Graham Allison and the Thucydides Trap Myth
Richard Hanania

More Is Not Always Better: Oversight of the Military
Marie T. Harnly

North America’s Imperative:
Strengthening Deterrence by Denial
Andrea Charron and James Fergusson

Will Emerging Technology Cause Nuclear War?:
Bringing Geopolitics Back In
Matthew Kroenig

Sophons, Wallfacers, Swordholders, and the Cosmic Safety Notice:
Strategic Thought in Chinese Science Fiction
Wendy N. Whitman Cobb

Six Steps to the Effective Use of Airpower
Joseph B. Piroch and Daniel A. Connelly

Cultivating Future Airpower Strategists
John G. Terino Jr.

A Case for Strategic Design
Brian R. Price

Leveraging Regional Partners
Sean P. Braniff

Outline of Strategic Aerial Culture
Mickaël Aubout
3 From the Editor

8 Strategic Studies Quarterly: From the Beginning
   Anthony C. Cain

SPECIAL COMMENTARY

13 Graham Allison and the Thucydides Trap Myth
   Richard Hanania

CIVIL-MILITARY RELATIONS

25 More Is Not Always Better: Oversight of the Military
   Marie T. Harnly

DETERRENCE

42 North America’s Imperative: Strengthening Deterrence by Denial
   Andrea Charron and James Fergusson

59 Will Emerging Technology Cause Nuclear War?: Bringing Geopolitics Back In
   Matthew Kroenig

CHINA

74 Sophons, Wallfacers, Swordholders, and the Cosmic Safety Notice: Strategic Thought in Chinese Science Fiction
   Wendy N. Whitman Cobb
89 Six Steps to the Effective Use of Airpower
On “The Drawdown Asymmetry: Why Ground Forces Will Depart Iraq but Air Forces Will Stay”
Joseph B. Piroch and Daniel A. Connelly

101 Cultivating Future Airpower Strategists
On “Developing Twenty-First-Century Airpower Strategists”
John G. Terino Jr.

106 A Case for Strategic Design
On “A Diplomatic Surge in Afghanistan”
Brian R. Price

112 Leveraging Regional Partners
Sean P. Braniff

118 Outline of Strategic Aerial Culture
Mickaël Aubout, translated by 1st Lt Casey Evans, Capt Marie Gaudreault, and Capt Cody Anderson
FROM THE EDITOR

Dear Reader,

In 1947, commanding general of Air University Major General Muir S. Fairchild established *Air University Quarterly Review*. In the inaugural issue, Fairchild wrote that the fledgling “journal of Air Power will not be just another news-magazine, nor is it intended as a periodical of interest only to the Air University. Rather, it will be a professional publication . . . and will reflect not only the highest scholastic standards and educational accomplishments of Air University, but also—and more important, perhaps—the best professional thought concerning global concepts and doctrines of air strategy and tactics.”

Over the years, the flagship journal of the Air Force has featured content ranging from the finer points of air and space operations to national security strategy and international security. A sample of articles from a 1977 issue of then-titled *Air University Review* includes “Nuclear Proliferation and U.S. Security,” “Air-to-Air Training under the DOC System,” and “Deterrence: Reckless Prudence.”

Until the turn of the century and, in particular, before the advent of the internet and social media, global events and a global, strategic interest on the part of readers tended to be confined to certain echelons within the Air Force and Air Force professional military education. At the same time, air and space operations became a priority focus for the journal, renamed a fifth time in 2002 to *Air & Space Power Journal (ASPJ)*. By the late 1990s and early 2000s, content in the journal was predominantly focused on the operational level, and the journal had a dearth of contributions by authors outside of the service. In 2006, Air University’s commander, then-Lieutenant General Stephen R. Lorenz, recognized this intellectual gap and directed the establishment of *Strategic Studies Quarterly (SSQ)*.

This winter issue marks the fifteenth year of this highly respected publication. What began as an effort to showcase strategic thought of leading national security intellectuals has developed into a well-respected strategic military journal on par with its long-standing sister-service strategic journals.

Over the past decade and in an accelerated manner likely unanticipated by most, social media has created access to and, by extension, expanded general interest in global affairs and national and international politics. Airmen and Guardians of today scroll through daily feeds that include news of new air and space platforms, People’s Republic of China military activities in the Taiwan Strait, Russia massing troops on the border of Ukraine, crowded low Earth orbit, and leadership and manage-
ment theories and lessons learned. Strategic interest is no longer solely the purview of scholars and practitioners including Air Force and national security leaders and their staffs. Indeed, lines between operations, national security strategy, and international relations are ever more opaque—intellectual seams increasingly less relevant to today’s informed Airman and Guardian.

A true flagship Air Force journal effort of the twenty-first century needs, then, to encompass the overlapping, interdependent realms of operations and national and international security. With this in mind, in 2022, Air University Press will bring SSQ and ASPJ under one web platform that will serve as the Air Force flagship journal effort. The platform will also give the Press the opportunity to feature additional digital content such as planned and potential partnerships with the Thomas N. Barnes Center for Enlisted Education, NATO and EU military air and space journals, and outreach to Historically Black Colleges and Universities.

Further, with the veritable explosion of online journals, it is increasingly important that journals signal immediately their association with their organization as well as their focus. Accordingly, ASPJ will take on a slightly altered name, Air & Space Operations Review (ASOR), highlighting that publication’s focus on operations. In the same vein, SSQ will become Æther: A Journal of Strategy and Airpower, making explicit its Department of the Air Force affiliation and acknowledging in advance the likely emergence of a US Space Force-sanctioned journal in the next few years. While the content of both will remain largely the same, SSQ’s new name embraces the journal’s emphasis on highlighting the relevance of national and international security policy concepts and theory to the Department of the Air Force.

