side at any time during the conflict. Brodie and Glenn Snyder reasoned that nuclear weapons fundamentally changed warfare, making the case that the only purpose of nuclear weapons was to deter.[14] Schelling further expanded the ideas of deterrence, indicating that although nuclear bombs were not exploded during conflicts such as the Korean War, the Cuban missile crisis, and the Vietnam War, the specter of nuclear weapons loomed large in the minds of great powers—deterring escalation to even greater levels of conflict.[15] Writing in 1996, Robert Pape considered the role of bombing campaigns in war and suggested that nuclear bombs were far better suited to threats than to actual attacks.[16] In many ways, the theme of deterrence remained remarkably consistent over time.

While scholars wrestled with deterrence, nuclear weapons technology morphed to create new and improved bombs, delivery systems, and defenses. Great powers went from mere nuclear bombs to thermonuclear bombs; weapon yields increased from tens of kilotons to tens of megatons, tripling or quadrupling the size of a city that could be obliterated.[17] Bombers leveraged a combination of novel guidance technologies and new configurable wing-design technology. This flexibility allowed them to switch between fuel-efficient, high-altitude flights over friendly territory and less risky, low-altitude flights over enemy territory—greatly increasing aircraft range and survivability.[18]

Ballistic missile technology stemming from World War II was adapted to the US nuclear arsenal. US ICBMs were operational in 1958, and the first sea-launched ballistic missiles were deployed in 1960.[19] A whole host of technologies went into improving the range and accuracy of ICBMs, including high-precision inertial components, transistors for miniaturized navigation computers, smart fuses to handle missile navigation errors, and rapid retargeting technologies to reduce the number of missiles needed to

attack targets.[20] Air-launched, nuclear-armed cruise missiles were operationally deployed in 1958 with a warhead about 10 times as powerful as the weapon used in Hiroshima.[21] Air-launched cruise missiles supported standoff attacks by bombers and small engine technology increased effective missile ranges.[22] New materials allowed higher engine operating temperatures, in turn increasing fuel efficiency and range.[23]

Defenses against nuclear weapons advanced as well. Soviet developments in radar, command and control, and fighter technology further threatened bombers, driving requirements for missiles with improved standoff attack range.[24] More accurate ICBMs potentially threatened the survivability of adversary ICBM forces.[25] The survivability of ICBMs was increased by developing systems and technology for hardening, redundancy, multiple warheads, concealment, and mobility.[26] Increasingly accurate delivery systems developed to thwart hardening and concealment are increasingly being offset by increased intelligence, surveillance, and reconnaissance (ISR) capabilities.[27] In-flight missile survivability improved through technologies such as decoys, chaff, alternate trajectories, radiation hardening, and electronic countermeasures.[28] In the 1980s, the US explored a host of new defenses against ballistic missiles under the aegis of the Strategic Defense Initiative.[29] Even after decades of unparalleled technological changes, the US nuclear arsenal remains postured in a way familiar to Brodie, Snyder, Schelling, and others. If defenses against and counters to hypersonic weapons eventually emerge, hypersonic weapons may not appreciably change strategic nuclear postures.

Writing in 1957, Kissinger evaluated the impact of coupling nuclear warheads to missile delivery systems, a relatively new technology at the time. His arguments can be grouped into two criteria.[30] First, technology should be evaluated relative to the advantages provided to one side, particularly in terms of existing systems. Second, eventually technological parity would be reached, and thus technology should be evaluated regarding the implications of both sides possessing the technology.

Using these benchmarks, Kissinger was skeptical about the utility of upgrading the missile-based delivery systems of the 1950s. Using the first criterion, he argued that once a missile was capable of traveling 5,000 miles in half an hour, additional increases in speed would "prove only marginally significant."[31] He added, "After a certain point, superiority in destructive power no longer pays strategic returns."[32] Kissinger reasoned that thermonuclear warheads were more advantageous to the Soviets than to the US because at the time, the US possessed a larger nuclear force.[33] Kissinger further noted that Soviet ballistic missiles were not a break-

through because short-range missiles provided minimal advantage relative to the existing Soviet bomber force.[34] Each argument compared the capabilities of new and improved weapons to existing weapons and concluded that the small increases in capability were evolutionary.

Using the second criterion, Kissinger evaluated the implications of both sides having upgraded weapons. He reasoned that survivability through concealment, mobility, or dispersion made a successful first strike impractical for both sides.[35] Kissinger acknowledged that technical parity did not always equal strategic parity, arguing that SLBMs were more threatening to a naval power than to a landlocked nation.[36] This second criterion complements the first by considering technological developments relative to the overall strategic situation rather than simply in a vacuum. Considerations such as relative force sizes and force postures influence whether a technology has evolutionary or revolutionary implications for deterrence thinking.

Kissinger also considered future technology developments when ascertaining the implications of current technological developments. He stated that as one side builds missiles, the other side would reach parity relatively quickly.[37] Technology levels, he reasoned, are not inherently stable. In his mind, there is no such thing as equilibrium in terms of technology-based capabilities because parity is a fleeting thing.[38] Colin Gray used similar arguments to conclude that arms races are rarely destabilizing because as one side gains a technological advantage, the other develops a countermeasure. While supportive of pursuing technological changes, Gray remained unconvinced that new technologies would fundamentally alter the principles of deterrence because each technological advance would eventually be countered.[39]

These contentions acknowledge that major changes in military technology like the development of nuclear weapons altered the way nations viewed weapons and warfare. However, subsequent changes in weapons and weapon delivery technology have been evolutionary and have not significantly changed thinking about nuclear deterrence. While each side pursued technological developments to gain some advantages, historical analysis shows that when changes were evolutionary, the resulting instability was temporary. This brief overview of nuclear weapon delivery technology shows that scientific developments have been an ongoing evolutionary process. Most of the aforementioned improvements were adaptions of existing technologies. Despite all these technology changes, the basic form of the nuclear triad for deterrence is still largely recognizable even well over a half-century later. In 2018, the United States *Nuclear Posture Review* declared that the US continues to use a combination of nuclear-

armed bombers, SBLMs, and ICBMs to deter nuclear attack, stating that US adversaries must understand that "any nuclear escalation will fail to achieve their objectives."[40]

This analysis clearly shows that most hypersonic component technologies are developing at an expeditious rate, especially compared to historical trends. However, while hypersonic technology is improving, technology improvements cannot be considered in a vacuum. Instead, determining whether hypersonic delivery technology is evolutionary or revolutionary requires a comparison with existing nuclear weapon delivery technologies. Reference points, such as the performance of cruise missiles or ICBMs, are needed against which to benchmark the progress of hypersonic technology.

Comparing Existing Nuclear Systems and Nuclear-Armed Hypersonic Weapons

One way to benchmark the impact of a potential future system is by analogy to existing systems. Many existing systems have a long history and have been thoroughly analyzed in terms of their impact on deterrence thinking. Second, any new system is going to potentially complement or replace existing systems, making them an appropriate baseline. ICBMs, cruise missiles, and missile defense are three analogs that have been analyzed regarding their impact on deterrence postures and thinking. Each component shares some similarities with a NAHW. A superposition of these three component features describes all the essential elements of NAHWs and provides key parameters that can be analyzed to determine if hypersonic technology is revolutionary or evolutionary for each component.

A *ballistic missile*, defined as "a projectile that assumes a free-falling trajectory after an internally guided ascent," travels very fast—a characteristic of all forms of hypersonic technology.[41] Thus, ICBMs are a good analog to future long-range NAHWs. There is keen interest in increasing the range of hypersonic weapons, suggesting that the long range of ballistic missiles is another reason to include them in the model.[42] However, ballistic missiles are inaccurate and generally follow predictable flight paths, indicating that additional elements are needed for a working model useful for analyzing the range of future capabilities promised by NAHWs.

A *cruise missile* is defined in part by the Intermediate-Range Nuclear Forces (INF) Treaty of 1987 as "an unmanned self-propelled guided vehicle that sustains flight through aerodynamic lift for most of its flight path."[43] Cruise missiles are maneuverable, meaning they can make course adjustments to improve their accuracy or to avoid obstacles and defenses.[44]

The maneuverability of cruise missiles makes it difficult for defenders to determine their destination, potentially reducing the reaction time of missile defense systems. On the other hand, cruise missiles are relatively slow, making their maneuvers easy to track and leaving substantial time for defenders to react.

Missile defense systems are extremely complicated.[45] Just as there are many types of missiles each with its own combination of vulnerabilities and defenses, there are multiple types of missile defense systems, each a complex collaboration of sensors and shooters. Coordinating between the various elements presents huge technical challenges, especially for targets defended by layers of missile defense systems. Coordination between the various layers means missile defenses need enough time to operate effectively—a luxury that hypersonic weapons may not allow.

From a mathematical point of view, NAHWs can be seen as a superposition of three elements: ICBMs, cruise missiles, and the negative (opposite) of missile defense systems. The stability implications of NAHWs should be considered relative to changes in these three elements. The central features of NAHWs (speed, range, accuracy, and missile defense) are evaluated by considering whether hypersonic systems are a revolution or an evolution in the key technology associated with each element. The first three factors are evaluated quantitatively while the evolution of missile defense is evaluated qualitatively. The analysis concludes by considering whether the individual elements might combine synergistically to create a new path to a revolutionary technology combination.

Intercontinental Ballistic Missiles

ICBMs, the first component analog of hypersonic technology, will be analyzed in terms of changes to speed, range, and accuracy. It is important to consider ICBMs in the context of the other legs of the nuclear triad, so the development of SLBMs is considered here as well. The bomber leg of the nuclear triad is considered later in this section in terms of the technological evolution of cruise missiles. Gravity weapons employed as part of the nuclear triad are not considered in this analysis because gravity bombs have less in common with potential NAHWs than with existing missiles.

Figure 2 plots the speed of US ICBMs and SLBMs as a function of the year various systems obtained an initial operating capability (IOC). The system IOC is used because it provides a useful historical marker noting when a technology transitions to operational employment. Other measures of technological progress are possible, such as dates and results of key missile test flights. However, there is often considerable additional devel-

opment necessary to go from a proof-of-concept test system to an operationally fielded system. For example, operational systems represent optimization between multiple contradictory requirements. Unlike prototypes, operational systems have additional requirements, such as terms of initial and recurring costs, usability, manufacturability, and sustainability in real-world environments. Furthermore, operationally deployed systems may have a different impact on deterrence than test systems that may fail to be operationally deployed.

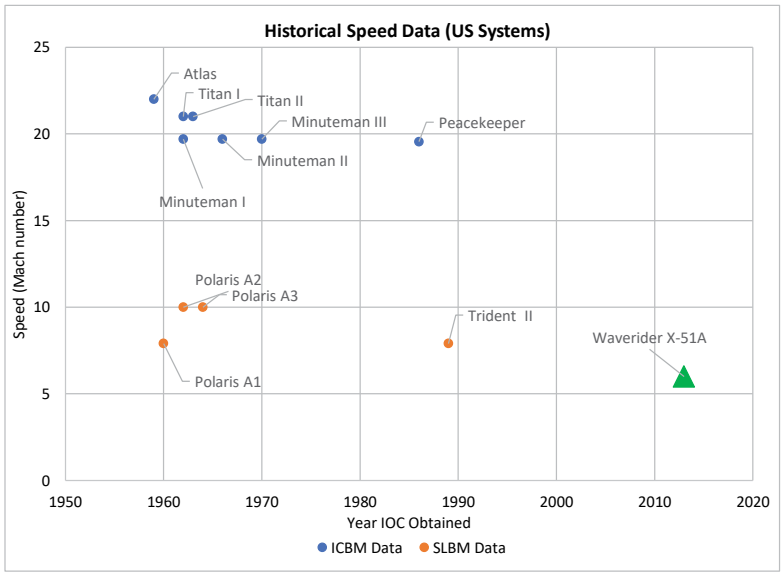


Figure 2. Speed of US ICBMs and SLBLMs as a function of the year IOC obtained[46]

Figure 2 further shows how sequential versions of ICBMs and SLBMs did not appreciably increase their speed over the last 50 years. Instead, speed remained relatively constant or even decreased. Increasing speeds may not matter for prompt nuclear strikes since SLBMs and other shorter-ranged ballistic nuclear weapons can already strike targets very rapidly.[47] This graph also shows that current US hypersonic technology lags ICBM and SLBM technology in speed. So far, hypersonics is an evolutionary technology relative to the speed of existing ballistic nuclear weapon delivery systems.

Ballistic missiles have a much longer trajectory than hypersonic weapons, suggesting that raw speed is a poor comparison. James Acton et al. estimated that early warning satellites could provide up to 30 minutes of warning time for an attack by ICBMs and slightly longer in the case of boost-glide hypersonic weapons. While satellites might provide a much shorter (16 minute) warning against hypersonic cruise missiles, this time

frame would be similar to the warning times provided against intermediate-range ballistic missiles.[48] SBLMs are estimated to be able to reach their targets in about five minutes in many scenarios.[49] NAHWs are unlikely to change warning times of the overall arsenal. Furthermore, it is difficult to see the strategic advantage of even faster attack times, especially considering the aforementioned analyses estimate it will take at least seven or eight minutes to notify the president of the United States of an impending nuclear attack. While there are other potential advantages to hypersonic delivery systems, there is little to gain by simply increasing speed relative to that of existing nuclear weapon delivery systems.

Another key component of a nuclear weapon delivery system is range. The historical evolution of cruise missiles, ICBMs, and SLBMs provides context for evaluating the range of potential hypersonic nuclear delivery systems. As seen in figure 3, the range of US ICBMs did not substantially increase over time. Since they could already cross continents, additional range improvements were gratuitous.[50] Technology did improve the relatively short ranges of US SLBMs, and they grew to eventually equal the range of US ICBMs. This development arguably increased the operational attack range of submarines, making them harder to find in a vast ocean and thereby increasing their survivability. Overall, historical technological changes in ICBM range were small while the range of SLBM technology increased steadily over time. However, current US hypersonic technology is dwarfed by existing ballistic technology in terms of range.

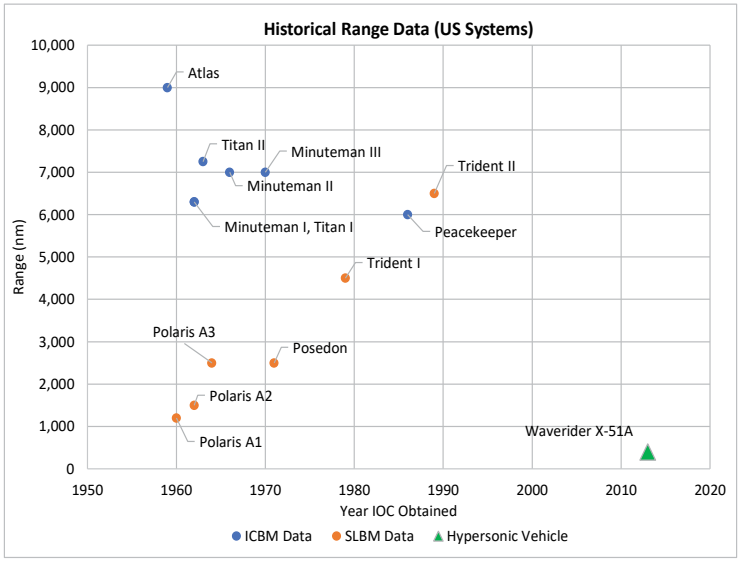


Figure 3. Range of US ICBMs as a function of the year IOC obtained[51]

One virtue of conventional hypersonic weapons is that they may be accurate enough to destroy individual vehicles, suggesting they may eventually provide significantly greater accuracy than provided by ICBMs. Increasing weapon accuracy by a factor of two is functionally equivalent to increasing yield by a factor of eight, meaning an accurate bomb is often better than a bigger bomb.[52] Missile accuracy is defined in terms of circular error probability (CEP), the range described by a circle within which a missile has a 50 percent probability of striking.[53] As figure 4 depicts, US ICBMs generally increased their accuracy through each technology upgrade. A similar analysis shows that Russian and Chinese ICBMs and SLBMs were characterized by evolutionary changes in speed, range, and accuracy.[54] While little data is available about US hypersonic delivery systems, the accuracy of ICBMs is significantly less than that of cruise missiles. Therefore, a NAHW will probably be significantly more accurate then ICBMs.

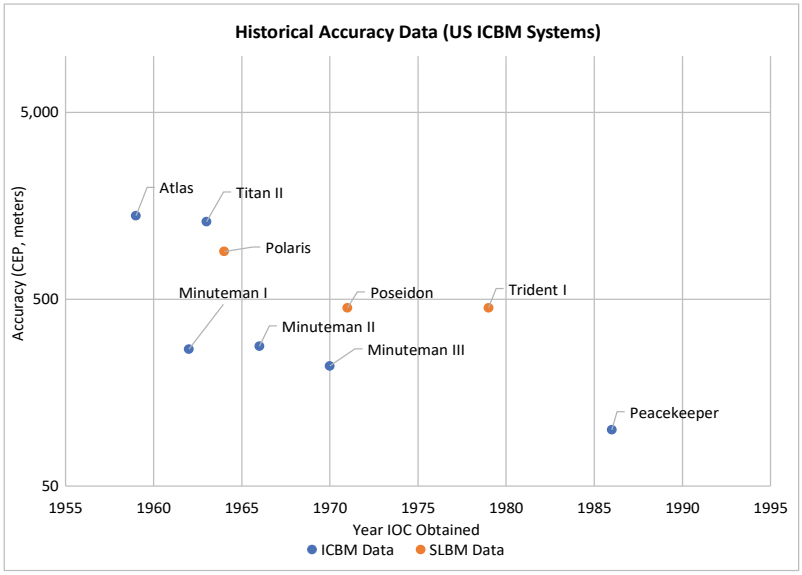


Figure 4. Accuracy of US ICBMs and SLBMs as a function of the year IOC obtained[55]

The range and speed of US ICBMs has been relatively stable over time.[56] US hypersonic missiles show little sign of leapfrogging ICBMs in terms of these parameters. However, hypersonic missiles will likely prove to be more accurate than ICBMs. On one hand, increasing the accuracy of a nuclear weapon delivery system by a few tens of meters may be inconsequential for strategic nuclear weapons with blast radii measured in miles. On the other hand, increased accuracy allows smaller-yield nuclear warheads to be con-

sidered for various missions. Accuracy is an important characteristic of cruise missiles, the next element considered in our model of NAHWs.

Cruise Missiles

Figure 5 shows the range of cruise missiles as a function of the year they obtained IOC. The range of cruise missiles has remained relatively fixed over the last 70 years. While limited range presents some operational constraints, many cruise missiles are released from mobile platforms like ships or aircraft capable of independently maneuvering close to their targets. Historically, acquiring bases close to the Soviet Union was an important consideration in overcoming the range limits of aircraft. In any case, the current unclassified range of hypersonic weapons is well within the range of existing cruise missiles. From the perspective of range, NAHWs are an evolution—rather than a revolution—relative to cruise missiles.

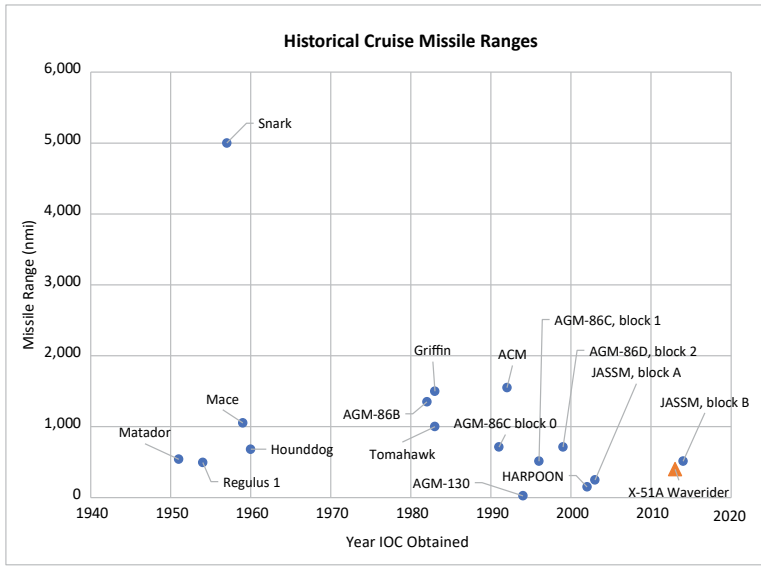


Figure 5. Range of US cruise missiles as a function of the year IOC obtained[57]

Figure 6 shows that US cruise missiles have operated at speeds of less than Mach 2 for the last 50 years. Figure 6 contains fewer data points than figure 5 because the operational speed of several current US missiles remains classified.[58] It should be noted that this data compares an experimental test system (X-51A) to operational systems, an inevitable limitation since hypersonic weapons have not been fielded. Also, the classified nature of most recent cruise missile data may mask a recent evolution in cruise missile technological capability. However, while hypersonic missile speeds

exceed the known speeds of existing cruise missiles, the reported speed of some hypersonic weapons is still considerably slower than for ICBMs.

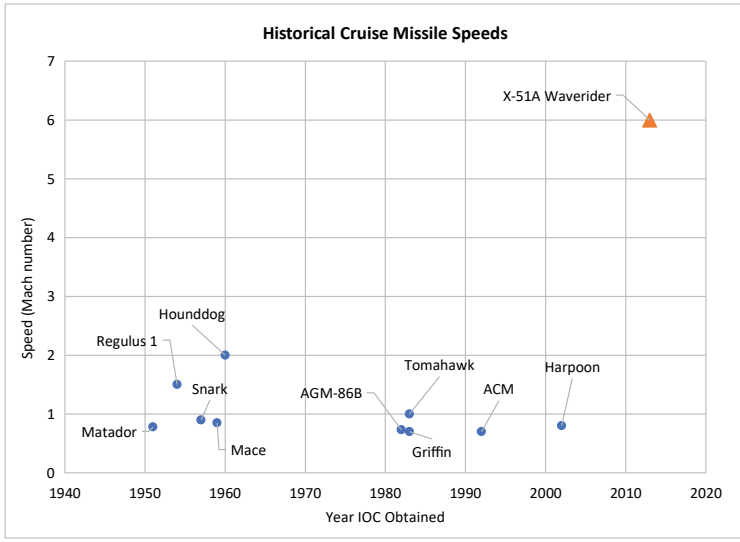


Figure 6. Speed of US cruise missiles as a function of the year IOC obtained[59]

While little data is available on the accuracy of cruise missiles, a comparison between figure 4 and figure 7 (below) shows that the accuracy of cruise missiles is much greater than for ICBMs and that cruise missile accuracy increased over time. However, as with nuclear weapons, it is unclear why increases in accuracy on the order of meters might prove decisive with weapons whose blast radius is measured in kilometers.