Moreover, Chief of Staff of the Air Force General Charles Q. Brown Jr. recently called on Air University to be a leader in innovative thought and theory, noting that in addition to traditional deterrence measures, cultural understanding was vital to combating our adversaries. Understanding an adversary’s mind via culture sometimes requires an unconventional approach. Æther will be looking for such content—rigorous, scholarly, and intellectually creative.

A discussion of content is incomplete without a note about sanctioned military journals. Although taxpayer funded, by charter and regulation and residing under professional military education institutions, these service journals operate with academic freedom. One only need look back through the pages of almost 75 years of the Air University Press family of journals to see that authors and editors have shown no reluctance to ad-
dress sensitive issues or to air well-deserved criticism of Air Force leaders and initiatives—a true loyal opposition when warranted.

The first editor and editorial board of Air University Quarterly Review noted in the inaugural issue that “if the appearance here of articles which may not agree with accepted policy, or even with majority opinion, will stimulate discussion and provoke controversy, an important part of this journal’s mission will have been accomplished: to induce airmen to have original thought on these matters and to give these thoughts expression.”

In addition to continuing the 75-year tradition of academic freedom, both journals will continue to be refereed, and issues will be published quarterly and limited in page count. Due to significant budget cuts, ASOR will move to a digital-only publication, something that journal has done in the past. Æther will continue a print distribution. The journals will also continue to solicit content from authors associated with the Air Force as well as voices from outside the service. The platform will host a single book review effort, and submission requirements for both journals will be largely identical.

As the last issue under the current name, this SSQ winter edition features a selection of content including a special commentary, articles, and retrospectives. The issue begins with a contribution from SSQ’s first editor, Anthony C. Cain, which details the journal’s conceptualization, launch, and first few years. In our special commentary, Richard Hanania questions the application of Graham Allison’s notion of the Thucydides Trap to the contemporary US-China relationship.

Leading our articles, Marie Harnly finds the Air Force experiences varying degrees of autonomy when Congress and the president differ on policy issues, which is at odds with classic tenets of civil-military relations theory. Our forum on deterrence features two articles. Andrea Charron and James Fergusson advocate for a joint Canada-US deterrence-by-denial approach to continental defense. Matthew Kroenig argues strategic stability can only be guaranteed by encouraging technology-related policies that bolster US, Ally, and partner militaries. In our final article, Wendy Whitman Cobb analyzes four strategies in a popular Chinese science fiction trilogy to highlight the ways in which Chinese strategic thinking differ systematically from Western modes of thinking.

Our Reconsidered forum includes four contributions primarily from Air University faculty. Dan Connelly and Joseph Piroch contend then-Lieutenant Colonel Clinton S. Hinote makes critical assertions that underpin airpower efficacy today. Continuing in the vein of airpower strategy, John Terino Jr. highlights the prescience of a 2008 article by
Major General R. Michael Worden considering future challenges to American airpower strategists such as technology, information, funding, and recapitalization.

Turning to planning, Brian Price considers Daryl Morini’s 2010 call for a diplomatic surge in Afghanistan and the lack of strategic design in the Coalition approach to the Afghanistan War. Closing our forum, Sean Braniff finds that while Robert Ross’s 2013 prediction that China would be the main focus of US grand strategy has come to fruition, his opposition to engaging mainland East Asian nations in that strategy in meaningful ways was shortsighted.

Our Winter issue also includes a recently added element, Par Avion, which features translated articles from our NATO and EU Allies and partners—a section that will continue in Æther. In this issue, Par Avion features a contribution from Mickaël Aubout, a French Air Forces officer, discussing the notion of a shared, international aerial strategic culture.

In closing, the editor would like to recognize the hard work of the former editors of SSQ, Dr. Anthony C. Cain, Colonel, USAF, retired, and Mr. W. Michael Guillot, Colonel, USAF, retired; myriad contributing editors and reviewers; and the skilled and dedicated staff of the journal, some of whom have retired and those still with the Air University Press family, namely Tawanda Eaves, Jeanne Shamberger, Megan Hoehn, Donna Budjenska, and Nedra Looney. Without these individuals’ vision, commitment, excellence, and teamwork, SSQ would not have succeeded. We are grateful for the efforts of the SSQ team, past and present, and we intend to honor and continue that commitment to excellence and rigor as the Air Force flagship journal effort sets its sights on its 100th birthday in 2047.

When he became Chief of Staff of the Air Force, General Brown called on the Department to “accelerate change or lose.” Air University Press takes this charge seriously and welcomes leadership that advocates for dynamic adaptation to the realities of today’s complex physical and intellectual universe.

~The Editor
Letter from the Editor

Notes

2. Editor and Editorial Board, “Editorial,” 94.
I
n January 2007, then-Lieutenant General Stephen R. Lorenz, the Air University commander, asked me to start a new journal for the Air Force focused on strategic matters. He described a journal that would include articles from military, government, and academic professionals interested in how air, space, and cyber power contributed to national security. General Lorenz said he wanted something like *Foreign Affairs* (the long-standing publication of the Council on Foreign Relations), *Parameters* (the US Army War College’s quarterly journal), and the *Naval Institute Proceedings* (a publication of the independent Naval Institute).

I told him the three journals he described targeted very different audiences—one was clearly a flagship journal for the international relations community; the other aimed to reach graduates of the Army War College; the third was a forum for readers interested in matters related to operations and strategy for the Sea Services. Creating a journal that was sustainable and that would attract readers from a wide array of professional perspectives was a challenge that seemed both exciting and daunting—a challenge I enthusiastically took on.

Immediately after I told General Lorenz I would accept the challenge of becoming the founding editor in chief of the new journal, which he had already named *Strategic Studies Quarterly* (*SSQ*), he asked me when I would publish the first issue. I thought of all the things we needed to do: hire a staff (there would be a managing editor, an editorial assistant, and me), design the cover, procure a printing contract, determine the number of pages in each issue, design the interior format, coordinate production schedules with Air University Press, develop distribution lists, create peer-review processes, and, most importantly, collect articles aimed at General Lorenz’s vision of the target market. I told him we would publish the first issue on September 1, 2007, nine months from the date I was hired. In true Lorenz fashion, he said, “Let me know what I can do to help.” He also added, “I’ll hold you to it.”

My first task was to hire my staff. Luetwinder (Tawanda) Eaves was the new journal’s managing editor. Tawanda and I had worked together at *Air & Space Power Journal*, and in my opinion, there’s no one better at managing the thousands of tasks required to publish a professional journal. To
many of our authors, Tawanda was the journal. She accepted their draft articles, explained our peer review processes, assigned peer reviewers, served as the conduit between the reviewers, the authors, and me, and made sure we met all our timelines to publish the journal on time.

Later, as we added contributing editors and an editorial advisory board, she coordinated the meetings of those groups. Through her efforts, SSQ established a solid reputation of never missing a publication date, and, more importantly, our contributing authors had a positive experience with the publication process.

Betty Littlejohn was our editorial assistant. Betty did everything imaginable to make SSQ successful. She managed our book review program by requesting the latest books on strategy, national security policy, international relations, economics, history, and air, space, and cyber power from publishing houses. She advertised the books we had available, provided them to reviewers, and passed along the completed reviews to me and Tawanda for inclusion in the journal. She planned and coordinated the travel, lodging, and meals for our editorial advisory board meetings. But her most challenging task was to handle my travel schedule.

To reach the potential contributors and readers in the military, government, and academic communities General Lorenz had identified as our target audience, I had to travel extensively—on average three weeks a month for the three years I served as editor in chief. Betty planned and deconflicted all the travel arrangements—often connecting events in different places on the same trip.

She prepared and shipped copies of SSQ and our promotional materials to the places I was visiting to make sure I had tangible products to use to reach out to our potential partners. When things changed on the fly, as they often did, Betty activated her impressive list of contacts to adapt to the change. Her cheerful attitude and her willingness to take on any task was a key ingredient that helped make SSQ successful and sustainable from the outset.

From the beginning, we planned for SSQ to have print and electronic subscribers. I believed the ability to place a printed copy of the journal in someone’s hand, to have a copy on someone’s desk or in a senior leader’s waiting area, and to give authors the ability to pass along a copy of the journal with their article was a foundational element to starting the new publication. I also believed electronic media would become the normal way for most readers to access content in the future.

In early 2007, libraries had begun to include eBooks in their collections, and by November 2007, Amazon released its first Kindle e-reader.
I wanted *SSQ* to be positioned to reach any reader in any format. We sent out an announcement advertising the launch of *SSQ* with the first issue to appear in September 2007—and offered early sign-up to those who wished to subscribe electronically. The print version of the journal would reach 5,000 readers; by September 2007, we already had about 20,000 electronic subscriptions, and these e-subscribers have remained a key part of *SSQ*’s readership.

Professional journals rely on peer-review processes to ensure quality and credibility. *SSQ* developed an impressive list of peer reviewers starting with the Air University faculty and including professionals within the Defense Department, other branches of the government, and US and foreign faculty members in key disciplines. Our challenge to the peer reviewers was to provide clear, actionable feedback to authors on what they could do to improve their draft articles. In some cases, the peer reviewer recommended an article was a better fit for publication in another journal. We welcomed that kind of feedback because our ultimate approach was to help authors get their ideas published—whether it was with *SSQ* or not.

Contributing editors also helped instill rigor and quality in *SSQ*. I wish I could take credit for this idea, but it came from Dr. John T. LaSaine, a long-serving member of the Air Command and Staff College faculty. The idea was to provide me with a group of counselors from all the academic disciplines at Air University. When articles had gone through the peer review and refinement process, and I deemed them worthy of consideration for publication, the contributing editors, initially all full professors or senior faculty members in Air University graduate degree-granting programs, met to discuss the slate of prospective articles and make a recommendation. These meetings were always fun and interesting and, most importantly, gave me invaluable perspectives on preserving *SSQ*’s quality and credibility.

Our editorial advisory board comprised key senior military, government, and academic leaders who provided advice and perspectives on *SSQ*’s direction. Typically, I reviewed *SSQ*’s charter and operating philosophy with the board before plunging into the details of articles published since the last meeting, the types of articles scheduled for publication during the next year, budget planning and execution, and, most importantly, the ways in which the board could help make *SSQ* more effective. Invariably, the members generated more ideas than I could execute in a single year, but their perspectives as senior leaders helped me keep *SSQ* aimed at the strategic level that General Lorenz had set as the target.

*Strategic Studies Quarterly* appeared for the first time in September 2007, as I had forecast in January of that year. General Lorenz had con-
tributed his perspective, “Ideas—The Essential Elements for Strategic Security in an Uncertain Future.”