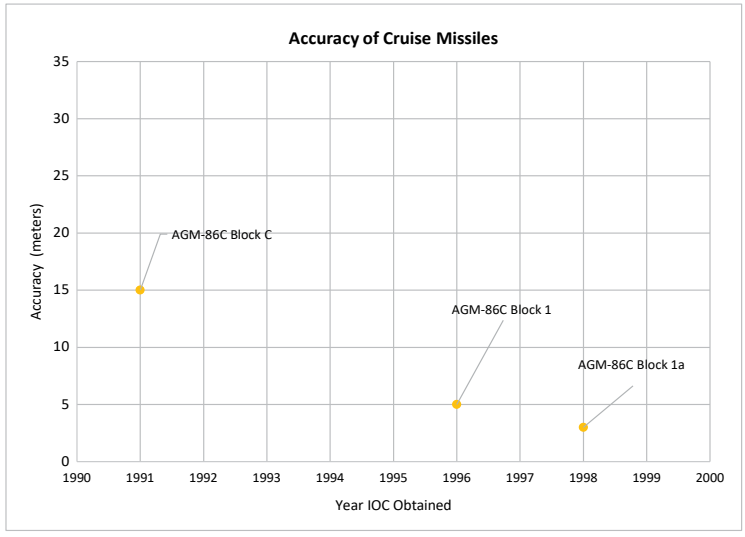


Figure 7. Accuracy of US cruise missiles as a function of the year IOC obtained[60]

As shown in the previous graphs, technology to improve the speed, range, and accuracy of ICBMs and cruise missiles took years or even decades to develop. This growth is an evolution compared to how nuclear weapons increased bomb yields by a factor of 700 over the course of a few years.[61] This revolutionary improvement dramatically changed national strategies and policies. Even though it has been researched for decades, hypersonic technology has not yet leapfrogged existing nuclear missile delivery technology. This suggests that hypersonic technology is not revolutionary. Therefore, a NAHW can be described using existing deterrence thinking. This does not imply that hypersonic technology is meaningless. As discussed earlier, technology evolution is an important part of nuclear weapon delivery system technology. The evolutionary nature of hypersonics is simply a strong argument that coupling nuclear weapons with hypersonic delivery vehicles is consistent with historical technological developments and with current US thinking about nuclear deterrence.

In terms of range, speed, and accuracy, sequential versions of ICBMs are best described by evolutions in technical capability rather than by revolutions in technology. In terms of range and speed, hypersonic technology did not leapfrog the capabilities of existing ICBMs and cruise missiles in three of the four metrics used in this study. The individual strategically relevant component technologies have been evolving slowly. Furthermore, the component technologies will likely combine in a linear way to form NAHWs. Therefore, NAHWs are an evolution relative to existing nuclear delivery technology. Using hypersonic technology for nuclear weapon delivery may provide strategic advantages, but it will likely not prove to be destabilizing.

Missile Defense

Hypersonic technology is remarkable because it provides another means to improve missile survivability. The combination of speed and maneuverability may give hypersonic weapons the potential to mitigate existing missile defense technologies. Missile defenses need time to observe a launch, deduce the object is a missile, classify the missile flight parameters, distinguish the missile from decoys or other noise, and continue tracking.[62] Hypersonic systems reduce the amount of time available for all of these tasks. In the case of a ballistic missile, once the defender has identified it as such, it has a good idea of where the missile is going and can use that knowledge to cue midcourse and terminal defenses. Since hypersonic weapons are maneuverable, defenders are not sure which of their systems will be positioned to defeat the incoming missile.

However, early nuclear cruise missiles, such as the Snark, were vulnerable to antiaircraft fire.[63] After cruise missiles sank an Israeli destroyer in 1967, the US began to develop antimissile ship defense.[64] Air defenses motivated Britain to move from a bomber-based to a missile-based nuclear force.[65] Congress authorized the first US ballistic missile defense system in 1969.[66] ICBM designs and tactics dealt with the problem of survivability by incorporating decoys, chaff, alternate trajectories, radiation hardening, electronic countermeasures, and launch-on-warning postures.[67] Multiple independently targetable reentry vehicle (MIRV) systems packing multiple warheads onto a single missile were another response to missile defense technology.[68] These examples show how missiles and missile defense technology tended to coevolve.

From a historical point of view, hypersonic delivery systems for nuclear weapons can be viewed as a response to a long line of developments in the competition between missiles and missile defense. Technology advancements often provided an evolutionary technological edge, but the temporary advantage lasted only until a compensating technology was developed. While hypersonic weapons again promise that the missile will always get through, history suggests that new defenses against them will eventually thwart these new technologies. NAHWs are unlikely to prove revolutionary enough to catalyze the development of a new class of deterrence thinking.

This analysis assumed hypersonic delivery technology to be a linear combination of its constituent elements. This assumption is justified as there is a significant overlap in missile defense against cruise missiles and missile defense against ICBMs. This point is important because it suggests that hypersonic weapons are a combination of cruise missiles and ICBMs. Since defenses against both ICBMs and cruise missiles exist, it seems reasonable that defenses against hypersonics are quite possible.

According to the latest US *Missile Defense Review*, multiple missile defense systems are capable of defending against a mix of ballistic and cruise missile threats. For example, the Patriot PAC-3 missile defense system is capable of defending against cruise missiles and short-range ballistic missiles.[69] The F-35 is currently capable of defending against cruise missiles, and there are plans to include a capability to defend against boost-phase ballistic missiles.[70] The SM-6 missile of the Aegis system is also capable of defending against both ballistic and cruise missiles.[71] Since several existing systems can defend against cruise and ballistic missiles, it is reasonable to expect that future systems will be capable of defending against missiles that are a combination of the two.

The reason that existing missile defense systems can defend against cruise and ballistic missiles is that the dividing line between them is ambiguous.[72] While there are clear differences, there are significant similarities. For example, short-range ballistic missiles spend a significant portion of their time in the atmosphere and have more aerodynamic features than longer-range missiles that spend more of their flight time in space.[73] Furthermore, ballistic missiles and cruise missiles generally have a similar flight path in the terminal phase. This is important because many missile defense systems are designed to attack missiles in their terminal phase.

Finally, ballistic missiles do not always follow a completely ballistic trajectory. When Terminal High Altitude Area Defense missiles execute an energy management maneuver to burn fuel as required for short-range engagements, the missile executes a very non-ballistic loop.[74] MIRV weapons are designed to attack multiple targets, demonstrating that a limited maneuvering capability has previously been incorporated into ballistic missile delivery systems. Since maneuverability is possible in some ballistic missiles, maneuverable NAHWs can be viewed as an evolution in ICBM technology rather than as a revolution.

One final evidence of the evolutionary nature of NAHW technology is the fact that missile defenses for hypersonic weapons are already being developed. The 2019 US *Missile Defense Review* states that the US is currently working on developing systems to defeat hypersonic weapons.[75] DARPA's recently announced Glide Breaker project is one example.[76]

This analysis has shown that NAHWs will constitute an evolution rather than a revolution in technology. Missile defense technology is capable of defending against maneuverable weapons such as cruise missiles and high-speed threats such as ICBMs. Reasonably, it follows that missile defense technology may evolve to address weapons like NAHWs that combine both capabilities.

Implications for Both Sides Having Weapons

In terms of speed, range, accuracy, and missile defense, there seem to be few differences between both sides having NAHWs and both sides having significant numbers of ICBMs, SLBMs, and nuclear-armed cruise missiles. Missile defense is incapable of defeating the hundreds of nuclear weapons in the Chinese arsenal, much less the thousands of nuclear weapons possessed by the US and Russia. NAHWs do not increase the first-strike advantage against powers with large, diverse nuclear arsenals. Even if a hypersonic weapon successfully defeats existing missile defenses and

delivers a nuclear weapon, China, Russia, and the US can still deliver an overwhelming retaliatory strike.

Acton and others suggest that conventional hypersonic weapons introduce significant risks specific to these fast, maneuverable missile weapons.[77] The risk that a conventional attack is confused with a nuclear strike (warhead ambiguity) and the risk that a country mischaracterizes an attack on a neighbor as an attack on itself (destination ambiguity) should be considered in terms of the relative level of risk posed by other legs of the nuclear triad.[78] Deploying NAHWs while simultaneously deploying conventional hypersonic weapons may significantly exacerbate concerns regarding warhead and destination ambiguity. However, US bombers were capable of carrying both nuclear and conventional weapons as early as 1956.[79] The current US aircraft fleet also includes dual-capable aircraft.[80] The US believes that Russia has a "large, diverse, and modern" set of dual-capable weapon systems.[81] Likewise, bombers and cruise missiles can change course, meaning they have the potential for destination ambiguity, although their smaller speeds make this less of a concern. Dual-capable bombers have long been part of the strategic environment without proving hugely destabilizing, suggesting that warhead ambiguity may not be an issue for NAHWs.

Conclusion

This article examined whether a future NAHW can be understood by existing deterrence logic by considering hypersonic weapons as an evolution in nuclear weapon delivery technology. The analysis considered a NAHW to be a superposition of existing technologies analogous to hypersonic missiles: ICBMs, cruise missiles, and missile defense. The advancement of each of these systems was analyzed through the perspective of historical development and compared with unclassified information describing hypersonic systems. Key quantitative parameters such as range, speed, and accuracy were used alongside more qualitative data. Analyzing these analogous technologies suggests that the relevant elements of hypersonic technology will evolve slowly enough to remain consistent with existing thinking about nuclear deterrence.

There are several limitations of this study. First, operational data on hypersonic systems and their capabilities (range, speed, and accuracy) are not widely available because these systems are still under development and potentially classified. As more information on operational hypersonic systems comes available, it is possible that new systems may provide notable improvements relative to existing nuclear weapon delivery systems. Sec-

ond, by comparing hypersonic weapons to existing systems, we implicitly assumed that NAHWs will be used the same way as existing systems and ignored the possibility that NAHWs might be used differently. In novel applications, range, speed, accuracy, and avoiding missile defenses may not be paramount considerations. As Kissinger pointed out, technology is not everything. Instead, real advantages stem from "subtler and more discriminating uses rather than adding to [weapon] power or speed."[82] Perhaps this is the case with hypersonic weapons. This does not imply that hypersonic delivery systems are a useless military innovation. The risks of nuclear retaliation described by nuclear deterrence are more relevant to nuclear weapons than to conventional weapons.

However, policy makers do not have the luxury of choosing a development path based on a perfect, full-fledged knowledge of future fielded systems and how they will be used. Other authors have investigated the ways hypersonic technology may shape strategy and policy. For example, in their *War on the Rocks* commentary, Heather Venable and Clarence Abercrombie predict that hypersonic technology will face technological countermeasures and will not be destabilizing.[83] On the other hand, nuclear strategy and emerging technology researcher Alan Cummings argues that simply having the capability to launch rapid strikes may provide strategic advantages.[84] The Defense Intelligence Agency director predicts that "developments in hypersonic propulsion will revolutionize warfare by providing the ability to strike targets more quickly, at greater distances, and with greater firepower."[85]

Our analysis uses existing systems and operational concepts as a starting point to consider the policy implications of NAHWs. While NAHWs may evade missile defense systems in small numbers, ICBMs will probably defeat missile defense systems if used in large numbers.[86] NAHWs cannot prevent nuclear reprisal by Russia or China unless they are used as part of a massive first strike capable of destroying adversary second-strike capability. In terms of missile defense, a large salvo of NAHWs is nearly identical to a large salvo of ICBMs as both may overwhelm a missile defense system, suggesting NAHWs offer little advantage for large-scale nuclear strike missions relative to ICBMs or SLBMs. Even if revisionist powers possess highly robust defenses against existing ICBMs and nuclear-armed cruise missiles, their nuclear arsenals are too small to offer a credible second-strike capability, suggesting NAHWs offer little advantage.

One way that NAHWs may tangibly affect US deterrence policy is in their potential role for "tactical" nuclear weapons, otherwise known as low-yield nuclear weapons. The 2018 *NPR* discusses the need for having a flexible option and ensuring that there is no adversarial misperception about US capabilities. Hypersonic nuclear weapon delivery systems may provide advantages for delivering tactical nuclear weapons.[87] Consider a scenario in which a small "tactical" nuclear warhead is employed to destroy a difficult target with only minimal collateral damage. If nuclear weapons are viewed as gargantuan classical bombs, tactical nuclear weapons are simply another form of war fighting in which accuracy is extremely important. Cruise missiles are more accurate than ICBMs, and so cruise missiles might be more useful than ICBMs for highly precise nuclear strikes. Since cruise missiles are more vulnerable to missile defenses than are hypersonic delivery systems, a NAHW may be advantageous relative to existing weapons for tactical nuclear strikes.

However, nuclear weapons are not simply bigger, more effective conventional bombs.[88] Tactical nuclear weapons carry a risk of nuclear retaliation if used against a nuclear-armed adversary. Extended deterrence suggests there is a risk of nuclear retaliation if nuclear weapons are used against a nation allied to a nuclear power. The risk of retaliation by a nuclear superpower is not mitigated by using a NAHW to "guarantee" successful delivery of a small number of tactical nuclear weapons because the arsenals of the world's great powers are probably too large and diverse to allow a successful first strike.[89] Using a NAHW against a nation possessing a small nuclear arsenal only provides an advantage if the adversary also retains an effective missile defense system. Otherwise, nuclear-armed cruise missiles or nuclear-armed ballistic missiles are presumably equally as effective as NAHWs.

There appears to be little advantage to upgrading the existing US nuclear arsenal to include hypersonic delivery systems. Based on the research conveyed in this article, one conclusion is that there is little advantage to upgrading the existing US nuclear arsenal to include hypersonic delivery systems as their advantages in speed, range, and accuracy are on the margins and the monies required may be better used elsewhere. A NAHW provides few advantages relative to cruise missiles or ICBMs in terms of speed, range, or accuracy. While hypersonic delivery systems appear to provide some capability to defeat missile defense systems, this potential advantage may only be temporary—especially if current efforts to develop missile defenses against hypersonic weapons continue. SSQ

Lt Col Nathan B. Terry, PhD

Colonel Terry is a physicist with the United States Air Force. His previous assignments include operations analyst, research physicist, assistant professor at the United States Air Force Academy, and branch chief in the Air Force Research Laboratory.

Paige Price Cone

Dr. Cone is an assistant professor at the Air War College's Center for Strategic Deterrence Studies where she teaches courses on nuclear deterrence and arms control. Prior to her appointment, she was a postdoctoral research fellow at The University of Chicago and the first nuclear proliferation fellow at the Chicago Project on Security and Threats. Her research and teaching interests are in international security, with a specific focus on issues of nuclear deterrence and nonproliferation.

Notes

1. John A. Tirpak, "The Great Hypersonic Race, *Air Force Magazine*, 27 June 2018, http://www.airforcemag.com/.

2. John L. Dolan, Richard K. Gallagher, and David L. Mann, "Hypersonic Weapons: A Threat to National Security," *Real Clear Defense*, 23 April 2019, https://www.realcleardefense.com/.

3. Former CJCS Martin Dempsey said that the only way to defeat large salvos of ICBMs is through preemptive strikes, implying that missile defense is inadequate to deal with large salvos of missiles. Martin E. Dempsey, *Joint Integrated Air and Missile Defense: Vision 2020* (Washington, DC: The Joint Staff, 2013), 3, https://www.jcs.mil/. The *2019 Missile Defense Review* states that the US is "protected against a limited ICBM attack." Department of Defense, *2019 Missile Defense Review* (Washington, DC: Office of the Secretary of Defense, 2019), X, https://www.defense.gov/. President Bush viewed ballistic missile defense as necessary to protect the United States against rogue nations and terrorists, emphasizing the utility of ballistic missile defense against small salvos of ICBMs. See, for example, Terence Neilan, "Bush Pulls Out of ABM Treaty; Putin Calls Move a Mistake," *New York Times*, 13 December 2001, https://www.nytimes.com/. Lawrence Freedman makes a similar observation in his book *Deterrence* (Malden, MA: Polity Press, 2004), 37.

4. Richard H. Speier, George Nacouzi, Carrie A. Lee, and Richard M. Moore, *Hypersonic Missile Proliferation: Hindering the Spread of a New Class of Weapons* (Santa Monica, CA: RAND Corporation: 2017), xii, https://www.rand.org/; and Kyle Mizokami, "China Conducts New Hypersonic Weapons Test," *Popular Mechanics*, 7 August 2018, https://www.popularmechanics.com/.

5. James Acton, *Silver Bullet? Asking the Right Questions about Conventional Prompt Global Strike* (Washington, DC: Carnegie Endowment for Peace, 2013), https://carnegieendowment.org/; Eleni Ekmektsioglou, "Hypersonic Weapons and Escalation Control in East Asia," *Strategic Studies Quarterly* 9, no. 2 (Fall 2015): 43–68, https://www.airuniversity.af.edu/; Joshua H. Pollack, "Boost-Glide Weapons and US-China Strategic Stability," *Nonproliferation Review* 22, no. 2 (2015): 155, published online February 2016, https://doi.org/10.1080/10736700.2015.1119422; and Speier et al., *Hypersonic Missile Proliferation*, iii.

6. Ekmektsioglou examines how China might perceive US conventional hypersonic weapons as a threat to its nuclear forces and discusses ways China might respond. Speier

et al. also consider the impacts of proliferating hypersonic technology. Acton considers the relevant missions of conventional hypersonic weapons, as well as various modes of instability they might prompt. Ekmektsioglou, "Hypersonic Weapons," 43–68; and Speier et al., *Hypersonic Missile Proliferation*; and Acton, *Silver Bullet?*

7. Acton, *Silver Bullet?*, 5–7; Ekmektsioglou, "Hypersonic Weapons," 43; and Pollack, "Boost-Glide Weapons," 155. However, this definition technically includes ICBMs and SLBMs as hypersonic weapons because they travel faster than Mach 5. See, for example, Office of the Secretary of Defense, *Nuclear Posture Review* (Washington, DC: Department of Defense, 2018), 45, https://media.defense.gov/.

8. Andrew Knoedler, "Hypersonic Air-Breathing Weapon Concept (HAWC)," Defense Advanced Research Projects Agency, accessed June 2019, https://www.darpa.mil/.

9. We examined the rate of change of specific technology areas by reviewing the number of publications reporting advances in each technology as a function of time. We considered three technology groups: scramjet engines, high-temperature hypersonic materials, and hypersonic aerodynamic structures. Beginning with 1960 data, we counted the number of journal publications, patents, and conference materials for each of the six search term sets. We chose 1960 as the start date to correspond to the ICBM and cruise missile data shown later in the article. Advanced manufacturing was not included because this technology is broad enough to apply to many other technologies. We queried the EBSCO database using the following six sets of search terms: (1) "hypersonic," (2) "hypersonic flight," (3) "high temperature" and "hypersonic" and "material," (4) "hypersonic" and "vehicle configuration," (5) "scramjet" or "supersonic ramjet" or "hypersonic propulsion," and (6) "hypersonic aerodynamics." The first two search terms were included to account for general trends in overall hypersonic research.

10. For a brief description of the X-20 project, see Jay Miller, *The X-Planes, X-1 to X-45* (Midland, MI: Midland Publishing, 2001), 231–39; for a brief description of the X-24 project, see 261–71. Examining the number of patents filed referencing the aforementioned search terms shows a similar story. A similar analysis of the record of conference materials shows a more ambiguous story. However, journal articles are more likely to be peer reviewed than conference materials. Patents are expensive to acquire and renew. Therefore, journal publications and patents probably show a more reliable summary of the evolution of hypersonic component technology than conference materials. There are limitations to analyzing this type of metadata because the search technique does not assess the relative quality of the research. Large technology advances are weighted equally with small advances, and flawed analyses are counted the same as sound ones. The data shows significant previous interest in hypersonic technology in the 1960s and 1970s. These short but intense explosions in research correspond to the development of the X-20—also known as the Dyna-Soar project—and the X-24. The X-20 was cancelled in December 1963, and the X-24's last flight was in 1975. Even though the X-20 was never built into a flight-test-capable model and the X-24 never achieved speeds of greater than Mach 1.8, both projects tremendously advanced the state of the art of hypersonic research. Developments of hypersonic technology over the last two decades may be the result of researchers building off a single large technology breakthrough. Or this increase in research could simply be mirroring greater government interest and/or funding.

11. Bernard Brodie, "The Absolute Weapon: Power and World Order," in *War in the Atomic Age*, ed. Bernard Brodie (New Haven, CT: Yale University Institute of International Studies, 1946), 25. Brodie cites a post–World War II survey study suggesting it

would have taken 730 bombers to do the same damage done by the one nuclear bomb dropped on Hiroshima.

12. Brodie, 76.

13. Thomas C. Schelling, *Arms and Influence* (New Haven, CT: Yale University Press, 2008), 24–27.

14. Glenn H. Snyder, *Deterrence and Defense: Toward a Theory of National Security* (Princeton: Princeton University Press, 1961), 15; see also Brodie, "The Absolute Weapon," 76.

15. Schelling, *Arms and Influence*, 166–68. Kissinger reached a similar conclusion; see Henry Kissinger, *Nuclear Weapons and Foreign Policy* (New York: Harper & Brothers, 1957), 48. Rosemary Foot concedes that while nuclear deterrence wasn't central to ending the conflict, nuclear threats were a significant factor. Rosemary J. Foot, "Coercion and Ending the Korean War," *International Security* 13, no. 3 (Winter 1988–89): 92–112.