We look forward to seeing an active, enthusiastic exchange of ideas that contributes directly to making our nation more secure, and we encourage you to bring ideas into this forum—to offer your views and proposals here, where they will receive the serious consideration they deserve. We expect that some of the answers to the tough problems we will confront in the rapidly changing strategic environment will appear in the pages of this journal.¹

Congressman Terry Everett (R-AL) wrote about the need for a comprehensive space protection strategy. Dr. Edwina Campbell wrote about the collapsing transatlantic consensus; Dr. Jeffrey Record provided a perspective on the Powell Doctrine; Dr. Jim Forsyth wrote the first of several articles that would appear in SSQ over the years on great power competition; and Dr. Phil Meilinger wrote about restoring the influence of Clausewitz for military strategists.

When I placed the first copy in General Lorenz’s hands, he laughed, presented me a coin, and said, “I never thought you would do it!” Then, the more daunting issue hit me: What had taken us seven months to accomplish (it took nearly two months for Air University Press and the printer to edit, typeset, print, and ship each issue) now had to be done in less than two! Nevertheless, SSQ had launched. For the next three years, Betty, Tawanda, and I worked, learned, and grew as professionals in the journal business.

As the Air University team evolves Strategic Studies Quarterly into Æther: A Journal of Strategy and Airpower to serve as a forum for contributors to offer and exchange ideas to accelerate change to meet our increasingly complex security challenges, I believe that SSQ’s continuing legacy will be to encourage its readers to think, to encourage authors to produce their best ideas for public consumption, and, as General Lorenz wrote in that inaugural issue, to find answers to some of our toughest national security problems. I trust that today’s leadership team at the professional journals division will enjoy as much success and as much fun with Æther as the SSQ team did throughout our time creating new and compelling professional publications for our Air Force. SSQ

Anthony C. Cain
Dr. Cain, Colonel USAF, retired, served as the editor for Strategic Studies Quarterly from 2007 to 2010.
Notes

China is rising, and tensions with the United States have increased in recent years. Yet the core theory informing much of US policy toward China today—the idea of a Thucydides Trap popularized in Graham Allison’s *Destined for War*—suffers from three major flaws: unclear definitions, omitted variable bias, and selection bias. Because any one of these problems is fatal to an attempt to use data in order to make predictions or any causal inferences, Allison’s findings on power transitions should not be used as a guide for understanding the US-China relationship.

**Introduction**

The idea the United States and China are potentially headed for war has become commonplace among knowledgeable observers. This possibility is often articulated through the concept of the Thucydides Trap, which says that when one power seeks to displace the other, war is, if not likely, at least a serious possibility. Given China’s economic and military capabilities are rising relative to those of the United States, this concept has become the lens through which many see great power competition. Despite China’s three decades of remarkable economic growth, the more pessimistic view of its rise has been a recent development. According to Google Scholar, between 1989 and 2012, the number of works in which the phrases “China” and “Thucydides Trap” were both mentioned ranged between 0 and 7 each year. In 2013, there were 23 works with both phrases, and by 2019 that number had reached over 500 (fig. 1).

*Destined for War*, published in 2017 and named a notable book of the year by the *New York Times*, *Financial Times*, and *Times of London*, is the most influential work pushing the Thucydides Trap as a way to understand current international politics. The book has garnered praise from the likes of former CIA director David Petraeus, former Secretary of State Henry Kissinger, former Senator Sam Nunn, and former Secretaries of Defense Ash Carter and William Cohen. It also inspired the Harvard Thucydides’s Trap Project, an ongoing effort to expand on and facilitate discussion about Allison’s findings, created by the scholar himself. Cited nearly 800 times as of this writing, perhaps no international relations book of the last decade has had as much impact.
Despite China’s rise and persistent tensions with the United States over human rights and other issues, the analysis in Destined for War suffers from major flaws, some of which have been pointed out in academic responses to the book.¹ Three in particular—unclear definitions, omitted variable bias, and selection bias—make moot any attempt to use data in support of predictions or causal inferences. The findings in Destined for War on power transitions should therefore not be used as a framework for interpreting the US–China relationship.

A Structural Explanation

Those who see a threat coming from Beijing differ over the ultimate source of tensions. In his classic Man, the State, and War, Kenneth Waltz presents three images to explain the causes of war.² Conflict is caused by bad actors, that is, leaders with an unusual propensity toward aggression; by bad states, usually nondemocracies; or by the structure of the international system, the explanation favored by most academic realists. In the case of China, the first image puts the blame on President and Chinese Communist Party Leader Xi Jinping and his consolidation of power.³ An even less compromising school of thought holds to the second image wherein dictatorships are necessarily aggressive even if power is not consolidated in the hands of one individual.⁴

The third image is the lens through which the Thucydides Trap understands great power relations. The first and second images tend to be less

---

**Figure 1. Google Scholar "Thucydides Trap" mentions**

![Graph showing the number of Google Scholar mentions for "Thucydides Trap" over years from 1990 to 2020.](image)

Richard Hanania
credible to those with a thorough understanding of history. Democracies can be belligerent, and leaders are usually constrained by domestic politics, the international system, and their own sense of self-preservation. The idea that modern China is particularly aggressive is popular among commentators but not among academics and those who take a more historically informed perspective.5

Throughout history, democracies have demonstrated hegemonic behavior—the United States and the United Kingdom both established global empires. Dictatorships can also coexist without major conflict, as demonstrated by the Concert of Europe, the name given to the arrangement through which five mostly nondemocratic powers on the continent managed to settle disputes peacefully for most of the nineteenth century. Thus, for those who argue we must treat China as an adversary, the structural explanation, Waltz’s third image, applies.