16. Robert Anthony Pape, *Bombing to Win: Air Power and Coercion in War* (Ithaca, NY: Cornell University Press, 1996), 6–11.

17. Kissinger, *Nuclear Weapons*, 13–14.

18. Dietrich Schroeer, *Science Technology and the Nuclear Arms Race* (New York: John Wiley & Sons, 1984), 110–12, 115.

19. See figure 2 and its associated references.

20. Schroeer, *Science Technology*, 151. See also Ronald Huisken, *The Origin of the Strategic Cruise Missile* (New York: Praeger Publishers, 1981), 16; and Keir A. Lieber and Darryl G. Press, "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence," *International Security* 41, no. 4 (Spring 2017): 24, https://doi.org/.

21. Huisken, *Origin of the Strategic Cruise Missile*, 9–10, 16–17; and Ted Nicholas and Rita Rossi, *U.S. Missile Data Book*, vol. 1, 36th ed. (Huntington Beach, CA: Data Research Associates, 2012), tables 1-2 and 3-2.

22. Schroeer, *Science Technology*, 114–15.

23. Schroeer, 109.

24. Thomas G. Mahnken, *The American Way of War* (New York: Columbia University Press, 2008), 30; and Schroeer, *Science Technology*, 107.

25. Schroeer, 201.

26. Lieber and Press, "New Era of Counterforce," 16.

27. Lieber and Press, 10.

28. See Ronald L. Tammen, *MIRV and the Arms Race: An Interpretation of Strategy* (New York: Praeger, 1973), 85–86.

29. For a review of how deterrence thinking evaluated missile defense, see Ashton B. Carter and David N. Schwartz, eds., *Ballistic Missile Defense* (Washington, DC: Brookings Institution, 1984).

30. Kissinger, *Nuclear Weapons*, 17–18, 120.

31. Kissinger, 119.

32. Kissinger, 126.

33. Kissinger, 120. Herman Kahn made similar arguments that a "symmetrical" use of nuclear weapons would be more advantageous for China than for the US. See Herman Kahn, *On Thermonuclear War* (Piscataway, NJ: Transaction Publishers, 2007), 505.

34. Kissinger, *Nuclear Weapons*, 122–23.

35. Kissinger, 124, 185, 199.

36. Kissinger, 120.

37. Kissinger, 123.

38. Kissinger, 16–17.

39. Colin Gray, "Nuclear Deterrence and Technological Change: Retrospect and Prospect," in *On the Logic of Nuclear Terror*, ed. Roman Kolkowicz (Winchester, MA: Allen & Unwin, Inc., 1987), 155–85.

40. Office of the Secretary of Defense, *Nuclear Posture Review*, II, VII.

41. Peter J. Mantle, *The Missile Defense Equation* (Reston, VA: American Institute of Aeronautics and Astronautics, 2004), 10.

42. Acton points out the range of some planned hypersonic weapons is at least 5,000 miles Acton, *Silver Bullet?,* 78.

43. Mantle, *Missile Defense Equation*, 10.

44. Lieber and Press, "New Era of Counterforce," 16.

45. For a review of ballistic missile defense technology, see National Research Council, *Making Sense of Ballistic Missile Defense: An Assessment of Concepts and Systems for U.S. Boost-Phase Missile Defense in Comparison with Other Alternatives* (Washington, DC: National Academies Press, 2012), https://www.nap.edu/; and Ashton B. Carter, "BMD Applications: Performance and Limitations," in Carter and Schwartz, *Ballistic Missile Defense*, 124–29.

46. The graph was developed from Nicholas and Rossi, *U.S. Missile Data Book*, tables 1-2 and 3-2; Schroeer, *Science Technology*, 176; and Jonathan E. Medalia, *Strategic Forces: MX ICBM* (*Weapons Facts*), Report no. 1884046 [Washington, DC: Congressional Research Service, August 1987], 1. Hypersonic data comes from US Air Force, "X-51A Waverider," fact sheet, 2 March 2011, https://www.af.mil/.

47. SLBMs travel at roughly the same speed as ICBMs but can be fired from much closer, suggesting they can be significantly more prompt than ICBMs.

48. Acton, *Silver Bullet?*, 68–70.

49. Speier et al., *Hypersonic Missile Proliferation*, 16. Other studies estimate nine minutes of warning time. See Lisbeth Gronlund and David C. Wright, "Depressed Trajectory SLBMS: A Technical Evaluation and Arms Controls Possibilities," *Science and Global Security* 3, nos. 1–2 (1992): 110, 117.

50. Kissinger, *Nuclear Weapons*, 119.

51. The graph was developed from Nicholas and Rossi, *U.S. Missile Data Book*, tables 1-2 and 3-2; and Schroeer, *Science Technology*, 176. However, Nicholas and Rossi report the speed of the Peacekeeper as Mach 4, inconsistent with information provided by other sources. See National Museum of the US Air Force, "Martin Marietta LGM-118A Peacekeeper," fact sheet, 13 July 2015, https://www.nationalmuseum.af.mil/. The database "Military Periscope" also reports the speed as approximately Mach 20. See "Military Periscope," accessed April 2019, https://www.militaryperiscope.com/. Both estimates convert to a Mach number by assuming the speed of sound is approximately 767 miles per hour. A speed of Mach 20 is used for the graph. For hypersonic data, see US Air Force, "X-51A Waverider."

52. Schroeer, *Science Technology*, 202.

53. Schroeer, 144.

54. For example, the Russian SS-11 (mod 2) had a range of 13,000 km in 1973. In 1988, the SS-18 (mod 6) had a range of 16,000 km. Several other variants including the SS-17, SS-19, SS-24, SS-25 and several variants of the SS-18 had shorter ranges than the SS-11 (mod 2). The range difference between SBLBs increased from 7,800 km in

1973 (SS-N-8) to 8,300 km in 1985 (SS-N-23). Again, several intermediate variants had less range than the SS-N-8. The accuracy of Russian ICBMs decreased from 1,675 meters CEP in 1973 (SS-13 mod 2) to 200 meters in 1987 (SS-24). See Thomas B. Cochran, William M. Arkin, Robert S. Norris, and Jeffrey I. Sands, *Nuclear Weapons Databook*, vol. 4, *Soviet Nuclear Weapons* (New York: Harper & Row, 1989), 111, 114; and Schroeer, *Science Technology*, 56, 63–65, 68–72.

55. The graph was developed from Nicholas and Rossi, *U.S. Missile Data Book*, tables 1-2 and 3-2; and Schroeer, *Science Technology*, 176. For information on the accuracy of the Atlas missile including CEP, see "SM-65 Atlas," GlobalSecurity.org, accessed 24 February 2019, https://www.globalsecurity.org/. Note the *Air and Space Power Journal* (*ASPJ*) article citing the accuracy of 10 miles but not indicating whether this indicates CEP. See *ASPJ* staff, "Atlas: The Godfather of ICBMs and Space-Launch Vehicles," *Air and Space Power Journal* 17, no. 1 (Spring 2003): 68, https://www.airuniversity.af.edu/. For cruise missile data, see National Museum of the US Air Force, "Martin TM-61A Matador," fact sheet, 29 May 2015, https://www.nationalmuseum.af.mil/; National Museum of the US Air Force, "Martin CGM-138 Mace," 29 May 2015, https://www.nationalmuseum.af.mil/; US Air Force, "AGM-86B/C/D Missiles," fact sheet, 24 May 2010, https://www.af.mil/; and US Air Force, "AGM-130 Missile," fact sheet, 18 June 2003, https://www.af.mil/. For hypersonic data, see US Air Force, "X-51A Waverider."

56. In one way, this is not surprising because the physics of ballistics make the range physically dependent on the speed of the missile. For a discussion of this, see, for example, app. C in Mantle, *Missile Defense Equation*.

57. The graph was developed from US Air Force, "AGM-86B/C/D Missiles"; US Air Force, "AGM-130 Missile"; IHS Jane's, *IHS Jane's Weapons: Strategic* (Coulsdon, Surrey: IHS Global, 2013), 155–56; and Missile Defense Project, "JASSM/JASSM ER (AGM-158A/B)," *Missile Threat*, Center for Strategic and International Studies, 6 October 2016, last modified 15 June 2018, https://missilethreat.csis.org/

58. See US Air Force, "AGM-86B/C/D Missiles"; and US Air Force, "AGM-130 Missile."

59. The graph was developed from National Museum of the US Air Force, "Martin TM-61A Matador"; Museum of the US Air Force, "Martin CGM-13B Mace," fact sheet, 29 May 2015, https://www.nationalmuseum.af.mil/; US Air Force, "AGM-86B/C/D Missiles"; and US Air Force, "AGM-130 Missile." For hypersonic data, see US Air Force, "X-51A Waverider."

60. The graph was developed from Nicholas and Rossi, *U.S. Missile Data Book*, tables 1-2 and 3-2; US Air Force, "AGM-86B/C/D Missiles"; National Museum of the US Air Force, "Martin TM-61A Matador"; and National Museum of the US Air Force, "Martin CGM-13B Mace."

61. Brodie, "The Absolute Weapon," 25.

62. Mantle, *Missile Defense Equation*, 174–75.

63. Huisken, *Origin of the Strategic Cruise Missile*, 16–17. While cruise missiles are not hypersonic, they are maneuverable. Maneuverability is what makes hypersonic weapons different from ICBMs.

64. Huisken, 29.

65. Lawrence Freedman, "The Small Nuclear Powers," in Carter and Schwartz, *Ballistic Missile Defense*, 253.

66. Schroeer, *Science Technology*, 236.

67. See Tammen, *MIRV and the Arms Race,* 85–86.

68. Alexey Arbatov and Vladimir Dvorkin, "The Impact of MIRVs and Counterforce Targeting on the US-Soviet Strategic Relationship," in *The Lure and Pitfalls of MIRVs*, eds. Wheeler Krepon, Travis Wheeler, and Shane Mason (Washington, DC: Stimson Center, 2016), 70–71, https://www.stimson.org/; Jeffrey G. Lewis, "China's Belated Embrace of MIRVs," in Krepon, Wheeler, and Mason, *Lure and Pitfalls of MIRVs*, 95, 100, 110; and Lawrence Freedman, "The First Two Generations of Nuclear Strategists," in *Makers of Modern Strategy: From Machiavelli to the Nuclear Age*, eds. Peter Paret, Gordon Alexander Craig, and Felix Gilbert (Princeton, NJ: Princeton University Press, 1986), 759–60. While Tammen suggests MIRVs are a response to ballistic missile defense technology, he indicates there were other motivations as well. Tammen, *MIRV and the Arms Race*, 104, 107, 113–14, 137.

69. See Department of Defense, *2019 Missile Defense Review*, 50, where air-breathing weapons refer to cruise missiles.

70. Department of Defense, 55.

71. Department of Defense, 48, 53.

72. Mantle, *Missile Defense Equation*, 8–16.

73. Mantle, 11–12.

74. Mantle, 13.

75. Department of Defense, *Missile Defense Review*, XIV, XVIII.

76. Tom Porter, "What Are Hypersonic Weapons and Why Is DARPA So Concerned about Stopping Them?," *Newsweek*, 13 November 2018, https://www.newsweek.com/.

77. In addition to the Acton references later in this paragraph, see Speier et al., *Hypersonic Missile Proliferation*, 8.

78. Acton, *Silver Bullet?*, 111–18.

79. "Telegram from the North Atlantic Council Ministerial Meeting to the Department of State," 14 December 1956, doc. 51, in *Foreign Relations of the United States, 1955–1957,* vol. 4*, Western European Security and Integration*, ed. William Z. Slany (Washington, DC: Government Printing Office, 1988), https://www.history.state.gov/.

80. Office of the Secretary of Defense, *Nuclear Posture Review*, II, X, 16, 49.

81. Office of the Secretary of Defense, 49.

82. Kissinger, *Nuclear Weapons*, 120.

83. Heather Venable and Clarence Abercrombie, "Muting the Hype over Hypersonics: The Offense-Defense Balance in Historical Perspective," *War on the Rocks*, 28 May 2019, https://warontherocks.com/.

84. Alan Cummings, "Hypersonic Weapons: Tactical Uses and Strategic Goals," *War on the Rocks*, 12 November 2019, https://warontherocks.com/.

85. Robert Ashley, "Statement for the Record: Worldwide Threat Assessment," Statement before Senate Armed Services Committee, Washington, DC, 6 March 2018, https://www.dia.mil/.

86. See Dempsey, *Vision 2020*, 3; Department of Defense, *2019 Missile Defense Review*, X; Neilan, "Bush Pulls Out of ABM Treaty"; and Freedman, *Deterrence*, 37.

87. See Kahn, *On Thermonuclear War*, 16. Others who believe nuclear weapons are useful for war fighting include Russian and US military planners who incorporate low-yield nuclear weapons into their military arsenals. See Office of the Secretary of Defense, *Nuclear Posture Review*, 22, 53–54. Freedman argued tactical nuclear weapons could be useful for war fighting by attacking logistical or other forces deep inside enemy territory;

see *The Evolution of Nuclear Strategy*, 106–8. Other scholars argue nuclear weapons have little military utility for directly attacks enemy military forces. For example, see Pape, *Bombing to Win*, 13–15, 49. Tactical nuclear weapons delivered via maneuverable hypersonic weapons will not prove any more or less destabilizing than tactical nuclear weapons delivered by other means.

88. For example, nuclear weapons are considered a weapon of mass destruction, distinguishing them from conventional weapons.

89. One counterargument is Ashton Carter's "ragged defense" theory, suggesting that missile defense provides a first-strike advantage because second-strike forces will be decimated so badly that their ragged response will be effectively defeated by the aggressor's missile defense system. See Ashton B. Carter, "Introduction to the BMD Question," in Carter and Schwartz, *Ballistic Missile Defense*, 22. Along similar lines, Kahn argued increased defenses or evacuations provide first strikes an increased advantage. Kahn, *On Thermonuclear War*, 14. However, as the adversary nuclear force becomes larger and more diverse, the risk of a successful first strike decreases, mitigating any theorized advantages posited by the ragged defense theory.

PERSPECTIVE

Power, Profit, or Prudence? US Arms Sales since 9/11

A. Trevor Thrall
Jordan Cohen
Caroline Dorminey

Abstract

By law and executive branch policy, every decision to sell American weapons abroad must reflect an assessment of strategic interests, economic considerations, and risk. Little work, however, has been done to determine how much relative influence each of these factors has on such decisions. This article evaluates arms sales in the post-9/11 era and finds evidence that strategic interests and economic considerations significantly impact arms sales but no evidence that risk assessment does so. It concludes with suggestions about how to better incorporate risk assessment by making the cost/benefit trade-offs more explicit in the arms sales decision process.*

*****

From 2002 through 2018 the United States sold over $200 billion in major conventional weapons to 169 countries. Thirty-one of those countries purchased at least $1 billion in arms.[1] By 2018 the United States extended its dominance as the leading exporter of weapons with a 36 percent share of the global market compared to 21 percent for Russia and 6.8 percent for France, the second- and third-ranked exporters.[2]

Previous research into US arms sales finds that they are driven mainly by strategic and economic factors.[3] The conventional view among international relations scholars is that strategic considerations loom largest. Writing about American decisions whether to provide alliance commitments or arms sales to client states, for example, Keren Yahri-Milo, Alexander Lanoszka, and Zack Cooper argue that "U.S. decisionmakers focused primarily on the commonality of security interests and the local military balance in determining which bundles of military assistance to give client states."[4] This view also appears to prevail in Washington, where

*The authors would like to thank Mike Guillot for his editorial guidance and several anonymous reviewers for their help in improving the final manuscript.

policy makers see arms sales as a valuable tool of foreign policy that can strengthen the military capability of allies, leverage the behavior of recipients, and promote regional stability in critical areas worldwide.[5]

The main competitor to the strategic narrative is the argument that economic considerations play a more decisive role in determining arms sales, at times outweighing strategic considerations.[6] Though selling weapons to other governments is a matter of US policy, the revenues from arms sales accrue directly to American companies. Presidents and members of Congress are well aware that arms sales represent jobs and corporate profits for Americans and American companies. President Trump has repeatedly echoed this rationale publicly. Reflecting on a major deal with Saudi Arabia, he stated, "It will create hundreds of thousands of jobs, tremendous economic development, and much additional wealth for the United States."[7] Trump's desire to create wealth from arms sales has been more transparent than most, but every president since Clinton has made clear in policy documents that economic benefits are a key consideration of the arms sales approval process.

Another factor receiving little explicit discussion in previous work about government decision-making is risk sensitivity—specifically, how risk is measured. All arms transfers, including those with important strategic rationales, carry the risk of negative strategic and humanitarian consequences. Recent history presents many examples. These cases run the gamut from weapons falling into the hands of criminals, terrorist groups, or rogue regimes; to weapons being used by recipient governments to commit human rights violations against their own people; to accidentally amplifying conflicts and arms races; and in several cases, to American weapons being used against American troops on the battlefield.[8] Every arms sale is thus a calculated gamble that the expected strategic and economic benefits will outweigh the potential costs. The risk does not have to be zero, but any negative consequences must be low enough to make the decision worth it in the long run.

In recognition of the risks of arms transfers, the US government has taken steps to tilt the odds in favor of positive outcomes—at least on paper. In 1976 Congress passed the Arms Export Control Act requiring the executive branch to conduct a risk assessment before approving a major arms sale to ensure that it "would not contribute to an arms race, aid in the development of weapons of mass destruction, support international terrorism, increase the possibility of outbreak or escalation of conflict, or prejudice the development of bilateral or multilateral arms control or nonproliferation agreements or other arrangements."[9]

The United States amplified the importance of risk sensitivity in the arms sales process when in 1997 Congress passed the first version of the Leahy Law, designed to prohibit security assistance—including arms transfers—to military units for which there was credible information implicating them in the commission of gross violations of human rights.[10] Since then the law has been expanded, becoming a permanent part of the Foreign Assistance Act in 2008. More generally, presidents from Jimmy Carter onward have all issued policy directives emphasizing the care that the United States must take when considering selling weapons abroad. The Trump administration's update of the Conventional Arms Transfer Policy, for example, adds new emphasis on preventing civilian casualties caused by American weapons in the hands of its clients.[11]

Though it makes sense to imagine that all three factors play a role in decisions about how much to sell to which countries, we are aware of no study attempting to assess their relative importance. How strategic is the arms export process? Other things being equal, one expects arms sales to be higher where the strategic value is high. It is less certain, on the other hand, what to expect in the absence of clear strategic benefits. In such cases, is the promise of economic benefits enough for the United States to approve arms sales even where the risks of negative consequences are significant? The existing literature offers little insight into how much impact the risk assessment process has on decisions to approve arms sales. One can imagine that pressing strategic considerations will overrule concerns about risk, especially in cases where the risks are long term, but how much does risk matter in cases where the only benefits are economic?

Based on our analysis of a range of strategic, economic, and risk factors, we find considerable evidence that strategic and economic factors drive arms sales, but no evidence to suggest that risk plays a meaningful role. Downside risks are rarely considered explicitly or appear to have ever affected the transfer of weapons. Moreover, by our measures, the United States has taken increasingly higher risks from 2001 to today. Our findings raise important questions about how well the United States is managing the trade-offs between strategy, economics, and risk. Critics have long argued that the United States is too quick to approve sales to nondemocratic clients with poor track records of human rights. Efforts in Congress to stop arms sales to Saudi Arabia over its intervention in the Yemen civil war are a recent illustration of the policy relevance of the debate.[12] And despite the new emphasis on preventing civilian casualties, the Trump administration has also highlighted economic security as a justifi-

cation for arms sales and is working to increase the federal government's role in promoting arms sales around the world.[13]

This article starts with a brief summary of patterns in US arms sales decisions and then develops our expectations about the roles of strategic, economic, and risk factors in decision-making. Next, we offer our analysis and findings and consider some potential objections. Finally, we present the arms sales risk-reward matrix to help policy makers consider the trade-offs between strategic considerations and risk.

Explaining Patterns of American Arms Sales

The United States government regulates the export of all weapons, which fall into one of three broad classes using 15 categories in the United States Munitions List (USML). The first class incorporates those things that the United States simply does not allow companies to sell to foreign customers. This includes anything having to do with nuclear, biological, and chemical weapons, in accordance with various international treaties. This class also includes some high-end weapons technology that is prohibited from export to preserve America's qualitative edge. The most visible system on this list is the Lockheed Martin F-22 Raptor. Advanced drones, until very recently, were also seen as too technologically sensitive to allow for widespread export.[14]

The second class includes small and light weapons (SALW), along with certain types of ammunition and equipment, that the government allows companies to sell directly to foreign customers with minimal government intervention through direct commercial sales (DCS). Even though small-arms sales are not always approved, critics have complained about the devastating effects of lightly regulated sales of small arms abroad. The implicit presumption behind these sales is that these weapons are unlikely to spawn large-scale negative consequences and thus need no risk assessment and only minimal oversight. SALWs exported via DCS fall in categories 1 through 3 of the USML, everything from ammunition for close assault weapons to flamethrowers.[15] In 2018, US companies sold almost $30 billion of these weapons to 29 different nations.[16]

The third class includes what are categorized as major conventional weapons (MCW). Categories 4 through 15 of the USML include planes, tanks, ships, missiles, and everything in between. Since World War II the United States has been the dominant exporter of MCWs, and its 36 percent global market share attests to this.[17] This weapons category not only is the greatest influencer of the global balance of power but also is the crux

of arguments about the strategic nature of arms sales. Thus, this category is the focus of our article.

For a country to purchase MCWs from the United States, it must first place a request with the Pentagon or State Department. The two agencies then work together to assess the strategic and political implications of the proposed purchase. The risk assessment required by the Arms Export Control Act is carried out by country-level teams, after which the administration makes a final decision. If approved, the State Department issues a notification of the sale to Congress.[18] Congress has between 15 and 30 days (depending on the recipient) to review the sale and, if there is opposition, to pass a resolution in both chambers to block the sale. After the review period has passed, the sale becomes official and the delivery of weapons can take place. To date, Congress has passed just one bill in an effort to block a sale (to Saudi Arabia in 2019) since the Arms Export Control Act was passed in 1976. However, it has used the threat of blocking a sale to alter the terms of a deal on a few other occasions. This record of approval seems to nullify any concern over the risk of inaction (non-sale). For perspective, Congress has been notified of 1,970 arms sales, 707 of which were of major conventional weapons.[19]

The Logics of Arms Sales: Power, Profit, and Prudence

As noted, the two most prominent explanations for the pattern of American arms sales are strategic and economic considerations. The role of risk assessment has received considerably less attention despite US law and policy. Making the debate interesting are the trade-offs and tensions within the process. Sole focus on maximizing any one of the goals in the arms sales process would require making concessions on other goals. Selling weapons without any concern for who receives them in the attempt to maximize economic benefits, for example, would inevitably raise the risk of negative outcomes and confound the pursuit of strategic goals. Minimizing risk, on the other hand, would limit the ability to pursue economic and strategic gains. By definition, then, US arms sales reflect these trade-offs, whether made strategically or by default.

Strategic Considerations

From a strategic perspective, arms sales have many purposes, but their impact occurs through two basic mechanisms: shifting the local or regional balance of power in favor of American interests and exerting leverage over the conduct of recipient nations.[20] By increasing the military ca-

pabilities of the recipient nation, arms sales can—in theory—help allies win wars, deter adversaries, and promote stability or buttress friendly governments against insurgencies and other internal challenges. The attractiveness of arms sales in these cases stems primarily from the fact that selling weapons to allies is less risky and less costly than basing American troops on foreign soil or having them fight alongside allies, though they also allow the United States to signal intentions to both allies and potential adversaries.[21] American foreign policy and the flow of American weapons indicate that the strategic logic of arms sales often played an important role during the Cold War. Throughout the Cold War, the United States used arms sales as one element of its strategy for defending Western Europe and containing the Soviet Union around the globe.[22]

Arms sales are also widely believed to provide the United States leverage over recipients, especially those who rely heavily on the United States for their military capability. The United States has used arms sales to gain access to overseas military bases, pressure countries to vote with the United States at the United Nations, discourage conflict, and encourage domestic political reforms. Andrew Shapiro, former assistant secretary of state for political-military affairs, notes that "when a country acquires an advanced U.S. defence system, [it is] not simply buying a product to enhance [its] security, [it is] also seeking a relationship" with the United States. . . . This engagement helps build bilateral ties and creates strong incentives for recipient countries to maintain good relations with the United States."[23]

There is good reason to expect that strategic considerations play an important role in determining where the United States is willing to sell major conventional weapons. And indeed, the historical record suggests that strategic motivations were a powerful driver of arms sales during the Cold War. US arms sales between 1950 and 1991 were restricted to allies and other nations the United States believed were useful partners in the struggle with the Soviet Union.[24] After the Cold War ended, however, the United States began selling weapons to a much broader set of customers, including nations formerly part of the Soviet bloc as the global strategic landscape changed.[25] In the wake of 9/11, the war on terror has been a major driver of American foreign policy, and arms sales patterns have shifted yet again as a result.

Economic Considerations

There is little doubt among scholars or policy makers that the pursuit of profit and other economic benefits has always driven arms sales. For advocates, the economic benefits from arms sales are obvious and come in the

form of exports, employment, economies of scale, and the general health of the defense industrial base.[26] The United States has long been the world's leading arms exporter.[27] Since 2002, American arms exports have ranged between $15 and $75 billion per year, representing as much as 4 to 5 percent of total American exports annually. The civilian defense industry employs almost two million people, and though American military spending provides the majority of revenue for most defense firms, exports can help sustain jobs and keep plants open. Arms sales advocates also argue that exports help lower costs for the Pentagon. By increasing the total number of orders for expensive weapons systems like the F-35, exports can in theory shrink the per unit cost for the US military and lower downstream costs for spare parts. Others also note that competing in the global market can help preserve the innovative capability and financial well-being of American defense firms.[28]

Economic motives for pursuing arms exports have been more explicit since the end of the Cold War.[29] For arms sales advocates, increased exports grew more important as the United States military budget drew down and domestic procurement of big-ticket items began to shrink.[30] Boosting arms exports also fits neatly with the Clinton administration's focus on the economy more generally. Presidents Bush and Obama continued in the same vein, with arms sales increasing considerably after the terrorist attacks of 9/11. In 2010, discussing reforms to arms export regulations, President Obama stated that "by enhancing the competitiveness of our manufacturing and technology sectors, they'll help us not just increase exports and create jobs, but strengthen our national security as well."[31] The most explicit statement of the importance of economic motives appeared in the Trump administration's update of the US Conventional Arms Transfer Policy in 2018, asserting that "when a proposed transfer is in the national security interest, which includes our economic security, and in our foreign policy interest, the executive branch will advocate strongly on behalf of United States companies."[32]

In short, the question is not whether economic motives affect US arms sales decisions, but how powerful they are relative to other considerations. Though up until the Trump administration American policy has always articulated the importance of strategic considerations, critics complain that strategy often seems to take a back seat to the profit motive. The defense industry spends a great deal of time and money lobbying Washington's policy makers to keep arms sales flowing. Beyond millions in campaign contributions and other soft-money contributions to both parties, the defense industry has worked hard to make sure the most visible benefits

of arms exports—factories and the jobs that go with them—are well distributed in congressional districts throughout the country.[33] The result, critics argue, is a tendency to sell weapons to almost any nation that wants them, regardless of whether the United States has a strong strategic interest in doing so and what risks might be associated with the sales.

Risk Considerations

Arms sales can generate undesirable strategic and humanitarian effects on three levels: direct negative consequences for the United States like blowback and entanglement; consequences for the buyer's neighborhood such as the dispersion of weapons, arms races, and increased instability; and consequences for the purchasing nation itself such as increased levels of corruption, social violence, human rights abuses, and civil conflict.[34]

History suggests that these risks are not simply far-fetched possibilities. A more common example is when American troops end up fighting other forces armed with American-made weapons that the United States had willingly provided, as happened in Somalia in 1991 with weapons exported during the Cold War.[35] Tens of billions of dollars in arms sales to Saudi Arabia and the UAE have also enmeshed the United States in the intervention in Yemen.[36] An extreme example of blowback is the 1979 Iranian Revolution, when the revolutionary government took possession of billions of dollars' worth of American fighter jets and other weapons, an arsenal that Iran has used ever since.[37]

Arms sales and transfers can also harm the regions into which American weapons flow. One danger is dispersion—when weapons sold to a foreign government end up in the hands of criminal groups or adversaries. This risk is highest with sales or transfers to fragile states that are unprepared, unwilling, or too corrupt to protect their stockpiles adequately. For instance, despite America's efforts to train and equip the Iraqi army, Islamic State fighters in 2014 captured three Iraqi army divisions worth of American equipment—including tanks, armored vehicles, and infantry weapons—fueling their campaign.[38] American arms sales can also prolong and intensify interstate conflicts. Although the goal might be to alter the military balance of a conflict to facilitate a speedy end, sending weapons can also encourage the recipients to continue fighting even with no chance of success, leading to more casualties.

Finally, US weapons sales in the name of battling terrorism and insurgency can undermine US national security when they are made to corrupt regimes and to nations with a history of human right violations. American firepower can enhance regime security and enable oppressive governments

to mistreat minority groups and wage inhumane actions against insurgents or terrorist groups. In countries where serious corruption is endemic, American weapons can be diverted from their intended recipients and wind up in the wrong hands. For example, as a result of military and police corruption, the small arms and light weapons that the United States sends to Mexico and to several other Latin American countries in support of the war on drugs often lead to increased gun violence and facilitate the very crimes they were meant to stop.[39]

The Arms Export Control Act and the Leahy Law are attempts to reduce the downstream risks of arms exports. As noted, these laws require the executive branch to assess every sale to ensure that the national security benefits outweigh the risk of sparking, amplifying, or enabling arms races, proliferation, conflict, or human rights violations. Since the Carter administration, every version of the Conventional Arms Transfer Policy has reiterated a list of potential risks to be avoided. Even the Trump administration, widely seen as the most pro-export administration since Nixon's term, explicitly identifies a host of risk-related criteria that will—at least in theory—guide US arms sales decisions in the 2018 arms transfer policy:

- The transfer's consistency with United States interests in regional stability.
- The recipient's ability to prevent the diversion of sensitive technology to unauthorized end users.
- The risk that the transfer will have adverse economic, political, or social effects with the recipient country.
- The risk that the transfer may be used to undermine international peace and security or contribute to abuses of human rights, including acts of gender-based violence and acts of violence against children, violations of humanitarian law, terrorism, mass atrocities, or transnational organized crime.
- Whether the United States has knowledge at the time of authorization that the transferred arms will be used to commit: genocide; crimes against humanity; grave breaches of the Geneva Conventions . . . [;] attacks internationally directed against civilian objects or civilians who are legally protected from attack; or other war crimes.
- The risk that the transfer could undermine the integrity of international nonproliferation agreements.[40]

Though declaratory policy does not always match reality, history does provide some evidence that risk matters for American decision-making under certain circumstances. The Nixon Doctrine and more recently the Obama administration's "light footprint" strategy were both efforts to reduce risk by substituting American weapons for American boots on the

ground. And as noted, the United States typically does not sell its most advanced technology outside the NATO alliance, nor does it allow export of sensitive nuclear weapons–related technology. Moreover, the United States currently bans 19 nations from purchasing American weapons—including not only obvious competitors like Russia, China, and North Korea but also countries like Sudan, South Sudan, Somalia, Syria, and Côte D'Ivoire—which present obvious risks thanks to ongoing civil conflicts and fragile political systems.

On the other hand, the American record since 9/11 casts doubt on the influence of such considerations. Since 2001, the United States has sold major conventional weapons to 169 countries; many of them are autocratic, have long records of human rights violations, or are involved in conflict. In those cases, it appears that strategic and/or economic factors have outweighed whatever risks have been identified. Making it difficult to adjudicate among these influences is that the fact that the State Department does not provide any public summary of the decision-making process or its assessment of the relative weight of strategic, economic, and risk considerations.

Roles of Strategic, Economic, and Risk Factors

To assess the relative importance of power, profit, and prudence, we conducted two sets of analyses on cumulative arms sales from 2002 through 2019 on these measures of strategic, economic, and risk factors, as well as a third analysis of annual sales over the period. We collected data concerning arms purchases, ally status, bilateral trade, military expenditures, risk indicators, and other control variables on 183 countries, 169 of which as noted purchased American weapons during the time period.[41]

Measuring the precise strategic value of any individual weapons deal is difficult. A comprehensive analysis would require not only an assessment of the strategic value of the American partnership with the customer but also the potential strategic benefits from the specific weapons being sold over the life span of the weapons system. Moreover, to be useful in a quantitative analysis, the assessment would have to be conducted in a manner that allowed comparison across cases. How, for example, should one quantify the benefits of selling weapons to Norway versus Taiwan? We are unaware of any granular analyses of this sort in the literature.

We took a more modest approach, beginning with the assumption that weapons sales to allies are more valuable strategically than sales to non-allies, other things being equal. Though clearly not all allies have the same importance, and though at times non-allies are quite important to American

national interests, using allies in our framework allows a commonsense starting point. After all, if allies are not more important to the United States than other nations, other things being equal, then the entire concept of alliances runs aground and any argument about the strategic value of arms sales is likely doomed. Moreover, using allies as a proxy sets a low bar for the argument that strategy matters and thus serves as a bulwark against conclusions that other considerations weigh more heavily than strategy.

Following previous work, we consider a nation an ally if it has signed a formal alliance or defense pact with the United States, if the US has designated the nation as a major non-NATO ally, or if scholars have typically included the country as an informal ally despite the lack of binding legal treaties.[42] The result is a list of 74 American allies. Though simple, this approach provides an explicit measure of strategic value and a straightforward first step in assessing whether strategic logic drives arms sales.

Measuring the impact of economic motivations on a specific arms sales decision is also difficult. Since most sales predicated on a strategic rationale will also have economic benefits, it is hard to know where one motive ends and the other begins. When the United States encourages its allies to buy the new and more expensive F-35 instead of the older and less expensive F-16, for example, it is difficult to know whether strategic or economic logic is at work, or simply both. In cases where there is no obvious strategic rationale, the clear default expectation from the arms sales literature is that the rationale is economic.

We measure economic incentives by assessing the correlation of bilateral trade and state military expenditures with American arms sales. Though heavily regulated, the global arms market remains a market—one with a strong and steady demand for the products the United States is selling. Given this, it is reasonable to imagine that the flow of arms will be higher between the US and countries with which it also conducts a good deal of other business. This might be due to a higher level of business contacts between the two nations, greater similarity of political and economic systems, or more experience dealing with the other country's business and political cultures.[43] We measure imports and exports between a given country and the United States since 2002 with data from the International Monetary Fund.[44]

Similarly, in a global market it makes sense that exporters will sell more products to customers with higher levels of demand. A straightforward proxy for the demand for major conventional weapons is a nation's annual military expenditures. Though some high-spending nations also have their own defense industries, and could thus spend large sums on defense with-

out buying from the United States, in practice the impact of this circumstance is muted by two factors. First, most nations buying weapons do not produce much major conventional weaponry. Second, even rich European nations that export weapons themselves also buy advanced weapons from the United States, particularly expensive aircraft. As a result, we measure potential demand for American weapons by collecting data from the Security Assistance Monitor on each nation's 2017 military expenditures.[45]

To date, few published efforts measure the risks of arms sales quantitatively. To assess risk, we use an updated version of the Arms Sales Risk Index (ASRI).[46] In the absence of a detailed historical record about the outcomes of American arms sales, good or bad, the ASRI is an effort to gauge the risk that weapons sold or transferred to any particular country will lead to the sorts of negative consequences outlined in the Arms Export Control Act and other federal policies. The index assesses the overall "riskiness" of each potential customer for US arms on a 0–100 scale based on the equal weighting of six factors (outlined below) that the literature suggests correlate with the likelihood of negative consequences occurring (fig. 1).[47]

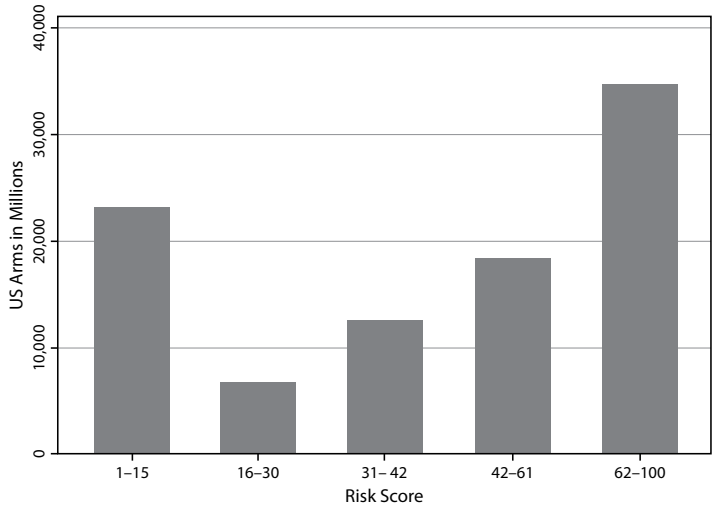


Figure 1. Risk and arms sales, 2002–19

We construct the ASRI by first identifying four underlying risk factors likely to lead to the kinds of negative outcomes noted above. The first is corruption. States with high levels of corruption should pose a much greater risk for diversion, that is, weapons being stolen and then sold to third parties including criminal gangs, insurgents, terrorist groups, or unauthorized local military units. To assess this factor, the index relies on Transparency International's Corruption Perceptions Index ranking 180 countries and territories by their perceived levels of corruption.[48]

The second risk factor we consider is the stability of the recipient nation. Fragile states have weak economies, lack the ability to deliver services effectively, have difficulty managing internal security, and are often beset by internal political divisions. Arms sales to these states pose a greater risk for a wide range of negative outcomes including diversion and the misuse of weapons by government forces, as well as for the amplification of existing conflicts. To measure fragility, we consult the Fragile States Index produced by The Fund for Peace.[49]

The third risk factor is a state's behavior toward its own citizens. States with a poor record of human rights or that regularly use violence against their own citizens pose a greater risk for human rights abuses and generally using American weapons in harmful ways. We include Freedom House's "Freedom in the World" Index to assess a state's commitment to human rights and freedom and the US Department of State's Political Terror Scale to account for state use of violence against civilians.[50]

Finally, conflict is a significant indicator of risk. States engaged in interstate conflicts or facing higher insurgency or terrorist threats likely pose much greater risks for dispersion, blowback, entanglement, arms races, regional instability, and human rights abuses. To measure state engagement in conflict, we rely on the Global Terrorism Index, which ranks countries according to the level and impact of terrorism on the domestic front, and the Uppsala Conflict Data Program/Peace Research Institute Oslo (UCDP/PRIO) Armed Conflict Dataset, which codes each state's participation in external conflicts on a simple scale (high level, low level, and no conflict).[51]

The 2020 ASRI scores range from a low of 2 to a high of 95, with an average of 39 and a standard deviation of 24.2. Though these risk metrics are commonsensical, they should be considered hypotheses in the absence of quantitative validation rather than actual measurements of risk. In the meantime, however, there is good reason to believe that nations scoring higher on this index are indeed riskier customers even though we cannot be certain about the precise weighting of different components. The world's least risky nations are Norway, New Zealand, Switzerland, Luxembourg, and Denmark. The riskiest nations are Syria, South Sudan, Yemen, Afghanistan, and Somalia. Countries scoring at the global average include Senegal, Armenia, South Africa, UAE, and Belarus.

American policy leads us to expect the United States to sell fewer weapons to the nations scoring highest on the risk index, other things being equal. Of course, in foreign policy things are rarely equal, and we do not expect that risk should have the same impact in all cases. As noted, a

rational approach does not require zero risk but simply ensuring that the potential benefits outweigh the potential risks. Sometimes, as in the case of Afghanistan and Iraq, the United States will sell (and give) billions of dollars of weapons to nations scoring very high on the risk index because decision makers believe the strategic case warrants doing so. Given this dynamic, it is impossible to determine a priori where the tipping point between the potential strategic and economic benefits and the potential risks might be, making it difficult to assess how much impact risk has on American sales decisions. Even so, if risk sensitivity is a significant feature of the arms sales decision process, and if selling weapons to allies carries greater strategic value than sales to non-allies, we should expect risk to have a greater impact on sales to non-allies.

Findings, Implications, and Objections

Our analysis reveals three broad findings. First, controlling for other factors, it finds strong support for the importance of strategic considerations in shaping the flow of American arms sales. Figure 2 breaks down arms sales by ally status, revealing that sales to allies clearly outpaced sales to other nations between 2002 and 2019. The 74 American allies purchased $135 billion of weapons compared to $75 billion for the 124 non-allies. Figure 2 also shows that, since 9/11, non-NATO allies have received considerably more weapons than those in the treaty organization. In our regression analysis, ally status correlated positively and significantly with both cumulative arms sales and annual arms sales.[52] According to our analysis, the United States sells over twice as much to allies as to non-allies, holding other variables at their mean.[53]

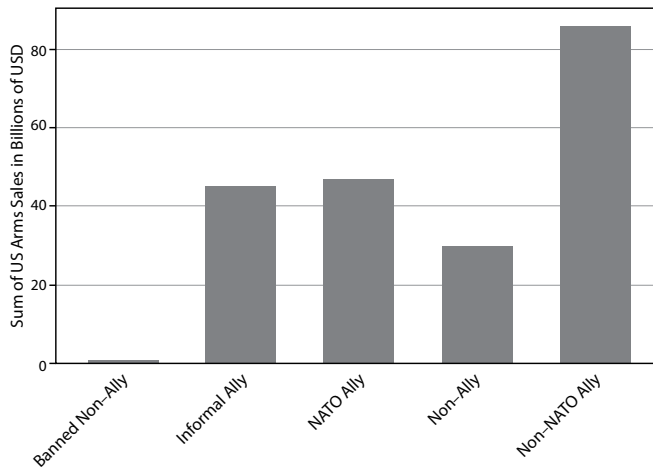


Figure 2. Allies and arms sales

Second, we find clear evidence for the importance of economic considerations. The United States sells more weapons to countries it trades most with and that spend more on their militaries. The trade and military expenditure variables correlated positively with arms sales in the regression analysis, indicating that trade and customer demand influence US arms sales even after considering the effect of alliances. Bilateral trade had a greater impact on cumulative arms sales than on annual variation in arms sales, suggesting that trade relationships may explain long-term arms sales patterns while variations in national military expenditures are more helpful in understanding year-to-year fluctuations. Marginal effects analysis shows that moving from the lowest to the highest value for bilateral trade, while holding other variables at their means, leads to a predicted increase of $491 million in arms sales. Doing the same for military expenditures leads to a predicted increase of $372 million in sales.

Finally, we find little evidence that risk has any important effects on arms sales decisions. The risk variables failed to reach statistical significance in our analysis. Our analysis suggests the proximate reason for this: both the least risky and most risky nations purchase more American weapons than nations scoring in the middle of the risk index. Table 1 shows that the list of leading customers includes both low-risk nations like Australia and Japan and higher-risk nations like Saudi Arabia, Egypt, and Turkey. Moreover, the regression analysis in figures 3 and 4 disavows the idea that the United States weighs risk more more heavily when considering sales to non-allies than to allies.

The most obvious implication of our findings is that the United States significantly privileges strategic and economic considerations over concerns about risk as it assesses potential arms sales. The most powerful explanatory factors for the pattern and volume of sales are whether a country is an ally of the United States, the level of trade conducted between a country and the United States, and how much a country spends on its military each year. Contrary to American statutory requirements—beyond compliance with the United Nations arms sales bans and limits on sales to obvious adversaries like Russia, China, and North Korea—we found no signs that the risk of negative downstream consequences impacts weapons sales even in cases where the strategic benefits appear to be marginal. Since 9/11 the United States has, on average, sold almost as much to the riskiest countries in the world as to the least risky, sold more to countries rated "not free" by Freedom House than to free or "partly free" countries, and sold almost twice as much to countries engaged in a conflict as to those that are not.