In 2006, John Mearsheimer predicted, “if China continues its impressive economic growth over the next few decades, the United States and China are likely to engage in an intense security competition with considerable potential for war.”6 He based this assessment not on the internal politics or culture of either of the two powers but on the idea that states cannot trust one another, and great powers will always compete for influence.

Recently, Stephen Walt has taken issue with those who think the United States and China are likely to coexist peacefully if America changes its posture. He writes, “because each is the other’s greatest potential threat, they will inevitably eye each other warily, go to considerable lengths to reduce the other’s ability to threaten their core interests, and constantly look for ways to gain an advantage, if only to ensure that the other side does not gain an advantage over them.”7

Perhaps because Allison has provided quantitative evidence for similar claims, Destined for War has become the most influential work in this genre. A former assistant secretary of defense under President Bill Clinton, Allison published his book while director of the Belfer Center for Science and International Affairs at the Harvard Kennedy School.

Despite his affiliation with a Democratic administration, Allison’s ideas were adopted by many Trump administration officials. For instance, the 2017 National Security Strategy of the United States informed the country that in recent years, “after being dismissed as a phenomenon of an earlier century, great power competition returned.”8 As demonstrated by the extensive media coverage of Allison’s work, perhaps no academic has done more to shape American understandings of the future trajectory of the US-China relationship.
Allison gathered 16 cases over the last 500 years where a rising power challenged an established power, from the rivalry between Portugal and Spain in the late fifteenth century to the rivalry between the UK/France and Germany at the end of the twentieth century. Of these, 12 cases ended up in war. A naïve analysis therefore suggests that if history is any guide, there is around a 75 percent chance the United States and China will go to war in the coming decades.

Although Allison does not conduct any more sophisticated statistical tests, such a record indicates if we accept his methodology and reject any bias in the analysis, using conventional measures of statistical significance, we can be more than 95 percent certain that the chances of a great power war between the United States and China are over 50 percent, a truly horrifying possibility. *Destined for War* does not recommend any particular course of action. Rather, the author presents four possible grand strategies that the United States may adopt. American leaders can accommodate China, seek to undermine it, negotiate a long peace, or redefine the relationship.

**Flawed Methodology**

As mentioned above, Allison’s work and his attempts to draw conclusions about the future course of US-China relations contain three interrelated problems, namely, unclear definitions, omitted variable bias, and selection bias. It is of note that Allison himself acknowledges his work might not withstand statistical scrutiny. His appendix 2 is titled “Seven Straw Men.” The fifth of these states, “the Thucydides’s Trap Case File offers too small a data set to support claims about laws or regularities, or for use by social scientists seeking to do so.” Allison responds, “Agreed. The purpose of this inquiry is to explore a phenomenon—not to propose iron laws or create a data set for statisticians.”

This admission is remarkable. In responding to this “straw man,” Allison creates his own, saying that he is not arguing for “iron laws.” Yet the sophisticated critique is not that 16 cases spread out over 5 centuries in a bivariate analysis does not lead to “iron laws.” Rather, it is that such an analysis provides no guidance to understanding US-China relations, a point Allison seems to agree with. His statement that he is not seeking to “create a data set for statisticians” implies there is one standard that those who engage in quantitative analysis should apply to judging a work and another for everyone else. Presenting numbers on a phenomenon and then saying it cannot meet the standards of statisticians is like presenting an argument about genetics and saying that it cannot be judged by the standards of biology.
Moreover, this seeming humility contradicts not only how others have used Allison's work but how he himself has promoted it. For example, in the *Atlantic*, after summarizing his findings, Allison writes:

> Based on the current trajectory, war between the United States and China in the decades ahead is not just possible, but much more likely than recognized at the moment. Indeed, judging by the historical record, war is more likely than not. Moreover, current underestimations and misapprehensions of the hazards inherent in the U.S.-China relationship contribute greatly to those hazards. A risk associated with Thucydides's Trap is that business as usual—not just an unexpected, extraordinary event—can trigger large-scale conflict.¹⁰

Allison made similar points in a 2017 *Foreign Policy* essay.¹¹ His book recommends the White House establish a Council of Historical Advisers, a group that would be analogous to the Council of Economic Advisers, and look at the past to draw lessons about the present.¹² In April 2017, he went to the White House and briefed a group of National Security Council staffers on the Thucydides Trap.¹³

China hawks at the highest levels of government have seized on the concept to support their preferred policies, including former National Security Advisor H. R. McMaster and former Secretary of Defense James Mattis.¹⁴ This means Allison, despite some equivocations, wants to have it both ways. He makes sweeping conclusions about what his results say for the future of the US-China relationship, while also heading off any rigorous analysis of whether the results support statements like those in the passage quoted above.

In fact, Allison misstates the nature of the problem of drawing conclusions from his cases. The issue is not a small data set per se; a study with 16 observations can be valuable if it is well designed. Assuming there are no problems with data collection, whether one can make a predictive claim about the likelihood of a binary outcome depends on two factors: the ratio of hits to misses across observed cases and the total sample size.¹⁵

As discussed above, if the methodology were sound, 12 cases of armed conflict out of 16 observations would pass conventional tests of significance. This result would provide a high level of confidence that when a rising power challenges an established power, war is more likely than not to result. Unfortunately, the study is not well designed.

**Unclear Definitions**

In the field of psychology, the replication crisis has shown how even well-intentioned analysts can introduce bias into their scholarship when
they have too much flexibility in research design. Consciously or not, there is a human tendency to pick cases and measure variables in ways that support one’s theory. Consequently, psychology has seen the rise of the preregistration revolution in which scholars explain every step of their research project before conducting an experiment or analyzing data. The lessons learned from other areas of social science urge caution when interpreting empirical data in the field of international relations. If studies in which the experimenter has complete control of the environment can be cherry-picked or p-hacked to produce certain results, historical analysis provides many more opportunities for subjectivity unless the researcher is careful.

It is thus worth exploring how Allison defines his cases. On his website explaining the methodology, he writes that he includes each case where “a rising power threatened to displace a major ruling power.” In addition, “these histories use ‘rise’ and ‘rule’ as conventionally defined, along with synonyms emphasizing rapid shifts in relative economic and military strength.” Nearly every substantive word in these sentences is ill-defined. We are not told what the “conventional” definitions of “rise” and “rule” are. The term “rapid shift” in the context of geopolitics can mean anything from one or two years to several decades.

Moreover, how exactly are economic and military strength measured, and how large does the shift have to be? Is economic strength measured by GDP, or does the calculation also consider the production of militarily important sectors such as steel? In other words, is military strength actual or potential? Scholars have compiled empirical measures of these things, but Allison provides no details about which measures he used, if any. We have no way of determining whether a 20 percent reduction in the GDP gap between two powers over 10 years would count as one of his cases or whether the same reduction over 20 or even 50 years would. Additionally, what does “threaten to displace” mean? Does it account for the intentions of each side, and if so, how are those measured? The selection process seems to be completely anecdotal.

Throughout his data set, it is unclear why Allison includes certain cases but omits others. Thus we are told Germany displacing Great Britain and France in Europe since the 1990s is a power transition that did not lead to war. Why not also consider the end of colonialism when Great Britain and France gave up positions in Africa and Asia, and the United States took their place?

For example, after their defeat at the hands of the Viet Cong, the French left Indochina, and the United States assumed many of the responsibili-
ties of the former colonizer. One may respond that the voluntary relinquishment of power does not count. Yet it is hard to see why Germany taking control over European affairs should count, given that all the powers involved were treaty allies within NATO and the European Union.

**Omitted Variable Bias**

Omitted variable bias has been called perhaps “the most serious and pervasive threat to the validity of social science research.”\(^{19}\) If an analyst finds one variable predicts another, the question becomes whether this outcome is because \(x\) itself causes \(y\) or because one or more unobserved variables that correlate with \(x\) actually cause \(y\). If those unobserved variables are no longer present, we may find the relationship between \(x\) and \(y\) disappears.

As an example, throughout most of history, economic downturns have been associated with an increase in the death rate. In twentieth-century America, however, this relationship did not hold. Health indicators improved during the Great Depression, while mortality increased during the economic boom years of the 1920s.\(^{20}\) The reason for this incongruency (compared with longer-term historical data) is that because Americans live in an industrialized country, they are so wealthy that even a large decrease in economic output does not necessarily lead to more deaths. A relationship that held throughout human history disappeared or even reversed when circumstances changed.

Allison conducts a bivariate analysis in which one independent variable predicts a dependent variable. Yet if we control for other variables that could determine whether rival powers end up in war, the results look much less impressive. Perhaps the most important omitted variable Allison does not consider is time. It may seem too obvious to point out, but the world has changed quite a bit in the last 500 years. Do international relations in the sixteenth century have anything to say about the twenty-first century, given the social, political, and technological changes that have occurred?

For international relations theorists, nuclear weapons have fundamentally changed world politics and made war between great power unthinkable.\(^{21}\) Still, this factor is only one possible explanation as to why the past is not a good guide for the present. Crediting nuclear weapons for the decline in great power war requires setting aside massive societal and technological changes including secularization, the explosion of wealth since the Industrial Revolution, the increasing political power of women, the emergence of mass media, and the improved ability of world leaders to communicate quickly with one another. Also, satellite imagery and other
breakthroughs in information and intelligence gathering lower uncertainty regarding power disparities.\textsuperscript{22}

The Thucydides’s Trap Project website indicates 14 more cases are being considered for inclusion in the data set. Of these cases, only seven ended up in war.\textsuperscript{23} Had these cases been included in the original book, the results would have looked much less impressive, with only 19 of 30 power transitions resulting in armed conflict. Moreover, if one simply stuck to cases from the second half of the twentieth century and later, then only one out of seven cases led to war—indicating the past might not be such a valuable guide to understanding modern international relations.

One may also add the US-China relationship, which thus far has not resulted in war, giving us only one out of eight recent power transitions that led to armed conflict. None of this is to say that Allison should have included the additional cases. Without clearer definitions of what is being measured, the lesson is that one cannot determine which cases should be included.

\textbf{Selection Bias}

There are two ways to understand the Thucydides Trap. In the first—the more ambitious version of the theory that Allison presents—China and the United States may fight a war because of changes in relative economic strength. In the second—the one Allison uses when he is more careful—the two countries may fight a war because of how they perceive their interests combined with shifts in objective measures of power. Yet one cannot conclude states are “destined for war” based on a material shift in power from a data set that selects for countries based on whether they are rivals, which is a state of affairs determined by the intentions of the actors in question.

In technical terms, selection bias occurs when one attempts to make conclusions based on a nonrandom sampling of the data.\textsuperscript{24} If a researcher wants to understand public opinion on a specific issue, it would be a mistake to rely on a demographically unbalanced sample or a survey in which respondents seek out participation.