Table 1. Top 20 customers of US major conventional weapons, 2002–19

Country	Arms sales ($US millions)	Ally status	2020 risk index score
Saudi Arabia	31,380	Informal ally	71
Egypt	17,640	Non-NATO ally	78
Taiwan	16,010	Informal ally	11
Israel	15,790	Non-NATO ally	52
Australia	11,700	Non-NATO ally	9
Iraq	10,680	Non-ally	85
Japan	10,360	Non-NATO ally	12
South Korea	9,252	Non-NATO ally	8
United Kingdom	6,825	NATO ally	16
United Arab Emirates	6,660	Non-ally	39
Greece	6,381	NATO ally	25
Turkey	6,282	NATO ally	77
Kuwait	5,552	Non-NATO ally	36
Canada	4,222	NATO ally	9
Poland	4,072	NATO ally	19
Pakistan	4,051	Non-NATO ally	78
Singapore	3,671	Non-ally	12
Netherlands	3,525	NATO ally	8
Jordan	3,090	Non-NATO ally	42
Germany	2,935	NATO ally	12

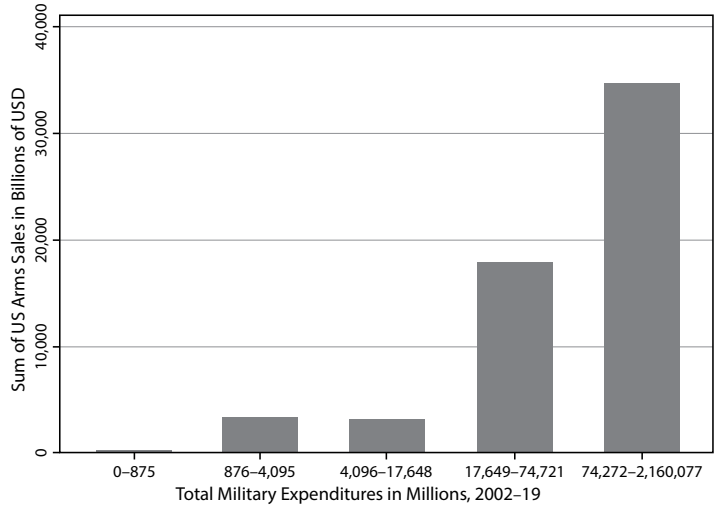


Figure 3. Total military expenditures and US arms sales, 2002–19

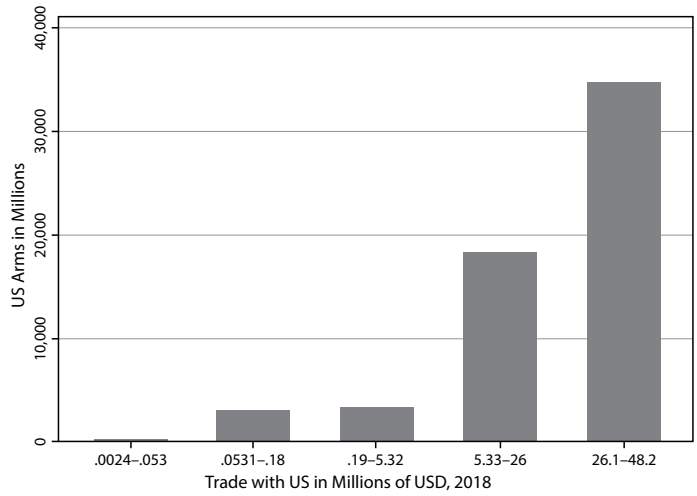


Figure 4. Bilateral trade and US arms sales, 2018

Interpreting our findings, however, requires caution. The apparent lack of risk sensitivity can be explained in at least two different ways. One view, common among critics, is that arms sales decisions privilege short-term strategic and economic benefits while discounting potential downstream negative consequences. This interpretation not only aligns with the analysis but also fits with a good deal of literature on political and governmental decision-making. After all, though the immediate benefits of arms sales are quite obvious—and presidents, policy makers, and CEOs of defense contractors can take credit for actions taken in the present—the future is difficult to predict. Any negative consequences will occur on someone else's watch many years from now.

It would be too hasty to conclude that the US government simply dismisses concerns about risk. However, the presumption that US strategic and economic concerns outweigh risk concerns raises an important challenge to overly simplistic interpretations of our findings. An advocate of American arms sales might point out, for example, that the risks of arms sales are often correlated with the strategic benefits. Several of the riskiest consumers of American weapons since 2002—Saudi Arabia, Iraq, Egypt, and Pakistan, for instance—have also been involved in conflicts and other situations that raised the potential strategic benefits of American engagement (or to put it another way, each situation presented the possibility of serious negative consequences if the United States did not get involved). Since 9/11, the United States has chosen to transfer billions of dollars of weapons to some extremely risky clients largely due to the greater risk of inaction.

Similar care should be exercised when interpreting our findings about strategic considerations. For instance, the data make clear that the United States simply sells much more to allies (an average of $2.3B) than to non-allies (an average of $600M). But ally status is a blunt measure of strategic value that could be problematic in either of two directions. First, it might fail to measure the potential strategic benefits of selling weapons to non-allies. Though on balance it makes sense that sales to allies yield higher benefits, clearly there are cases where selling weapons to other nations will make strategic sense. For example, the United States has sold $120 million of weapons to Nigeria since 9/11, mostly with the aim of enabling the government to combat Boko Haram. Therefore, it is possible that our analysis underestimates the strategic nature of American arms sales to some degree.

Assigning strategic value to an arms sale just because the customer is an ally is also problematic. Some critics of American foreign policy have argued that the United States has too many allies—there are 74 in our data set—and that many are simply free riders seeking protection rather than true allies furthering American national security.[54] By one accounting, the United States is responsible for coming to the aid of 25 percent of the world's population should their homelands come under attack.[55] And yet, thanks to its favorable geography, size, wealth, and military capability, none of these allies adds much to the actual security of the United States.[56] Moreover, many analysts believe that some nations we coded as allies—Saudi Arabia, Pakistan, and Turkey, for example—should not be considered allies at all. And though the United States may care about other strategic interests like regional stability, the protection of friendly regimes, and terrorism, a single alliance variable cannot capture the range of strategic benefits involved across those 74 nations. Thus, our analysis possibly overestimates the impact of strategic considerations by assuming that all sales to allies have strategic value.[57]

Finally, our analysis clearly indicates that economic incentives matter. Bilateral trade and customer demand for American weapons influence sales to allies and non-allies alike and to risky and nonrisky nations. On one hand, it should surprise no one to discover that economic concerns help drive arms exports. On the other hand, when considered in light of the discussion about risk, our findings do little to quell the concerns of critics who believe that the United States and other arms exporting nations too often ignore downstream consequences to make money in the short run. Since 9/11, the United States has sold to many countries scoring

high on the risk index but for which the strategic benefits are dubious or the track record of negative outcomes is already long.

Balancing Costs and Benefits: The Arms Sales Risk-Reward Matrix

If, as our analysis suggests, the United States is doing too little to incorporate risk assessments into arms sales decisions, how might the process be improved? Though the law requires the government to weigh risks and benefits, it does not define how to make the calculations. The problem for policy makers is that several challenges make assessing the expected costs and benefits of arms sales difficult. First, policy makers often clash over defining benefits and costs. Strengthening a NATO ally, for example, might carry very different weight for Donald Trump than for previous presidents. Similarly, whether an arms sale affects the rate of gun violence in a client nation might matter a great deal to some but very little to others. Second, the government lacks the necessary historical data to identify and measure the potential benefits and costs in a way that encourages comparison and reasoned trade-offs. Third, forecasting is difficult under the best of circumstances, and forecasting arms sales risks is even more challenging given the complex interdependence of international affairs. US actions often cause unexpected reactions from others. Finally, the balance between risk and reward is a moving target. Just as the potential benefits vary widely from case to case, so does the amount of acceptable risk. There is no simple heuristic and no specific amount of benefit or risk that one can use to determine the point at which the United States should or should not sell weapons.

Despite these challenges, it is possible to think more rigorously about balancing the strategic benefits and potential risks of arms sales. Below we outline a simple tool that we believe offers a useful first step for policy makers trying to balance arm sales' risks and rewards. Given the considerable uncertainty on both sides of the equation, a useful decision-making tool, we believe, will encourage policy makers to take a more conservative "do no harm" approach that avoids overstating the easy-to-see benefits and underselling the hard-to-see risks of arms sales. By design such an approach would forgo maximizing the upside potential of arms sales in return for a reduction in the most common, predictable negative consequences.

We also argue that economic considerations should take a back seat in the calculus. Though the economic benefits from selling weapons are certainly positive, we believe that they pale in comparison to the potential strategic benefits on the one hand and to the potential negative outcomes

on the other. From an economic security perspective, the American defense industrial base is already so much larger and more robust than that of any other nation that the notion of arms exports as integral to its health rings hollow.[58] From an economic growth perspective—compared to the strategic value of strengthening NATO, for example—shaving a few dollars off the Pentagon's F-35 per unit acquisition costs, adding a few thousand jobs, or making a few billion dollars in sales for American companies is a rounding error. And when American forces in Iraq took fire from Islamic State fighters using stolen American weapons, any suggestion that economic benefits might have justified the harm is inappropriate.

With these assumptions in hand, we then used our data to construct a simple tool we call the Arms Sales Risk-Reward Matrix. To create the matrix, we began by using a nation's ally status to determine at a very basic level whether the potential strategic benefits are likely to be significant. As noted, according to the Department of Defense, the United States currently counts 74 nations as allies. One can—and should—argue about the value of specific allies or the benefits of the specific weapons being sold. However, in the interest of keeping things simple, as a first step most would agree that selling weapons to allies versus non-allies is more likely to bring strategic benefits.

We next looked at each nation's individual Arms Sales Risk Index components to identify red flags. We define a *red flag* as any instance where a nation scored in the riskiest category for a particular indicator. To earn a red flag, a nation had to do one of the following: score as "not free" in the Freedom House index (47 nations), fall in the "alert" category in the Failed States Index (31 nations), engage in any kind of military conflict (57 nations), experience "political violence everywhere" (nine nations) as rated by the US State Department, or suffer a "high impact" from terrorism as scored by the Global Terrorism Index (18 nations). Since many nations earned more than one red flag, this process identified a total of 76 red-flag nations.

Using these two measures, we classified nations into one of four categories, summarized in table 2. These categories, we believe, offer a useful starting point for discussion about the wisdom of exporting weapons. Low-risk allies are those for which one might most easily embrace a presumption of approval for arms sales. The strategic value of helping these nations maintain capable militaries is clear in many cases, none are engaged in active military conflicts, and all enjoy political systems stable and competent enough to manage and use their arms responsibly. The list of

these 60 countries includes Japan, South Korea, Mexico, all NATO members except Turkey, and several Caribbean nations.

Table 2. Arms sales risk-reward matrix

	Not an ally	Ally
No red flags	n = 47 2002–19 sales: $7.4B *Top customers* Singapore ($3.6B) Finland ($1.2B) Switzerland ($976M) Malaysia ($554M) Sweden ($424M)	n = 60 2002–19 sales: $101.8B *Top customers* Taiwan ($16B) Australia ($11.7B) Japan ($10.4B) South Korea ($9.3B) United Kingdom ($6.8B)
One or more red flags	n = 62 2002–19 sales: $24.5B *Top customers* Iraq ($10.7B) UAE ($6.7B) India ($2.2B) Oman ($1.8B) Indonesia ($703M)	n = 14 2002–19 sales: $81.7B *Top customers* Saudi Arabia ($31.4B) Egypt ($17.6B) Israel ($15.8B) Turkey ($6.3B) Pakistan ($4.1B)

In the lower left-hand quadrant, on the other hand, are the 62 nations to which it might make sense to stop selling weapons entirely. This group includes the 17 nations already banned from buying American weapons, as well as others suffering from a host of problems ranging from civil conflict and widespread terrorism to unstable governments and disastrous human rights records. The chances for negative downstream consequences in these cases are very high, while the strategic benefits are debatable.

The remaining two categories present somewhat less clear guidance and require more input from decision makers. In the upper left-hand quadrant are low-risk nations that are not allies. Advocates of arms sales might argue that this is precisely the group of nations where economic benefits could be the deciding factor. Since the risks are low, why not allow arms sales to proceed? From a risk-reduction standpoint, however, one might point out that there is no such thing as zero risk. Not only is there is a robust international black market for American weapons, but thanks to the lifespan of most weapons systems, the horizon for calculating risk is quite long. Nations that seem stable today might not seem so two decades from now. Adjudicating this tension will require policy makers to decide whether the modest economic gains from arms sales to this category of nations are worth the potential risk.

The final category, in the lower right-hand quadrant, is high-risk allies. Though the smallest category with just 14 countries, it holds the potential to generate some of the most heated debate over arms sales. This group

includes Saudi Arabia, Israel, Afghanistan, Turkey, Pakistan, Thailand, and Colombia. This category is difficult to assess because the cases come with both compelling strategic interests and potentially large downside consequences. Policy disagreements over these cases will hinge in part on views about American grand strategy and in part over retrospective assessments of policy success and failure. For those who believe that the United States must lean forward and take an active role in managing regional balances of power, these risks are likely to look more palatable. For those who advocate a more restrained grand strategy or who believe the war on terror has been a costly failure, for example, arms sales may be less risky but still not worth the gains. And in either case, the red flags suggest that the United States should look for approaches to achieving strategic goals involving those allies that would not involve the same level of risk.

This risk-reward matrix, we believe, can be useful regardless of one's own specific assumptions about the costs and benefits of arms sales. Starting with the initial classification, decision makers can then use their own criteria to reclassify nations with respect to risks and benefits. To assess specific arms deals, officials can create more detailed assessments of benefits and risks by accounting for the weapons in question and by incorporating current conditions and intelligence forecasts. The risk-reward matrix's utility lies primarily in encouraging decisions makers to identify risks and benefits more explicitly to weigh them against each other more effectively. Those calculations, in turn, can be used as a baseline for assessing the downstream outcomes with client states, both positive and negative, as data for improving future decisions about arms sales.

Much more research needs to be done to inform future arms sales decisions. Though the risk index and the risk-reward matrix are useful tools for thinking about why and where negative outcomes might occur, they represent a set of testable hypotheses for research aimed at establishing a more rigorous basis for future decision-making. The current debate relies more on assumptions than on evidence about the impacts of arms sales. Though there have been several efforts to assess the impacts of arms sales, there is little broad agreement about the conditions under which arms sales lead to either positive or negative outcomes.[59] Improving our understanding of these dynamics would be a major contribution to the practice of American foreign policy. SSQ

A. Trevor Thrall

A. Trevor Thrall is an associate professor in the Schar School of Policy and Government at George Mason University and a senior fellow at the Cato Institute. His publications include *US Grand Strategy in the 21st Century: The Case for Restraint* (London: Routledge, 2018).

Jordan Cohen

Jordan Cohen is a PhD student in the Schar School of Policy and Government at George Mason University.

Caroline Dorminey

Caroline Dorminey is a policy director at Women's Action for New Directions where she handles a portfolio of nuclear policy, arms control, and defense politics issues. She is an expert with the Forum on the Arms Trade and a contributor to the Pentagon Budget Campaign. She holds a master's degree in international relations from the University of Chicago.

Notes

1. This figure does not include sales of small arms and light weapons not requiring government approval for sale. Arms sales data is from Security Assistance Monitor, "Arms Sales Dashboard," accessed 1 April 2020, https://securityassistance.org/.

2. Peter D. Wezeman et al., *Trends in International Arms Transfers, 2018* (Stockholm: Stockholm International Peace Research Institute, March 2019), https://www.sipri.org/.

3. Exceptions include Shannon Lindsey Blanton, "Promoting Human Rights and Democracy in the Developing World: U.S. Rhetoric versus U.S. Arms Exports," *American Journal of Political Science* 44, no. 1 (2000): 123–31, DOI:10.2307/2669298; Shannon Lindsey Blanton, "Foreign Policy in Transition? Human Rights, Democracy, and U.S. Arms Exports," *International Studies Quarterly* 49, no. 4 (2005): 647–67; and Jennifer Erickson, *Dangerous Trade: Arms Exports, Human Rights, and International Reputation* (New York: Columbia University Press, 2015).

4. Keren Yarhi-Milo, Alexander Lanoszka, and Zack Cooper, "To Arm or to Ally? The Patron's Dilemma and the Strategic Logic of Arms Transfers and Alliances," *International Security* 41, no. 2 (Fall 2016): 93, https://doi.org/10.1162/ISEC_a_00250.

5. Robert M. Gates, "Helping Others Defend Themselves: The Future of U.S. Security Assistance," *Foreign Affairs* 89, no. 3 (May/June 2010): 2–6, https://www.jstor.org/; Stephanie G. Neuman, "Power, Influence, and Hierarchy: Defense Industries in a Unipolar World," *Defence and Peace Economics* 21, no. 1 (2010): 105–34, DOI: 10.1080/10242690903105398; and Andrew J. Shapiro, "A New Era for U.S. Security Assistance," *Washington Quarterly* 35, no. 4 (Fall 2012): 23–35, https://csis-prod.s3.amazonaws.com/.

6. Volker Krause, "Hazardous Weapons? Effects of Arms Transfers and Defense Pacts on Militarized Disputes, 1950–1995," *International Interactions* 30, no. 4 (2004): 349–71; William D. Hartung, *Prophets of War: Lockheed Martin and the Making of the Military Industrial Complex* (New York: Nation Books, 2012); Andrew J. Pierre, *The Global Politics of Arms Sales* (Princeton: Princeton University Press, 2014); and Yarhi-Milo et al., "To Arm or to Ally?"

7. White House, "Statement from President Donald J. Trump on Standing with Saudi Arabia," press release, 20 November 2018, https://www.whitehouse.gov/.

8. Ann Mezzell, "Deterring Terrorists Abroad: The Implausibility of Indirect Deterrence," *Strategic Studies Quarterly* 13, no. 1 (Spring 2019): 134–52, https://www.airuniversity.af.edu/; William D. Hartung, "Breaking the Arms-Sales Addiction: New Directions for U.S. Policy," *World Policy Journal* 8, no. 1 (Winter 1990): 1–26, https://www.jstor.org/; William D. Hartung, "U.S. Conventional Arms Transfers: Promoting Stability or Fueling Conflict?," *Arms Control Today* 25, no. 9 (November 1995): 9–13, www

.jstor.org/; James M. Scott, *Deciding to Intervene: The Reagan Doctrine and American Foreign Policy* (Durham, NC: Duke University Press, 1996); Cassady Craft, *Weapons for Peace, Weapons for War: The Effects of Arms Transfers on War Outbreak, Involvement, and Outcomes* (New York: Routledge, 1999); Cassady Craft and Joseph Smaldone, "Arms Imports in Sub-Saharan Africa: Predicting Conflict Involvement," *Defence and Peace Economics* 14, no. 1 (2003): 37–49, DOI: 10.1080/10242690302933; Rachel Stohl and Suzette Grillot, *The International Arms Trade* (Cambridge: Polity Press, 2009); Jennifer L. Erickson, "Stopping the Illegal Flow of Weapons: Compliance with Arms Embargoes, 1981–2014," *Journal of Peace Research* 50, no. 2 (2013): 159–74; Jesse Dillon Savage and Jonathan D. Caverley, "When Human Capital Threatens the Capitol: Foreign Aid in the Form of Military Training and Coups," *Journal of Peace Research* 54, no. 4 (July 2017): 542–57, https://doi.org/10.1177/0022343317713557; David Edelstein and Joshua Itzkowitz Shifrinson, "It's a Trap! Security Commitments and the Risks of Entrapment," in *US Grand Strategy in the 21st Century: The Case for Restraint*, eds. A. Trevor Thrall and Benjamin H. Friedman (New York: Routledge, 2018) 19–41; and A. Trevor Thrall and Caroline Dorminey, "Risky Business: The Role of Arms Sales in U.S. Foreign Policy," Cato Policy Analysis no. 836, 13 March 2018, https://www.cato.org/.

9. "Foreign Relations and Intercourse," *US Code*, Title 22, §2791(a), https://www.law.cornell.edu/.

10. US Department of State, Bureau of Democracy, Human Rights, and Labor, "Leahy Law Fact Sheet," 22 January 2019, https://www.state.gov/.

11. Donald J. Trump, "National Security Presidential Memorandum Regarding U.S. Conventional Arms Transfer Policy," 19 April 2018, https://www.whitehouse.gov/; and Eugene Gholz, "Conventional Arms Transfers and US Economic Security," *Strategic Studies Quarterly* 13, no. 1 (Spring 2019): 42–65, https://www.airuniversity.af.edu/.

12. Karoun Demirjian and Colby Itkowitz, "Trump Vetoes Congress's Attempt to Block Arms Sales to Saudi Arabia," *Washington Post*, 24 July 2019, https://www.washingtonpost.com/.

13. Gholz, "Conventional Arms Transfers," 42–65; and Jonathan Caverley, "America's Arms Sales Policy: Security Abroad, Not Jobs at Home," *War on the Rocks*, 6 April 2018, https://warontherocks.com/.

14. Rachel Stohl, "With Eye on Exports Boost, Trump Admin Revises Arms Transfers Policies," *Axios*, 27 April 2018, https://www.axios.com/.

15. "U.S. Munitions List," 58 Fed. Reg. 3928, 22 July 1993, *reprinted in* 22 U.S.C. sec. 121.1, https://www.ecfr.gov/.

16. Security Assistance Monitor, "Arms Sales Dashboard," 2019, https://securityassistance.org/.

17. Stockholm International Peace Research Institute (SIPRI), "Global Arms Trade: USA Increases Dominance; Arms Flows to the Middle East Surge, Says SIPRI," press release, 11 March 2019, https://www.sipri.org/media/.

18. Arms Export Control Act, 22 *US Code* 2778, §38, https://www.pmddtc.state.gov/.

19. Paul K. Kerr, *Arms Sales: Congressional Review Process*, CRS Report RL31675 (Washington, DC: Congressional Research Service, 1 April 2019), https://fas.org/; and Merrit Kennedy, "Trump Vetoes Bills Intended to Block Arms Sales to Saudi Arabia," 25 July 2019, https://www.npr.org/. Notifications data comes from the Security Assistance Monitor database, accessed 1 April 2020, http://securityassistance.org.

20. Krause, "Hazardous Weapons"; Thomas W. Zarzcki, *Arms Diffiusion: The Spread of Military Innovations in the International System* (London: Routledge, 2002); David Kinsella, "The Arms Trade," in *The Handbook on the Political Economy of War*, eds. Christopher Coyne and Rachel Matthers (Cheltenham, UK: Edward Elgar, 2011); Hugo Meijer et al., "Arming China: Major Powers' Arms Transfers to the People's Republic of China," *Journal of Strategic Studies* 41, no. 6 (2018): 850–86, https://doi.org/10.1080/01402390.2017.1288110; and Jeremy Pressman, *Warring Friends: Alliance Restraint in International Politics* (Ithaca, NY: Cornell University Press 2011).

21. Mezzell, "Deterring Terrorists Abroad"; Gates, "Helping Others Defend Themselves"; Shapiro, "New Era"; and Yarhi-Milo et al., "To Arm or to Ally?"

22. Hartung, "Breaking the Arms-Sales Addiction"; Stohl and Grillot, *International Arms Trade*; and Hugo Meijer, "Balancing Conflicting Security Interests: U.S. Defense Exports to China in the Last Decade of the Cold War," *Journal of Cold War Studies* 17, no. 1 (Winter 2015): 4–40, https://doi.org/10.1162/JCWS_a_00529.

23. Shapiro, "New Era."

24. Gregory S. Sanjian, "Promoting Stability or Instability? Arms Transfers and Regional Rivalries, 1950–1991," *International Studies Quarterly* 43, no. 4 (December 1999): 641–70, https://doi.org/10.1111/0020-8833.00140; and Gregory S. Sanjian, "Arms Transfers, Military Balances, and Interstate Relations: Modeling Power Balance versus Power Transition Linkages," *Journal of Conflict Resolution* 47, no. 6 (2003): 711–27, https://doi.org/10.1177/0022002703258801; Blanton, "Foreign Policy in Transition"; and Kinsella, "Arms Trade."

25. Anders Akerman and Anna Larsson Seim, "The Global Arms Trade Network 1950–2007, *Journal of Comparative Economics* 42, no. 3 (2014): 535–51, https://doi.org/10.1016/j.jce.2014.03.001.

26. Krause, "Hazardous Weapons"; and Gholz, "Conventional Arms Transfers."

27. Hartung, "Breaking the Arms-Sales Addiction."

28. Pierre, *Global Politics of Arms Sales*; Hartung, "Breaking the Arms-Sales Addiction"; and Jonathan Caverley and Ethan B. Kapstein, "Arms Away: How Washington Squandered Its Monopoly on Weapons Sales," *Foreign Affairs* 91, no. 5 (September/October 2012): 125–32, www.jstor.org/.

29. William D. Hartung, "Nixon's Children: Bill Clinton and the Permanent Arms Bazaar," *World Policy Journal* 12, no. 2 (Summer 1995): 25–35, https://www.jstor.org/.

30. David Gold, "The Changing Economics of the Arms Trade," in eds. Ann R. Markusen and Sean S. Costigan, *Arming the Future: A Defense Industry for the 21st Century* (New York: Council on Foreign Relations Press, 1999), 249–68; and Michael Brzoska, "The Economics of Arms Imports after the End of the Cold War," *Defence and Peace Economics* 15, no. 2 (2004): 111–23.

31. Paul Holden, ed., *Indefensible: Seven Myths That Sustain the Global Arms Trade* (London: Zed Books, 2017).

32. Trump, "Conventional Arms Transfer Policy"; and Gholz, "Conventional Arms Transfers."

33. William D. Hartung, *Peddling Arms, Peddling Influence: Exposing the Arms Export Lobby*, World Policy Institute Report (New York: World Policy Institute, October 1996), http://worldpolicy.org/; and Robert Levinson, Sopen Shah, and Paige K. Conner, *Impact of Defense Spending: A State-by-State Analysis* (New York: Bloomberg, 2011), http://www.ct.gov/.

34. For a more complete discussion of risks see Thrall and Dorminey, "Risky Business."

35. Robert M. Press, "Somali Civil War Is Fueled by Huge Stockpiles of Weapons," *Christian Science Monitor*, 14 October 1992, https://www.csmonitor.com/.

36. Anna Stavrianakis, "Playing with Words while Yemen Burns: Managing Criticism of UK Arms Sales to Saudi Arabia," *Global Policy* 8, no. 4 (2017): 563–68, https://doi.org/10.1111/1758-5899.12484; and Nima Elbagir et al., "Bomb That Killed 40 Childen in Yemen Was Supplied by the US," CNN, 17 August 2018, https://www.cnn.com/.

37. David Axe, "Iran's Air Force Flies American-Made F-14 Tomcats," *The National Interest*, 5 March 2018, https://nationalinterest.org/.

38. Amnesty International, *Taking Stock: The Arming of Islamic State* (London: Amnesty International, December 2015), https://www.amnesty.org/.

39. Carla Martinez Machain and Jeffrey Pickering, "The Human Cost of the Weapons Trade: Small Arms Transfers and Recipient State Homicide," *Journal of Global Security*, 2019, 1–20, https://doi.org/10.1093/jogss/ogz041; and Rafael Carranza, "U.S. Guns Used in Most Crimes in Mexico, Center for American Progress Says," *The Republic*, AZcentral.com, 6 February 2018, https://www.azcentral.com/

40. Trump, "National Security Presidential Memorandum."

41. We ran our first two analyses on cumulative sales rather than annual sales because most of the risk indicators are not available before 2015. We did, however, conduct a time series analysis from 2006 onward using the Fragile States Index on one of the risk index inputs that correlates fairly closely with the overall Arms Sales Risk Index. For space considerations we do not report the full regression analysis here. The data and analysis are available from the authors in a methodological appendix online at A. Trevor Thrall, Caroline Dorminey, and Jordan Cohen, "The 2019 Arms Sales Risk Index," Cato Institute Policy Analysis, 10 September 2019, http://cato.org/.

42. We also use Michael Beckley's method of identifying informal alliances that function as formal ones. For reference, see Michael Beckley, "The Myth of Entangling Alliances: Reassessing the Security Risks of U.S. Defense Pacts," *International Security* 39, no 4 (Spring 2015): 7–48, https://www.belfercenter.org/; and Brett Ashley Leeds et al., "Alliance Treaty Obligations and Provisions, 1815–1944," *International Interactions* 28, no 3 (2002): 237–60, DOI: 10.1080/03050620213653.

43. For example, see Joanne Gowa and Edward D. Mansfield, "Power Politics and International Trade," *American Political Science Review* 87, no. 2 (June 1993): 408–20, https://doi.org/10.2307/2939050.

44. International Monetary Fund, Direction of Trade Statistics, accessed 1 April 2020, http://data.imf.org/DOT.

45. We acknowledge the potential problem of regressing "x" on "x+e." Nonetheless, in our data set, US arms purchases are on average only 2.17 percent of a country's military expenditures. There are four outliers: Uzbekistan, Afghanistan, Antigua and Barbuda, and Egypt. Every other country falls within two standard deviations of the mean. Thus, US arms sales comprise a small enough proportion of other countries' military expenditures that we feel comfortable with this explanatory variable. Nonetheless, we also ran a time-series analysis with a temporal lag to avoid potential issues with this approach. For this research, see Thrall, Dorminey, and Cohen, "The 2019 Arms Sales Risk Index."

46. Thrall, Dorminey, and Cohen; and Thrall and Dorminey, "Risky Business."

47. To combine each of the risk factors—themselves constructed using a variety of different scales—into a single index, we began by normalizing the scores using a Min-

Max method. After normalization, every country has a score of between 1 (the lowest risk) to 100 (the highest risk) for each measure. We then weighted each of the six components equally, averaging them to find the composite risk score for each country (where a country was missing a variable, we averaged the remaining factors). We weigh the components equally because we do not yet have any empirical reason to consider different weighting systems.

48. Transparency International, "Corruption Perceptions Index," 2018, https://www.transparency.org/.

49. The Fund for Peace, "Fragile State Index," 2018, http://fundforpeace.org/fsi/.

50. Freedom House, "Freedom in the World Index," 2018, https://freedomhouse.org/; and US Department of State, "Political Terror Scale," 2017, http://www.politicalterrorscale.org/Data/.

51. Institute for Economics and Peace, *Global Terrorism Index 2019: Measuring the Impact of Terrorism* (Sydney: Institute for Economics and Peace, November 2019), http://visionofhumanity.org/reports/; and Peace Research Institute Oslo, "UCDP/PRIO Armed Conflict Dataset, version 19.1," accessed 1 April 2020, http://ucdp.uu.se/downloads.

52. Following Beckley, "Myth of Entangling Alliances," we coded Saudi Arabia and Taiwan as allies, though they are technically not formal allies. The regression analysis changed very little, however, when we ran the analysis coding Saudi Arabia and Taiwan as non-allies.

53. We use the Stata program "marginscontplot" to examine marginal effects of our log-transformed covariates.

54. Christopher A. Preble, *The Power Problem: How American Military Dominance Makes Us Less Safe, Less Prosperous, and Less Free* (Ithaca, NY: Cornell University Press, 2009); and Barry R. Posen, *Restraint: A New Foundation for U.S. Grand Strategy* (Ithaca, NY: Cornell University Press, 2014).

55. Beckley, "Myth of Entangling Alliances."

56. A. Trevor Thrall and Benjamin Friedman, *U.S. Grand Strategy in the 21st Century: The Case for Restraint* (London: Routledge, 2018).

57. A powerful argument that security assistance tends not to be very effective is Mezzell, "Deterring Terrorists Abroad."

58. For a discussion, see Gholz, "Conventional Arms Transfers."

59. Mezzell, "Deterring Terrorists Abroad"; Sanjian, "Promoting Stability or Instability?"; Sanjian, "Arms Transfers"; John Sislin, "Arms as Influence: The Determinants of Successful Influence," *Journal of Conflict Resolution* 38, no. 4 (December 1994): 665–89, https://www.jstor.org/; and Charles W. Kegley Jr. and Steven W. Hook, "U.S. Foreign Aid and U.N. Voting: Did Reagan's Linkage Strategy Buy Deference or Defiance?," *International Studies Quarterly* 35, no. 3 (1991): 295–312, https://doi.org/10.2307/2600701.

PERSPECTIVE

American Grand Strategy for an Emerging World Order

Scott Lawless

Abstract

Since the end of the Second World War, the United States has secured its core national interests primarily through the creation and maintenance of the liberal international order. Today, this order is being challenged in ways that will define the twenty-first century context. America's most pressing foreign policy challenge is finding strategies to counter a potentially illiberal global order. Neo-authoritarian states are seeking to establish spheres of influence by violating territorial norms, undermining the liberal order via coercive economic measures, and weakening democratic regimes through unconventional political warfare. The current liberal order is ill-equipped to face these challenges because of two global trends: the erosion of its legitimacy and the shifting global balance of power. In a changing environment such as this, where the ends of American grand strategy remain fixed while its relative means are eroding, the US must revise the ways in which it seeks to achieve its strategic objectives. The shifts in geopolitics today necessitate a revitalization of American grand strategy and the establishment of a new security order—namely, a Concert of Democracies—to secure American interests, reestablish liberal legitimacy, and shape the emerging international order toward a stable future.

*****

The liberal international order that emerged triumphant over fascism and communism during the twentieth century is a testament to the institutions, alliances, and norms US statesmen established to avoid the revival of great power conflict. Though these structures have granted the United States and its allies several decades of unparalleled security and prosperity, it is unclear as to what is invoked by the term *liberal international order*. The modern world is characterized by what is referred to as the *international system* or the global assemblage of sovereign nation-states that is the primary structuring mechanism for interstate relations. *Order*, however, requires that this international system operates within two basic conditions: "a set of commonly accepted rules that define

the limits of permissible action and a balance of power that enforces restraint where rules break down."[1] Rather than being a monolithic structure with which the United States has promoted stability, economic opportunity, and freedom around the world, the liberal order exists as a collection of many *suborders* depending on the type of interactions taking place. The liberal order includes three suborders—the *security order*, the *economic order,* and the *political order*—that have come to define the nature of international relations. Yet today, each of these areas is being challenged in ways that will define the emerging context, and countering strategies to subvert the liberal order is our most pressing foreign policy dilemma.

The liberal security order implies that the international system is inherently rules based and not simply determined by power relations. Rather, international laws and norms restrain the action of states to bring an end to global disorder that has too often been "organized into rival blocs or exclusive regional spheres."[2] The economic order builds upon this notion of rules-based interaction and embraces international markets defined by openness, manifesting "when states trade and exchange on the basis of mutual gain."[3] Moreover, the economic order is increasingly becoming difficult to distinguish from globalization or the "breaking down of artificial barriers to the flow of goods, services, capital, knowledge, and people across borders."[4] Lastly, the political order is "a kind of fusion of two distinct order-building projects."[5] The first dates back to the creation of the modern state system. Encapsulated within the treaties known as the Peace of Westphalia were the concepts of state sovereignty, the inviolability of national borders, and noninterference in another state's domestic affairs. The second element is built upon the ascension of liberal values, such as political, civil, and universal human rights as a collective standard, characterized by the rise of liberal democracies across the world.[6] Thus, the liberal order we conceive of today was defined by the formation of the Westphalian system, "on top of which various forms of order have developed that have become gradually more liberal over time."[7]

Challenges to the Evolving Liberal Order

After the Cold War, many American strategists envisioned a world in which former communist states, devoid of their ideological foundation, would converge with the West, facilitating the "end of history" and the emergence of a "new world order."[8] In such a world, geopolitical rivalries would dissolve as states would converge around universal values and a united conceptualization of global order. Though much of the world continues to benefit from the liberal order the United States champions, the

global order is evolving and not necessarily in the predicted or desired direction. Countries espousing neo-authoritarianism, the belief that societies are best served by stability rather than political and economic liberalization, are actively subverting the current liberal order. According to Will Marshall, president of the Progressive Policy Institute, "for much of the 20th century, the main threat to liberal and democratic societies came from militant and totalizing ideologies: fascism and communism, or revolutionary socialism."[9] However, the current context is characterized by the active targeting of liberal societies "to undermine [them] from within and overwhelm [them] from without," thereby stripping "liberal democracy of its moral allure" and elevating authoritarianism as "a plausible, alternative path to national development and prosperity."[10] While authoritarianism is nothing new, the objective and tactics wielded by neo-authoritarian regimes today define the twenty-first century context and pose new strategic quandaries. Altogether, what authoritarians seek is the creation of an illiberal and multipolar global order—thereby making the world safe for neo-authoritarian regimes—by establishing spheres of influence, undermining the liberal order, and weakening democratic regimes. It is essential, then, to identify the specific ways in which the liberal order is being challenged to devise a sound counterstrategy.

Security Challenges

The liberal order faces a number of security challenges from both revisionist and revanchist neo-authoritarian states. Most notably, the current context is "characterized by [a] decline in the long-standing rules-based international order—creating a security environment more complex and volatile than any we have experienced in recent memory."[11] Russia has violated the sovereignty of Georgia and annexed Crimea, all the while breaching arms control treaties and modernizing its nuclear arsenal. Putin has thrust himself into conflicts in the Middle East and South America to support dictators and to revive Russia's standing on the international stage. Meanwhile, China continues to militarize the South China Sea so as to facilitate implementing an antiaccess/area denial (A2/AD) strategy. Both Russia and China are attempting to carve out spheres of influence through the "acquisition and consolidation of territory using force and in violation of international law."[12] Iran and Saudi Arabia continue to support proxy wars in the Middle East in their bout for regional hegemony while North Korea has developed nuclear and ballistic missile capabilities able to threaten two of America's strongest allies. Each of these states is undermining and seeking to alter the security order to "shape a world consistent

with its authoritarian model—gaining veto authority over other nations' economic, diplomatic, and security decisions."[13]

Economic Challenges

Economic statecraft has always been a tool to influence foreign policy; however, the methods being used today undermine the economic order by overtly coercive measures. Countries like China have experienced successes in this regard due to the relative size and importance of the Chinese market, and "it does so to bolster its territorial claims and national sovereignty or to advance other core interests."[14] States are employing "hybrid economic measures" including "politically conditioned loans and business deals" and coercive business, trade, and investment restrictions to target those competitors espousing critical political perspectives. China's Belt and Road Initiative (BRI) is using conditional loans, deals in the form of development projects, and other economic measures as its principal foreign policy tool in Eurasia. Moreover, China often utilizes predatory trading practices such as import and export restrictions, popular boycotts, and tariff and nontariff trade barriers directed toward democratic states to "target politically influential constituencies."[15] The effectiveness and scale of these tools make it likely that neo-authoritarian states will continue to supplement their security strategies with coercive economic measures to undermine the liberal order.

Political Challenges

The liberal order faces political challenges that involve unconventional political warfare to weaken democratic regimes as well as a "clash of social models" between liberal and neo-authoritarian states.[16] Political warfare "refers to the employment of military, intelligence, diplomatic, financial, and other means short of conventional war to achieve national objectives."[17] Its current "hybrid" adaptation is unconventional in that it involves tactics such as the weaponization of traditional and social media, sophisticated propaganda, and the widespread use of disinformation campaigns to sway public opinion, discredit liberal politicians, and sow distrust for democratic institutions. Russia, for instance, is actively "exploit[ing] European and transatlantic fissures and support[ing] populist movements to undermine European Union and NATO cohesion."[18] Meanwhile, China is targeting the United States' companies, government, and allies as part of its ongoing cyber-espionage campaign to steal trade secrets, intellectual property, and advanced technology. What's more,

the political order is experiencing a clash of social models in which states such as Russia and China "believe in the virtues of a strong central government and disdain the weaknesses of the democratic system."[19] Thus, collective political convergence has never been realized because autocratic leaders "concluded that if the liberal order succeeded globally, it would pose an existential threat to their regimes."[20]

Legitimacy and Power

However, neo-authoritarian states are not solely responsible for the evolving context in which we find ourselves. The current liberal order is experiencing a lack of cohesion from a combination of two global trends: the erosion of its legitimacy and a shift in the perceived balance of power. The first occurs "when the values underlying international arrangements are fundamentally altered, abandoned by those charged with maintaining them."[21] As the leader of the liberal order, the United States is experiencing a crisis of legitimacy on the world stage. The overt hostility with which the Trump administration regards the liberal order and its utility is illustrated by the administration's withdrawal from numerous international agreements, opposition to multilateralism more broadly, and its "conditional approach to once inviolable US alliance commitments in Europe and Asia."[22] However, this sentiment—defined by disdain for globalism and president Trump's rise as the paragon of American nationalism—is a reflection of a rationale that has been fraying since the end of the Cold War. President Trump is not an aberration in American foreign policy; rather, he is the culmination of a 30-year trend toward American disengagement in global affairs.

Orders are created by powerful states to suit their interests, and the same is true for the United States in its creation of the liberal order. Throughout the Cold War, the United States opened US market to foreign exports, ensured the freedom of navigation to protect free trade, and established security guarantees in Europe and Asia. These were American investments made to subsidize a global alliance with the primary aim of combating and deterring the Soviet Union. However, once the Cold War ended, these various security commitments no longer seemed indispensable. Without the Soviets lurking as an existential threat to the United States, justifications for the continuation of the liberal order began to be challenged. Thirty years later, the liberal order is being contested by an American populace that views these commitments as burdensome and costly without any tangible benefits. However, the various suborders are not eroding at equal pace. While the economic and political orders still

maintain incentives for further integration, the global security order is fracturing in part due to US disinterest. A new global order is emerging in which the United States has enduring global reach but waning global interests. This slow retreat from its role as the guarantor of global security spreads doubt about US reliability moving forward, and its foreign policy missteps since the end of the Cold War have only exacerbated and eroded its legitimacy.

Though the West claims to operate within an open, rules-based liberal order, it was the US that often "broke the rules" of the security order during the post–Cold War era. NATO's military intervention in Kosovo, without authorization from the United Nations Security Council (UNSC), was perceived as a violation of international law by Russia and China. Western military action in Iraq and Libya, as well as the passivity with which the West stood by and permitted the invasions of Georgia and Ukraine, undermined the stated core principles of the political order such as the preservation of sovereignty and inviolability of national borders. Additionally, the UNSC has largely failed in its mission "to maintain international peace and security and . . . to take effective collective measures for the prevention and removal of threats to the peace."[23] As a result of increasing American disinterest and a breakdown in a rules-based security order, neo-authoritarian leaders are forecasting a world in which the United States continues to disengage and are choosing to become the guarantors of their own well-being.

The second trend occurs when a global order "proves unable to accommodate a major change in power relations. In some cases, the order collapses because . . . a rising power may reject the role allotted to it by a system it did not design, and the established powers may prove unable to adapt the system's equilibrium to incorporate its rise."[24] The US-dominated unipolarity of the post–Cold War era is slowly transitioning to increasing multipolarity, defined by a more equal distribution of global power. This emerging multipolarity is characterized by a "militarily and economically dominant, but not all-powerful, United States; a rising China and India; a resurgent Russia; an economically potent but militarily declining Europe; an unstable and violence-prone Middle East; and a proliferation of weak and failed states."[25] States such as Russia are relying on its increasing military capabilities to intimidate and coerce political concessions from its neighbors while China, due to its economic successes, is more inclined to challenge the economic and political orders. What's more, the emergence of multilateral institutions such as the Eurasian Economic Union, the Shanghai Cooperation Organization, and the Asian Infrastructure Invest-

ment Bank signifies that these states seek to establish spheres of economic and military influence at the exclusion of the liberal West. These tactics are being utilized with the intent of creating a multipolar global order in which neo-authoritarianism has a more influential role within the global balance of power. Due to these two global trends, the liberal order will continue to face meaningful structural opposition in the twenty-first century. However, through a clear-eyed realization that the global balance of power is shifting as well as an honest effort to restore its legitimacy according to liberal principles, the United States may be able to usher in a more sustainable global order.