Similarly, Allison’s claims about whether nations are destined to fight cannot rely on choosing cases where states have subjectively perceived overlapping interests. Drawing such a conclusion is like finding the most aggressive individuals in a bar and measuring how often they end up fighting to draw conclusions about the likelihood of a conflict between any two random individuals with high levels of upper body strength.
There are two ways to avoid conflict between the United States and China. In one scenario, China challenges the United States in East Asia, and we somehow avoid war—the optimistic outcome of Allison’s Thucydides Trap. In the other scenario, there is no challenge in the first place. Imagine if the economic rise of China mirrored that of Japan in the second half of the twentieth century, and Beijing did not become militarily more assertive. By Allison’s definition, there would be no Thucydides Trap because neither side is challenging the other. Moreover, the same would be true if the United States decided it was no longer interested in maintaining its military position in East Asia.

Allison collects cases based on one side threatening another; consequently, his analysis is biased by the fact he selects countries that are antagonists and then checks how often they find themselves at war. Advocates of restraint want to push the United States toward taking a less militarized approach to foreign policy. If they succeed, there is no longer a Thucydides Trap, just as there was none when the United States replaced France as the dominant power in Southeast Asia.

Great power antagonism is in that sense not only a choice; it is the crux of the discussion focused on how the United States should meet the rise of China. The concept of the Thucydides Trap skips the entire debate and assumes the United States wishes to remain an established power in East Asia and is being challenged by China. If these propositions are true, they are due to choices made by both sides, not the result of circumstances outside of human control.

**Conclusion: Misusing History**

*Destined for War* has helped transform how foreign policy elites and much of the educated public think about the US-China relationship. While Allison studiously avoids making strong recommendations, many in the press have done so on his behalf, and some conclusions seem to follow naturally from the underlying analysis. After all, if one believes there is a 75 percent chance the United States and China will end up in war, does it not make sense to increase military spending just in case?

Thus, while Allison explicitly rejects the idea that accepting the Thucydides Trap means the United States should adopt an aggressive posture toward China and presents accommodation as one possible strategy, practically all prominent analysts who have accepted his framing have advocated for more confrontational policies. The idea of the Thucydides Trap is appealing to hawks because it skips the question of whether the
United States *should* be a rival to China and assumes the laws of history have decided that it already is, or must be in the near future.

The only remaining question is whether American leaders can manage to make good decisions that would lead to the rare situation in which a power transition does not result in war. While it is possible to argue the Thucydides Trap calls for accommodation, presenting the US-China relationship as naturally antagonistic has provided rhetorical and political ammunition for advocates of more confrontational policies.

Luckily, things are not so dire. Allison’s analysis does not follow the most basic rules of statistical modeling. There is practically no attempt to clearly define how he chooses his cases. Nor is there any attempt to account for omitted variables, even in the simple form of dividing the data by historical era. Finally, Allison selects cases where countries have conflicting interests as subjectively perceived by leaders. Thus, he short-circuits the continuing debate about how the United States should respond to a rising China by ignoring a potential path to peace wherein American leaders move away from a confrontational posture in East Asia.

Allison asks the reader not to judge his work from the perspective of statistics. Unfortunately, it is incoherent to present data to the world, argue that it should shape our predictions about how the US-China relationship will unfold, and then ask that we do not judge the theory by the most rigorous standards. Either the data set Allison presents should guide US thinking and behavior, or it should not.

Whether it is possible to use history to derive statistical predictions about the likelihood of war is an open question. What is certain is that doing so must, at the very least, avoid the problems highlighted here. A rigorous historical analysis intended to frame the US-China relationship and support the foreign policy process must clarify the standards of inclusion; consider other variables that might influence the likelihood of war; and avoid endogeneity problems that conflate the dependent variable and the independent variable of interest.

A narrower historical focus on international politics since the second half of the twentieth century provides a more optimistic lens for understanding the future of great power relations. As measured by GDP, a handful of power shifts have occurred over the last several decades. Among these are China relative to Japan and Russia, and Germany relative to other European nations. None of these cases has led to war. The disappearance of interstate conflict more generally offers hope that even if power transitions may have created a substantial risk of war in the past, they do not do so today.
In the end, however, history may be of limited utility in understanding the US-China relationship. Instead of employing historical analogies that may or may not apply or using data sets that cannot meet basic standards for establishing causal inference or reasonably predict behavior, American foreign policy should proceed by considering the interests, politics, and material capabilities of both sides. Questions such as what does China want, can the United States live with its claims, and what is worth going to war over should be at the forefront of the minds of American leaders.

The Thucydides Trap begins by assuming the two superpowers are engaged in a rivalry, all but foreclosing a more restrained American foreign policy by presenting such a view as hopelessly naïve. When it comes to power transitions, it is not enough simply to say that studying previous centuries reveals no iron laws. Rather, scholars have yet to show that conclusions about the likely course of future events derived from the distant past can withstand basic scrutiny.

Richard Hanania
Dr. Hanania is the president of the Center for the Study of Partisanship and Ideology and a research fellow at Defense Priorities.

Notes


14. Crowley, “Greek History.”


More Is Not Always Better: Oversight of the Military

Marie T. Harnly

An analysis of three case studies reveals the Air Force does not necessarily gain autonomy when government principals are divided over policy, contradicting current scholarship on the issue. Varying levels of service autonomy under divided principals requires tailored approaches to policy development and implementation.