An American Grand Strategy for the Emerging World Order

Whether we like it or not, a new global order is emerging: one that is increasingly multipolar and characterized by a growing number of neo-authoritarian states seeking to expand their influence. The current liberal order is ill-equipped to face these primary challenges. Neo-authoritarian states are carving out spheres of influence through the violation of territorial norms because the West has also broken the established rules of the security order, thereby diminishing its legitimacy. Economic coercion undermines the openness of the economic order while still operating within it to build exclusive economic relationships and to exert further political influence. Meanwhile, unconventional political warfare is effective because it erodes confidence in liberal institutions and democratic governance, allowing for the prospect of alternative political models to take root and spread. In light of these realities, our objective cannot simply be promoting the current liberal order in a context where it is deemed illegitimate and is actively being undermined. Instead, "the world's democracies need to begin thinking about how they can protect their interests and defend their principles in a world in which these are once again powerfully challenged."[26] The shifts in world politics today necessitate a revitalization of America's grand strategy for the twenty-first century by redefining its legitimacy as the leader of the liberal order and leveraging its power and influence to shape the emerging global order in its favor. What we seek is an evolution in American foreign policy, and the "current reflexive opposition to multilateralism needs to be rethought" to make sure the transition from one order to another does not result in crisis.[27] Just as the creation of the liberal order thwarted the totalizing ideologies of fascism and communism, our current strategy must reflect the emerging threats to free societies and evolve alongside them.

Grand strategy is "the use of all instruments of national power to secure the state."[28] An effective strategy encompasses the desired political *ends* to be achieved through the utilization of societal *means* and *ways*, and it must be "based on a set of overarching premises and principles that will allow us to chart a consistent general course in the world."[29] The ends of American grand strategy, otherwise known as core national interests, have remained consistent since the establishment of the liberal order, namely the protection of the American people and way of life by securing the homeland, preserving an open and dynamic global economy, and fostering a stable international environment. Means, however, involve all manifestations of a society's power, including but not limited to military, economic, political, and cultural influence. Examples of traditional American means include "a strong and survivable nuclear deterrent, capable military forces that can project power globally, and intelligence services that can ensure global situational awareness."[30] Moreover, these are "intrinsically linked to a powerful economy and industrial base, advanced technology, an educated and technically skilled population, and a political system based on classically liberal democratic values."[31] Most notably, the ways in which a society chooses to resist threats to its core national interests are the most vital aspect of grand strategy because these involve effectively understanding the political environment and employing prudent action when seeking to alter it. For the past 70 years, the principal way in which the United States has secured its core national interests has been through the creation and maintenance of the liberal order. The first step to revitalizing a grand strategy is to put the ends in context and assess how the remaining elements either prevail or transform accordingly.

Securing the Homeland

The primary aim of any grand strategy should and must be the security and defense of the homeland. This encompasses several vulnerabilities that are increasingly threatened. An effective strategy must be able to safeguard the United States from territorial conquest by a foreign adversary, attacks against its citizens and infrastructure—both physical and virtual, and assaults to its institutions vital to sound governance and the advancement of civil society.[32] In an era characterized by the return of great power rivalry, conventional military capabilities continue to threaten the homeland while unconventional political warfare will become more prevalent. Moreover, the threats associated with international terrorism have not subsided with the relabeling and the de-emphasizing of the war on terrorism. In fact, the "threat of nuclear terrorism looms greater than any other nuclear threat

because of the limits of traditional concepts of deterrence."[33] Beyond the purview of direct and violent attacks, various infrastructure networks are more vulnerable than ever before, including "our economy, our utilities, our health care system, and our principal means of communication from a catastrophic cyber-attack."[34] What's more, rival states will continue to utilize virtual platforms to erode confidence in democratic forms of governance, our constitutional values, and multilateral institutions promoting a rules-based security order.

Preserving an Open and Dynamic Global Economy

Essential to both US national security and the prosperity of its citizenry is the preservation of a global economy characterized by openness and dynamism. Of all the suborders that have emerged since the end of World War II, the global economic order is most ubiquitous, having expanded to nearly every state. The lessons learned from the interwar period are that "economic hardship can be immensely destabilizing" and, by contrast, that "global economic development and international economic integration contribute to stability and peace within countries and regions."[35] The inclusiveness of this order has brought about an era of unprecedented economic growth and bound states to practices enshrined within global economic institutions. Increasingly, however, states that consider these rules to be inherently beneficial to American and Western interests seek to undermine regulations they deem illegitimate. Moreover, with the share of American economic influence in decline, the rise of new economic powers such as China and India is cause for concern. Ensuring that these states continue to seek mutually beneficial opportunities through an open economic order rather than exclusive economic advantages will be a central challenge of the twenty-first century. Also, the economic development of growing economies such as Brazil, Mexico, Nigeria, and South Africa offers "enormous opportunities to the world's consumers and producers alike . . . [,] but managing these countries' growth, integrating them fully into evolving regional and global economic institutions, and addressing their concerns will be a challenge that we must meet."[36]

Fostering a Stable International Environment

Following the calamity of the Second World War, the United States learned an important lesson that it had been grappling with since its founding. Americans learned that the security of the homeland and the American way of life are not isolated from circumstances around the

world. Rather, "we learned that aggressors in faraway lands, if left unchecked, would someday threaten the United States."[37] We decided once and for all to play a leading role in global politics. In doing so, we used the nascent liberal order to build a stable international community, understanding that security and prosperity for the nations of the world would help shape "a world environment in which the American system can survive and flourish."[38] We implemented this vision through the establishment of worldwide alliances to alleviate regional security anxieties, the building of international economic institutions to assist in revitalizing the global economy, and the advancement of liberal democratic values to extend the breadth of like-minded states—all making it easier to pursue our interests. However, if states like Russia, China, Iran, and the United States continue to violate principles of international law that prohibit the "use of force against the territorial integrity or political independence of any state," then the rules-based security order we have come to appreciate will soon devolve into a more volatile global arena where states once again resort to unbounded power relations to secure their interests.[39]

With our ends put in context, it is clear that the American advancement of the liberal order as a grand strategic imperative has produced profound and lasting benefits. Not only has it sustained the primary interests of the US, but it has bestowed international legitimacy to American efforts and leadership. At the broadest level, there is no reason not to retain the defense of the liberal order as the foundational goal. However, the global order is changing, and what is needed to secure the liberal order is changing with it. For this reason, the United States must reformulate how it achieves its strategic objective, that is, how it pursues its grand strategy.

Strategic Options: Selective Retrenchment versus Engagement

The first strategic question then is, Do global trends necessitate restraint and the curtailing of US international commitments, or are these trends more favorable for effectively sustaining the liberal order to secure American interests? One possible answer is that due to the adverse effects of American overreach, the path to preserving the liberal order lies not in expansion but in crafting more prudent strategic choices to ensure liberal outcomes and legitimacy, albeit limited in scope. Another advocates for a concerted effort by the US and its liberal partners to make the world safe for democracies by deepening and advancing the liberal order. These two responses encompass the debate regarding US selective retrenchment versus further engagement.[40]

The central argument for selective retrenchment is that the US, while still preeminent, faces limits on what it can achieve in a more challenging international environment. Thus, supporters of the liberal order must be more prudent in selecting when, where, and how to engage. Pursuing a strategy that advocates a scaled-back US presence overseas might "undercut support for anti-American terrorism and reduce the need for other powers to develop their own weapons of mass destruction."[41] Moreover, proponents of selective retrenchment look to the failures of Iraq and Libya, "where liberal inclinations produced decidedly illiberal and counterproductive results," as case studies for overreach and the erosion of US legitimacy.[42] Though the preservation of the liberal order is in the best interest of the United States and the world, "liberal overreach . . . is likely to generate damaging blowback that will weaken the liberal order abroad and undermine its political support at home."[43] Instead, according to champions of retrenchment like John Mearsheimer, the United States should engage militarily only when local powers are unable to effectively balance against an emerging regional hegemon, particularly in Europe, Northeast Asia, and the Persian Gulf due to the strategic significance of these regions.[44]

There are, however, detrimental side effects to retrenchment that may be worse than the risk of overreach. Initially, paring down the US defense posture around the world not only would make it more difficult to preserve existing security commitments but also could further embolden states inimical to US interests.[45] As Robert Gilpin claims, "retrenchment by its very nature is an indication of relative weakness and declining power, and thus retrenchment can have a deteriorating effect on allies and rivals. . . . Rivals are stimulated to 'close in,' and frequently they precipitate a conflict in the process."[46] Most importantly, the signaling of retrenchment to US allies could douse their support for maintaining the liberal order, thereby exacerbating instances of regional instability as well as emboldening the encroachment of neo-authoritarian social models. Retrenchment might simply accelerate the challenges to the liberal order, generating an erosion of rules-based behavior that will prove costlier to address in the future. Without the United States leading a global order that assures stability and inclusiveness, it runs the risk of creating power vacuums that other, less benign forces will happily fill.

Proponents of an engagement strategy argue that though American predominance has indeed declined since the early post–Cold War era, states that support the liberal order maintain geopolitical dominance. The "liberal coalition still commands a clear majority of that power in economic and military terms alike, and at a share far greater than that of any

conceivable illiberal counter-coalition."[47] Additionally, engagement advocates assert that the global allure of liberal ideas and values remains substantial and more resilient than critics claim. The "democratic recession" the world has experienced over the past decade does not "represent a fundamental historical turn away from the liberal ascendancy, but rather a set of difficulties that can be overcome via a sufficient investment of effort and resources by the United States and its liberal partners."[48] Further engagement, therefore, builds upon the many successes we have achieved and plays to our strengths. But to do so effectively, states within the liberal order cannot simply rely on their collective power to serve as a mandate for action in international affairs. Rather, power must be perceived as legitimate if it is to yield a sustainable global order.

Thus, the proponents of retrenchment are correct in that the key to preserving the liberal order is for the United States to be more prudent in its strategic choices, thereby mitigating the consequences of overreach and exhaustion. Careless US interventions without much strategic foresight have validated this main critique of engagement by needlessly inviting the condemnation of much of the international community. But the liberal order need not retrench. Instead, it could take this critique into account when devising a more thoughtful, deliberate engagement strategy—one that seeks to reestablish domestic and international legitimacy. Such a foreign policy agenda would strive to avoid past pitfalls to yield more liberal results. In doing so, not only would more tangible and realistic successes strengthen the liberal order, but collective participation could mitigate domestic exhaustion, enhance engagement's legitimacy globally, and increase the likelihood of pushing back the proliferating influence of neoauthoritarianism. Consequently, "a reinvigorated liberal offensive appears a plausible and potentially rewarding course."[49]

There are inevitable trade-offs with any strategic approach. However, "in the end a forward strategic presence . . . is very useful for American interests," and the US must continue to engage the global order to thwart the challenges of the emerging context.[50] Though the ends of American grand strategy have essentially remained unchanged, the means required to implement them are indeed evolving. American advantages in the global share of economic and military power, though significant, are diminishing relative to regional powers and revanchist regimes.[51] Thus, in a changing environment such as this, with our ends fixed and our relative means eroding, the US must become more clever in its ways to achieve the objectives of its grand strategy. The United States must reform and reinvigorate the liberal order so that it may adapt to the myriad challenges of the twenty-

first century. Success in this endeavor will enable the United States to leverage the full influence of the various suborders in a way that restores domestic and international legitimacy to its foreign policy. Therefore, the US must establish a new and transformational security order, namely a *Concert of Democracies,* as part of a renewed engagement strategy to simultaneously sustain, deepen, and advance the liberal international order.[52]

Concert of Democracies

A Concert of Democracies is not a new idea. During President Clinton's second term, Secretary of State Madeleine Albright established an international coalition known as the Community of Democracies with the principal aim of strengthening democratic institutions, norms, and values around the world. Ivo Daalder and James Lindsay advocated for a Concert of Democracies during the George W. Bush administration to bring together "the world's most capable states in terms of military potential, economic capability, and political weight . . . to prevent and, when necessary, respond to threats to international security."[53] Even Senator John McCain during the 2008 presidential campaign proposed the creation of a global League of Democracies that would largely focus on bringing together "like-minded nations in the cause of peace."[54] However, rather than concentrating on values promotion or interstate aggression, this league would tackle a range of issues including deepening economic ties, managing humanitarian and health crises, and implementing environmental policies to mitigate the harm caused by climate change. Each of these initiatives is admirable and warrants the attention of the global democratic community. However, taking on such an extensive range of issues runs the risk of creating an institution that is utterly ineffectual. Instead, a Concert of Democracies should concentrate on the most immediate threat facing the liberal order: the disintegration of the global security order.

Building the Concert

In its efforts to secure its interests, reestablish legitimacy, and shape the emerging international order, the United States must spearhead the creation of a new global institution capable of reducing the volatility in the security environment it cannot and does not seek to solve unilaterally. The establishment of a Concert of Democracies would serve as the vanguard of a reinvigorated liberal order as US predominance gives way to a more equitable, multilateral global order. Such an institution would collectively manage security in a multipolar world and facilitate burden sharing among

democratic nations. Currently, the international community lacks institutions that are capable of prompt and effective action, and the states of the free world require new means of gauging and granting international legitimacy to its endeavors. Existing institutions like the United Nations Security Council fail to serve this purpose because "they have become hopelessly paralyzed by the split between its autocratic and democratic members."[55] However, the creation of a concert would not replace the authority or influence of current multilateral institutions such as NATO or the United Nations. It would ideally operate within these existing forums, but if they fail to defend and advance the liberal order—as they have often done—then the concert must act independently.

Characterized by shared values, decision-making procedures, and threat perceptions, states within this concert would constitute a "guiding coalition of states at the heart" of the emerging order, a "critical mass of like-minded states that form the center of gravity in international politics."[56] Such a coalition, representing a majority of global defense expenditure and GDP, would reinforce global security guarantees and diminish regional strategic anxieties. Moreover, the concert would serve as a collective forum to more effectively employ competitive and coercive measures to stem neo-authoritarian influence. It could help bestow the desired legitimacy the emerging order requires "on actions that democratic nations deem necessary but autocratic nations refuse to countenance."[57] Conversely, a concert may attempt to shape the behavior of revisionist states through cooperative initiatives as well. It must continue to engage challengers in both regional and global aspects of the liberal order. However, if the chance at cooperation proves unconvincing, the United States can rely on the members of the concert, with their shared interests and values, to make the world a safer place for free societies. It is the case that "orders grow out of broader realities in world politics," and it is time for the states comprising the free world to collectively defend and advance their interests.[58]

A Concert of Democracies would initially encompass a selective group of member states that are not only dedicated to the principles supporting liberal democracy but would also agree to a number of obligations, such as "pledg[ing] not to use force or plan to use force against one another; commit[ting] to holding multiparty, free-and-fair elections at regular intervals; [and] guarantee[ing] civil and political rights for their civilians enforceable by an independent judiciary."[59] This selective group could initially include the United States, NATO and non-NATO European democracies, Japan, South Korea, Australia, and New Zealand. Although these initial members are the most integrated into the security, economic,

and political suborders and will seek to further engage in order-building behavior, the concert need not be exclusionary. Mechanisms must be in place to facilitate the inclusion of emerging democracies seeking to join the liberal community. These emerging democracies might include Brazil, Argentina, India, South Africa, Mexico, Indonesia, and others. The concert's inclusion of these states would garner further legitimacy in that it would "constitute a major effort to integrate non-Western democratic powers into a global democratic order."[60] Though an ever-increasing membership would only benefit the concert, it must be able to enforce penalties or excommunication if member states fail to uphold the obligations outlined within its charter. Thus, the long-term strength of the concert would lie in its legitimacy as an institution to ensure democracy as the foundational element of membership rather than power or historical ties. These measures might serve as a structural framework for an effective concert. However, for it to reduce volatility in the security environment and restore order, a new set of commonly accepted rules would be required that define the limits of permissible behavior. Such a set of directives must be made explicit to garner legitimacy and signal to opposing states the concert's intentions and expectations. What follows is an outline of the specific roles required of such an institution and the strict guidelines for using military force.

A Renewed Security Order

The security role for our concert is twofold. It must sustain alliances by promoting security cooperation among liberal democracies to discourage neo-authoritarian states in their attempts to carve out spheres of influence. Furthermore, it must underwrite the reconstruction of a rules-based security order, one in which the concert serves "as the core military capability of a global veto on interstate aggression."[61] Consequently, several courses of action must follow. First, the United States should "sustain the military predominance of liberal democracies and encourage the development of military capabilities by like-minded democracies in a way that is consistent with their security interests."[62] Maintaining this military predominance is necessary to avert the military adventurism with which revisionist states like Russia, China, and Iran have conducted their foreign policies. Thus, reinforcing the global balance of power in favor of liberal democracies will require elevated defense budgets on behalf of all member states to prevent aggression. Additionally, to legitimately serve as a global veto on aggression, the concert must become an acceptable forum "for the approval of the use of force in cases where the use of the veto at the Security Council

prevented free nations" from defending the liberal order.[63] Codified within its charter, the concert could approve of the use of force by a supermajority of member states, with no veto power. Rather than undermining the Security Council in its efforts to maintain international peace and security, the concert would serve as a legitimate and viable alternative without the obstructionism often employed by neo-authoritarian states.

Though the security roles taken on by the concert will assist in protecting the American people, there will be instances in which the United States acts unilaterally to secure its fundamental interests. Within any institutional relationship there exists a trade-off between the advantages of independent engagement and the benefits of united action. However, the United States must not abuse this prerogative if it seeks to further its interests over the long term. In keeping its decisions to use military force closely tied to concert action, the US will demonstrate its credibility and bestow further credence to the concert as a whole. However, it is beyond the scope of this strategy to outline the utility of unilateral American action. Rather, it is important to stipulate under which circumstances the concert should authorize the use of military force. The concert must be capable of answering several questions provided by Henry Kissinger "to play a responsible role in the evolution of a twenty-first-century world order":

> What do we seek to achieve at all costs, and if necessary, alone?
>
> What do we seek to achieve, even if not supported by any multilateral effort?
>
> What do we seek to achieve only if supported by an alliance?
>
> When should we avoid military force, even if urged by multilateral groups or alliances?[64]

It is helpful to think of these questions under the framework of Miroslav Nincic's three functions of military power: defense, deterrence, and compellence.[65] Each of these functions serves to answer one of Kissinger's questions in a way that ensures the legitimacy of concert or US unilateral action while remaining true to the intent of the concert. Defense can be understood as simply the "repelling of foreign aggression" and involves the destruction of an adversary's capacity to do harm once its intent has been made clear by the application of force. Deterrence focuses on affecting an adversary's intent to use force by "ensuring through threatened retaliation . . . [or denial] . . . that acts against the country's national interest and security are not attempted." Compellence, then, is employed once a provocation has occurred and seeks to "alter, by force, an existing state of affairs in

pursuit of a policy objective."[66] With the functions defined, we can now answer questions in a manner that ensures prudent action.

First, what the concert seeks to achieve at all costs and alone, if necessary, is the *defense of concert members from foreign aggression*. Second, what it seeks to achieve, even if not supported by any non-concert multilateral effort, is *deterrence against foreign aggression through the sustainment of alliances, the promotion of security cooperation, and the adherence to its defense pact obligations*. Subsequently, what it seeks to achieve only if supported by an alliance is *compellence against foreign aggression toward non-concert states, if called upon by said states, to ensure international peace and security*. Lastly, the concert should avoid military force—even if urged by multilateral groups or alliances—*during calls for offensive engagement or cases of intrastate conflict, including civil war, regime change, or humanitarian intervention*. Under these circumstances of intrastate conflict, the concert would preferably intervene by other means, including the provision of economic and political assistance, to facilitate the reconciliation between warring parties. Thus, there is only utility in the application of military force under these limited circumstances where international legitimacy is preserved and power is wielded responsibly to achieve the reinvigoration of the liberal order.

While the three functions of military power justify the use of force to defend concert members, deter against neo-authoritarian aggression, and forcibly coerce states into abiding by the rules-based security order, there are limitations to its utility regarding offensive engagements and intrastate conflicts for several reasons. To start, if a concert were to become involved in these conflicts, it would result in an asymmetry of motivations and political will.[67] The justification for concert engagement would involve ends that it deemed limited, or "discrete policy goal[s] affecting some aspect of the [concert's] interest, not its core purposes."[68] Conversely, the adversary would be fighting for existential reasons such as territorial integrity, national survival, or political survival.[69] This would inevitably lead to a considerable difference in cost tolerance throughout the conflict and limit a concert's ability to achieve its ends. Secondly, there could be consequences resulting from conflicts that involve powerful states pitted against weaker opponents. Such asymmetries in relative power would result in strategic decisions that typically do not favor a powerful coalition. According to Ivan Arreguin-Toft, each side in an asymmetric conflict can choose either a "direct" strategy to eliminate an adversary's armed forces or an "indirect" one that focuses on weakening the opponent's political will. The more powerful state, especially a Concert of Democracies, is essentially incapable of adopting an indirect strategy because it would involve "depredations against

non-combatants," and such "barbarism" would not be tolerated by the international community.[70] Consequently, the concert would face a constraint in its strategic choices and thus be likely to lose an asymmetric conflict. Finally, there is often the assumption that external intervention in internal conflicts can solve problems that do not capitulate to force. This perspective tends to "view military victory as an end in itself, ignoring war's function as an instrument of policy."[71] This isn't to say that the concert should never intervene in internal conflicts, but rather that the application of force will fail to produce desired political outcomes. Instead, the remaining dimensions of societal power (political, economic, and cultural) are better suited to attain policy goals that are resistant to coercive action.

Conclusion

While serving as secretary of state, John Quincy Adams famously declared on 4 July 1821 that America "goes not abroad in search of monsters to destroy." However, he insisted that America would always champion the pursuit of liberty and that "wherever the standard of freedom and independence has been or shall be unfurled, there will her heart, her benedictions, and her prayers be."[72] This message underscores the manner in which strategy for the twenty-first century should be conceived. Constructing and leading the liberal international order has been the focal point of American grand strategy since the end of the Second World War. However, the current global order faces immense challenges, and the emerging context will not privilege American strategic interests. Without addressing the erosion of liberal legitimacy and the emergence of a more multipolar global order, the liberal international order as a grand strategic project cannot survive. By striking the balance between legitimacy and power, the US can lead a guiding coalition that represents a critical mass of states seeking to further engage in liberal order building. To do so effectively, this Concert of Democracies must galvanize the world's value-sharing democracies into action and seek deeper levels of cooperation with all states, depending on the issue and suborder at stake. It must work together to reconstruct a global order that is compellingly rules based, that is, free from interstate aggression. Only such an order can dissuade neo-authoritarian challengers, embolden the free world to advance its interests, and offer all states a critical and viable choice.

Reinhold Niebuhr often warned against the excessive use of American power in world affairs. Yet he also believed that "the world problem cannot be solved if America does not accept its full share of responsibility in solving it."[73] In this sense, he and Adams recognize that the United States is

truly indispensable in the defense and pursuit of liberty. The establishment of a new security order would be a foundational step in accepting this global responsibility while ensuring America shares the responsibility and burden with the rest of the free world. Moreover, "the future international order will be shaped by those who have the power and the collective will to shape it."[74] The creation of a coordinated, self-identifying Concert of Democracies would go a long way toward aggregating the necessary power and collective will needed to shape the emerging world order in our favor. SSQ

Scott Lawless

Scott Lawless currently works for Booz Allen Hamilton consulting its defense clients to fulfill their mission needs. He is a returned Peace Corps volunteer and earned his master of arts degree in international security from the University of Denver, Josef Korbel School of International Studies.

Notes

1. Henry Kissinger, *World Order* (New York: Penguin Books, 2015), 9.
2. Hans Kundnani, "What Is the Liberal International Order?," Policy Essay no. 17 (Washington, DC: The German Marshall Fund of the United States, April 2017), 4, http://www.gmfus.org/.
3. G. John Ikenberry, *Liberal Leviathan* (Princeton, NJ: Princeton University Press, 2011), 18.
4. Joseph E. Stiglitz, *Globalization and Its Discontents* (New York: Norton, 2002), 9.
5. Kundnani, "What Is the Liberal International Order?," 2.
6. Kundnani, 2.
7. Kundnani, 2.
8. Francis Fukuyama, "The End of History?," *The National Interest*, no. 16 (Summer 1989): 3–18; and Joseph S. Nye, Jr., "What New World Order?," *Foreign Affairs* 71, no. 2 (Spring 1992): 83–96.
9. Will Marshall, "Three Threats to Liberal Democracy," Commentary Prepared for the Biennial Colloquy on the State of Democracy, Loyola University Chicago, John Felice Rome Center, Rome, Italy, 10–11 April 2018, Progressive Policy Institute, 4–5, https://www.progressivepolicy.org/.
10. Robert Kagan, *The Jungle Grows Back: America and Our Imperiled World* (New York: Alfred A. Knopf, 2018), 10; and Marshall, "Three Threats to Liberal Democracy," 4–5.
11. Department of Defense, *Summary of the 2018 National Defense Strategy of the United States of America: Sharpening the American Military's Competitive Edge* (Washington, DC: Department of Defense, January 2018), 1, https://dod.defense.gov/.
12. Kundnani, "What Is the Liberal International Order?," 1–9.
13. Department of Defense, *Summary of the 2018 National Defense Strategy*, 1.
14. Peter Harrell, Elizabeth Rosenberg, and Edoardo Saraville, "China's Use of Coercive Economic Measures," Center for a New American Security, 11 June 2018, 2, https://www.cnas.org/.
15. Harrell, Rosenberg, and Saraville, 12.

16. Thomas Wright, "The Return to Great-Power Rivalry Was Inevitable," *The Atlantic*, 12 September 2018, https://www.theatlantic.com/.

17. Seth G. Jones, "The Return of Political Warfare," Center for Strategic and International Studies, 2 February 2018, 1–4, https://www.csis.org/.

18. Jones, 1–4.

19. Robert Kagan, *The Return of History and the End of Dreams* (New York: Alfred A. Knopf, 2008), 59.

20. Wright, "Return to Great-Power Rivalry."

21. Kissinger, *World Order*, 365–66.

22. Richard Haass, "How a World Order Ends: And What Comes in Its Wake," *Foreign Affairs* 98, no. 1 (January/February 2019): 28, https://www.foreignaffairs.com/.

23. U.N. Charter, chap. 1, art. 1, June 1945, https://treaties.un.org/.

24. Kissinger, *World Order*, 366–67.

25. R. D. Hooker, Jr., "American Grand Strategy," in *Charting a Course: Strategic Choices for a New Administration,* ed. R. D. Hooker, Jr. (Washington, D.C.: National Defense University Press, 2016), 3, https://ndupress.ndu.edu/.

26. Kagan, *Return of History*, 97.

27. Haass, "How a World Order Ends," 4.

28. Hooker, "American Grand Strategy," 1.

29. G. John Ikenberry and Anne-Marie Slaughter, *Forging a World of Liberty under Law: U.S. National Security in the 21st Century*, Final report of the Princeton Project on National Security (Princeton, NJ: Woodrow Wilson School of International Affairs, Princeton University, 2006), 14.

30. Hooker, "American Grand Strategy," 6.

31. Hooker, 6.

32. Ikenberry and Slaughter, *Forging a World of Liberty under Law*, 14.

33. Ikenberry and Slaughter, 14.

34. Ikenberry and Slaughter, 14.

35. Ikenberry and Slaughter, 15.

36. Ikenberry and Slaughter, 15.

37. Ikenberry and Slaughter, 16.

38. Ernest R. May, *American Cold War Strategy: Interpreting NSC-68* (New York: Bedford Books of St. Martin's Press, 1993), 40.

39. U.N. Charter, chap. 1, art. 2.

40. Hal Brands, *American Grand Strategy and the Liberal Order: Continuity, Change, and Options for the Future* (Santa Monica, CA: RAND Corporation, 2016), 17, https://www.rand.org/.

41. Colin Dueck, "The Strategy of Retrenchment and Its Consequences," Foreign Policy Research Institute, E-Notes, April 2015, https://www.fpri.org/.

42. Brands, *American Grand Strategy*, 17.

43. Brands, 16.

44. John J. Mearsheimer, "Imperial by Design," *The National Interest*, no. 111, (January/February 2011): 16–34, https://www.jstor.org/.

45. Brands, *American Grand Strategy*, 18.

46. Robert Gilpin, *War and Change in World Politics* (Cambridge: Cambridge University Press, 1983), 194.

47. Hal Brands, "Dealing with Allies in Decline: Alliance Management and U.S. Strategy in an Era of Global Power Shifts," Center for Strategic and Budgetary Assessments, May 2017, 13–29, https://csbaonline.org/.
48. Brands, *American Grand Strategy*, 13.
49. Brands, 14.
50. Dueck, "Strategy of Retrenchment and Its Consequences."
51. F. G. Hoffman, "U.S. Defense Policy and Strategy," in Hooker, *Charting a Course,* 41.
52. Michael J. Mazarr, *Summary of the Building a Sustainable International Order Project* (Santa Monica, CA: RAND Corporation, 2018), 18, https://www.rand.org/.
53. Ivo Daalder and James Lindsay, "Democracies of the World, Unite," *The American Interest* 2, no. 3 (January/February 2007): 1–2.
54. John McCain, "An Enduring Peace Built on Freedom: Securing America's Future," *Foreign Affairs* 86, no. 6 (December 2007): 1–17, www.jstor.org/.
55. Kagan, *Return of History*, 97.
56. Mazarr, *Building a Sustainable International Order Project*, 3–4.
57. Kagan, *Return of History*, 98.
58. Mazarr, *Building a Sustainable International Order Project*, 11.
59. Ikenberry and Slaughter, *Forging a World of Liberty under Law*, 26.
60. Ikenberry and Slaughter, 26.
61. Mazarr, *Building a Sustainable International Order Project*, 19.
62. Ikenberry and Slaughter, *Forging a World of Liberty under Law*, 29.
63. Ikenberry and Slaughter, 26.
64. Kissinger, *World Order*, 372.
65. Miroslav Nincic, *The Logic of Positive Engagement* (Ithaca: Cornell University Press, 2011), 3–4.
66. Nincic, 3–4, 7.
67. Nincic, 11.
68. Nincic, 11.
69. Nincic, 11.
70. Ivan Arreguin-Toft, *How the Weak Win Wars: A Theory of Asymmetric Conflict* (Cambridge: Cambridge University Press, 2005): 9.
71. Jeffrey Record, "Why the Strong Lose," *Parameters* 35, no. 4 (Winter 2005): 24.
72. John Quincy Adams, "Address on U.S. Foreign Policy," 4 July 1821, Presidential-Rhetoric.com, http://www.presidentialrhetoric.com/.
73. Reinhold Niebuhr, "American Power and World Responsibility," *Christianity and Crisis* 3, no. 5 (5 April 1943): 2, https://providencemag.com/.
74. Kagan, *Return of History*, 105.

Shadows on the Wall: Deterrence and Disarmament by Keith B. Payne. National Institute Press, 2020, 187 pp.

Shadows on the Wall: Deterrence and Disarmament is the latest installment of deterrence thought from Dr. Keith B. Payne and takes on the strategic nuclear deterrence policy debate from a unique perspective. This well-conceived and well-researched book reviews three competing philosophical viewpoints regarding expectations of human and state behavior vis-à-vis nuclear weapons and strategic deterrence within the current international system. These competing narratives share the same goal of precluding nuclear war but envision very different routes—from nuclear disarmament to the preservation of robust nuclear capabilities. These are the philosophical foundations for the contending arguments in the US nuclear policy debate. While Payne concludes as he begins, that "nuclear war must be prevented and deterrence remains a critical tool for this purpose," his assessment of these three narratives can educate the reader using a framework and acumen to inform effective nuclear war prevention strategies.

Dr. Payne is cofounder of the National Institute for Public Policy and professor emeritus at Missouri State University. He contributed to the 2018 *Nuclear Posture Review* and authored, coauthored, or edited over 100 published articles and 17 books and monographs.

The central thesis of this book is an assessment of the contending philosophies or "narratives" underlying the US nuclear policy debate from the 1960s to the present. This assessment is constructed around three narratives: a nuclear disarmament assertion reflecting idealist thought and two very different deterrence approaches that share some initial points of realist thought. Payne labels these latter two narratives "easy" deterrence and "difficult" deterrence.

First, Payne delivers perhaps the most efficacious and contextual understanding of realism and idealism as they pertain to competing world views and national priorities. These philosophies are informed by varying conjectural expectations of human and state behavior within the contemporary international system. He reminds us that for realists, the enduring interstate system is an anarchic "self-help" world that involves competition and the potential for aggression and conflict. Conversely, Payne reviews the idealist's anticipation of a cooperative global world order and goal of transforming the international system into one that facilitates and enforces peaceful resolutions of interstate conflict. These two divergent perspectives of the world form the context for his elegant presentation.

Next, Payne deconstructs the idealist's goal of international transformation and nuclear disarmament as the means to remove the omnicidal risk of nuclear war. The belief is that the current international order can be transformed via a rigorous, mutually complaisant effort so compelling that individual states willingly surrender their nuclear arsenals in favor of "alternative global security mechanisms." Payne surmises that to a nuclear idealist, the continued existence of nuclear arsenals poses a greater risk to global security than would their voluntary retirement, and a policy of nuclear deterrence is "an impediment to disarmament because it suggests a positive and important value for nuclear weapons." Unfortunately, Payne opines that international transformation and disarmament demand a preceding level of enlightenment, mutual trust, and cooperation that has not been seen in the history of mankind and generally is not deemed plausible by realists.

Payne then presents two alternatives to the idealist nuclear disarmament narrative. Couched as "easy" deterrence and "difficult" deterrence, Payne's bifurcated expressions of nuclear deterrence have common realist starting points but diverge from there. His assessment of these competing alternatives offers the reader a cogent understanding of deterrence that rivals the Kahn-versus-Schelling principles. Since international co-

operation cannot be expected and "the world lacks an overarching authority with sufficient power to regulate interstate behavior reliably and predictably," nation-states must act in their own national interests, sovereignty, and security. Consequently, states are "on their own" to pursue sufficient power to ensure their own existence and purpose. In the realist's worldview, nation-states generally act in their own survival interest first and foremost. For the realist, "nuclear weapons are a symptom of the enduring realities" of today's international system, according to Payne. His narratives of "easy" versus "difficult" deterrence provide a splendid framework by which to consume this expert's rationale.

Under "easy" deterrence, Payne posits that the "essential requirements for stable mutual deterrence are easy to understand, easy to meet, and are largely predictable and reliable." This narrative, derived from the works of Schelling and Waltz, relies on rational or "sensible" adversaries, "crystal ball" effects, and relatively modest second-strike nuclear capabilities. The key is an obviously easy mental transaction based on mutual fear of intolerable catastrophe or existential destruction. However, Payne carries this deterrence narrative into a clarity that any layman can comprehend. For the modern idealist, the disarmament narrative envisions the fear of nuclear war as a catalyst to enable global disarmament and enlightened transformation. Contrastingly, the "easy" narrative envisages the fear of nuclear war as a reliable means for minimizing the potential of actual nuclear war. These are two very different routes to the same goal of precluding nuclear war.

"Difficult" deterrence, Payne theorizes, shares the goal of precluding nuclear war but acknowledges that deterrence is a never-ending and messy pursuit of peace and stability, devoid of standard formulas or fully predictable behavior. Unlike "easy" deterrence, "difficult" deterrence does not assume all rational adversaries would behave in a foreseeable manner or necessarily calculate the costs and benefits of war akin to American values. Payne clearly describes "difficult" deterrence as an ongoing, complex challenge "with no fixed approach and no corresponding finite and fixed set of nuclear capabilities that can predictably provide the desired deterrent effects." Moreover, he states, to think otherwise would be a "fatal error." Thus, the lesson of this narrative is that deterrence strategies must be "tailored" to each adversary and account for each opponent's characteristics, values, and goals—an effort made difficult because it is imprecise and ever-changing.

Of the three narratives explored, idealism and "easy" deterrence offer society much greater comfort and perhaps a false sense of stability and security. Idealism projects a new and more peaceful world order without nuclear weapons while "easy" deterrence expects deterrence to preclude nuclear conflict without the need to transform the international order. Payne's clear-eyed assessment questions both the idealist solution of a timely, profound transformation of the international system and the "easy" deterrence expectation that all sensible leaders will respond with predictable caution if confronted with a nuclear deterrent threat to their societies. He adds, however, that the "difficult" deterrence narrative offers little comfort or ease; it alone confronts the two apparent realities that the timely, global, and cooperative transformation necessary for disarmament is unlikely and that effective deterrence ultimately is far from easy "because leadership decision-making is variable and unpredictable." Payne concludes that this is the challenge that must continually be met because "nuclear war must be prevented and national security preserved."

This book is a must-read for those serving in the nuclear enterprise or those interested in international relations. Dr. Payne's 187-page disquisition presents the most

cogent review of today's competing nuclear narratives, and his conclusions provide a new framework by which to devise a strategy to achieve a stabilizing deterrence effect.

Curtis McGiffin
Associate Dean, School of Strategic Force Studies
Air Force Institute of Technology

The Future of Strategy by Colin S. Gray. Polity Press, 2015, 150 pp.

Colin Gray is one of the most important strategy scholars of our time. He advised the US and UK governments and authored more than 30 books and numerous articles on a broad range of topics, from nuclear weapons to geopolitics to the theory of strategy. In many ways *The Future of Strategy* is a summary of Gray's previous ideas (he cites 10 of his own books in the 117 pages of text). If one did not want to read or assign to students 364 pages of *Modern Strategy* (Oxford University Press, 1999) or 257 pages of *The Strategy Bridge* (Oxford University Press, 2011), then *The Future of Strategy* captures many of Gray's previous ideas, though in much less detail.

Gray suggests that the book's purpose is to show the universal qualities of strategy. The introduction and chapters 1 and 2 discuss Gray's intent and arguments about the need for a theory of strategy. Chapters 3 and 4 emphasize the difference between the theory of strategy, which remains unchanged throughout history (and will remain so into the future), and the practice of strategy, which is altered by time, location, and technology, among other things. Chapter 5 touches on grand strategy and geostrategy. Chapter 6 and the conclusion focus on the threat of nuclear exchange and how that would void his arguments about future strategy.

On initial reading, the lack of new ideas was disappointing, and anyone familiar with Gray's work will likely react the same way. *The Strategy Bridge* considers the permanent nature but changing character of strategy and differentiates between a theory of strategy and strategies (aka plans), while *Modern Strategy* explores how the context of strategy changes as well as the relationship between strategy and politics. On closer inspection, though, and after rereading some of his previous work, although *The Future of Strategy* restates many of his ideas, it also shows some of the evolution of Gray's thinking.

Gray has written so much on strategy that some of his ideas in *The Future of Strategy* seem to contradict earlier thoughts, without mention of how his thinking evolved. The 23 dicta (pp. 47–48) that make up his general theory of strategy are incredibly useful for classroom discussions and are easily defensible as necessary elements of strategic thought. However, they are similar to the 21 dicta in *Strategy Bridge* and to many of the 40 maxims in *Fighting Talk* (Praeger, 2007). Many of those same 23 dicta reappear in the more recent *Theory of Strategy* (Oxford University Press, 2018), though they are now key principles and are organized differently. It is not a problem that his ideas evolve over time; in fact, that is admirable. But if the same concepts appear as maxims, dicta, and then principles, these revisions might confuse rather than clarify our understanding of strategy, and it would be useful to know what prompted the changes.

While Gray's definition of strategy appears to remain constant, there is a shift somewhere between *Modern Strategy*, where strategy involves only the military instrument of power, and *Strategy Bridge*, where it now includes any instrument of power (Gray has separate definitions for *strategy* and *military strategy* to account for this shift). This change suggests, probably accurately, that a strategy can involve nonmilitary instruments and still achieve political aims.

One weakness in Gray's work is his insistence on equating models with theory. A theory explains a phenomenon while a model is a representation of reality. Gray's general theory of strategy involves his 23 dicta as well as the ends/ways/means model of strategy.

By incorporating all of these into his theory, Gray both complicates and confuses the concept of a theory of strategy. Carl von Clausewitz claims that the nature of war is a violent, political, contest of wills; that is not a theory of war but a description of it. A theory of war is more akin to Thucydides's argument that war stems from fear, honor, and interest.

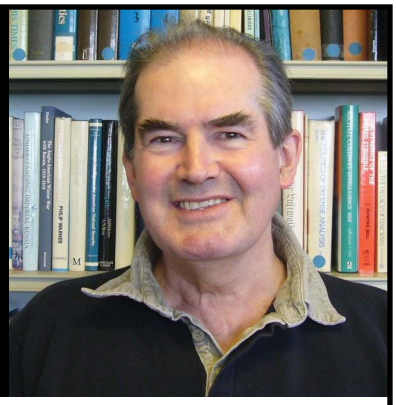
An appropriate parallel comes from physics. The laws of physics explain the unchanging nature of the world around us (its essence). Theoretical physics is a discipline intended to explain behavior as constrained by those laws. The theories rarely change, but could if confronted with new information or new technology (character). Applied physics and engineering put those theories to the test (conduct). To connect strategy with Clausewitz's view of war, I would offer that the nature or essence of strategy involves many of Gray's dicta—it is political, it is a bridge between politics and operations, and so forth. The ends/ways/means of strategy may be unchanging in the abstract, but that is not the essence of strategy—it is a model of strategic choices (or the character of strategy). Finally, strategies and plans are implementations of that model and represent the conduct of strategy.

One final issue is that Gray's discussion of strategy rarely addresses risk, and it is not a part of *The Future of Strategy*. His ends/ways/means model incorporates assumptions, and there is probably a link between those concepts; the more assumptions one makes, and the more heroic those assumptions, the greater the risk to one's strategy. More exploration of that connection would have been a welcome addition.

Like all of Colin Gray's books, *The Future of Strategy* is an important read for those who think about, teach, or create strategy. There is not much new material if one is familiar with his earlier work, but it is a more condensed version of his ideas and therefore more accessible to those first engaging the study of strategy. At the same time, this book seems to bridge his earlier work and the more recent *Theory of Strategy*, which further develops his ideas. While I disagree with some of Gray's views of strategy, his contribution to the field is immeasurable, and *The Future of Strategy* is another example.

Dr. Gregory D. Miller
Air Command and Staff College

In Memoriam
Colin S. Gray
1943—2020



We honor the memory of author, scholar, and strategist Dr. Colin S. Gray, one of the original *Strategic Studies Quarterly* advisers. As a dual US-UK citizen, he served as a strategy and policy consultant in Washington, DC, and London. Dr. Gray earned a bachelor's degree in economics from the University of Manchester in 1965 and a doctorate in international politics from Oxford in 1970. Among his published works are more than 30 books and 300 articles, including *Airpower for Strategic Effect* (2012), *The Airpower Advantage in Future Warfare* (2007), and *Understanding Airpower* (2009) with Air University Press.

Professor Gray taught at universities in Britain, Canada, and the United States; became assistant director of the International Institute for Strategic Studies (London); and worked with Herman Kahn at the Hudson Institute. In 1981 he founded the National Institute for Public Policy and then served on the President's General Advisory Committee on Arms Control and Disarmament. In April 1987 he was presented the Superior Public Service Award by the US Department of the Navy. In 1997–98 he served on the Panel of Experts on the UK Strategic Defence Review. Dr. Gray was also a member of advisory panels for the Congressional Office of Technology Assessment (strategic defense initiative and space weapons), the Department of the Army (tactical nuclear weapons), the Department of the Air Force (innovations), and the US Space Command (future of space forces).

Few scholars were ever more influential in the theory and practice of strategy, the dialogue about policy versus military force, or the value of history for educating policy makers. He will be sorely missed by Air University and *Strategic Studies Quarterly*.

Mission Statement

Strategic Studies Quarterly (*SSQ*) is the strategic journal of the United States 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 US Air Force, 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 US 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@us.af.mil

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@us.af.mil

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