The two government branches that delegate national security to the military, the executive and legislative, sometimes differ on policy preferences. Scholarly literature claims that autonomy increases for the military when these principals are divided on policy because the military can play one branch off the other, gaining latitude for its policy preference. An examination of three cases specific to the Air Force finds (1) the service does not reliably receive more autonomy from divided executive and legislative branches; (2) conditions other than those classically understood by civil-military relations theory contribute to variations in Air Force autonomy; and (3) the Air Force does not always desire more autonomy. These contributions offer practical insights for military advisers and policy makers.

Introduction

Three highly charged issues—the proposed retirement of the A-10 Thunderbolt II, the repeal of the combat flying ban for women, and the creation of the United States Space Force (USSF)—highlight the principal-agent dynamic between the US government and the Air Force in which the military provides national security expertise to the government. The preponderance of scholarly literature on civil-military relations claims that divided principals—instances in policy promulgation when preferences of the executive and legislative branches diverge—permit a less responsive military agent, resulting in more autonomy for the agent to act as it sees fit.¹ Scholars hypothesize that in these cases the military agent plays the principals off one another, providing the agent greater latitude to implement its preferred option.
Principal Agent Theory in Civil-Military Relations

Principal agent theory entails one party, the principal, delegating work to another party, the agent, to perform. Scholars first tailored principal agent theory to political science in 1975, determining the theory provided a new perspective of government and its policies. Almost three decades later, Peter Feaver applied the principal agent theory to American civil-military relations to better understand this relationship.

In Feaver’s formulation, the government principal delegates national security functions to the military agent due to expertise. In turn, the military agent presents its recommended policies and preferences to the government principal; together, the civilian principal and military agent work together to harmonize their respective preferences into national security policy.

Certainly, civil-military relations generate unique agency problems. The military agent traditionally prizes autonomy and prefers less intrusive monitoring mechanisms. Liberal rewards and minimal punishments provide the military agent with the autonomy to determine what tasks to complete and how to complete them. Feaver describes how closely military agents satisfy the government principal’s intent by using the terms working and shirking.

A military agent is working when the agent accomplishes tasks according to the government principal’s criteria. The government principal rewards working agents with greater levels of autonomy. A shirking agent, on the other hand, accomplishes tasks according to the military agent’s preferences instead of the government principal’s preferences. The government principal punishes shirking agents by reducing levels of autonomy.

While Feaver’s model simplifies the government actor as a unified principal, Deborah Avant, modifying Feaver’s model, establishes the executive branch and legislative branch as dual government principals for the military agent. Under the Constitution, the executive and legislative branches share civilian oversight of the military but possess different authorities. The executive branch develops military policy, and the president serves as commander in chief. Congress, as the legislative branch, balances these executive branch authorities by authorizing and appropriating military funding and retaining the authority to declare war.

Divided Principals and Autonomy

Avant elaborates on her discussion of two government principals by categorizing these principals as unified or divided on issues. Unified prin-
principals agree on how and what tasks the agent should perform, how to monitor the agent, and the incentive or consequence structure. Divided principals, by contrast, disagree on these items. In American civil-military relations, this disunity between the executive and legislative branches causes divided principals. Agent preferences may be most influential when principals disagree since the agent could potentially use the preference difference to gain support for its option. Most scholarship finds, therefore, the military agent tends to gain more autonomy in situations characterized by divided government principals.

In addition to discussing different authorities for divided principals, a distinction needs to be made for the two phases of agent autonomy—the advisory phase, prior to a policy decision, and implementation phase, after a decision has been made. In his modification of Feaver's principal-agent model, Jeffrey Donnithorne focuses on the different civil-military dynamics on either side of a policy decision. During the advisory phase, the military agent recommends a course of action based upon best military judgement. Once the principal makes a decision, the implementation phase begins, in which the military agent must carry out the decision. The advisory phase informs the implementation phase and the amount of flexibility an agent anticipates receiving to enact the policy in its preferred manner. Four attributes of the policy itself—specificity, imminence, durability, and enforceability—determine the degree of anticipated agent autonomy during the implementation phase.

The specificity of the policy language narrows the agent’s freedom of action as opposed to vague language that can be more broadly interpreted. The policy time frame—imminence—also impacts the military agent’s ability to maneuver. The durability of the policy determines whether the prescription is fleeting or enduring. Finally, the inherent enforceability of a policy portends how closely an agent will have to comply with the direction. The autonomy the military agent receives can thus vary from the advisory phase to implementation phase.

**Background**

The following research determined under what circumstances divided government principals led to more or less Air Force autonomy. The three cases occurred after the Goldwater-Nichols Act of 1986 and reflect the same Department of Defense (DOD) structure that exists today. The issues these cases revolved around occurred at the Air Force level and required policy decisions from the government principals. The analysis examined agent autonomy according to the two phases—the advisory phase and im-
plementation phase. In the cases that decreased agent autonomy, a deviation from the outcome most easily explained by scholarly literature exists.

In the advisory phase, the closer the chosen policy was to the Air Force’s preferred policy, the more autonomy the Air Force had. An enactment of the Air Force’s position increased agent autonomy and demonstrated support for the prevailing hypothesis. The government enacting an opposing position constrained agent autonomy.

In the implementation phase, four policy attributes—specificity, imminence, durability, and enforceability—helped determine the Air Force’s anticipated flexibility in implementation. A specific, immediate, binding, and enforceable policy decreased the service’s implementation autonomy. A vague, delayed, short-lived, and unenforceable policy, on the other hand, gave the Air Force wider latitude in implementation. A principal providing flexibility to the service to execute a policy in ways the Air Force sees fit provided autonomy, aligning with the preponderance of the literature.

Case 1: Proposed Retirement of the A-10

The proposed retirement of the A-10 fleet must be understood within the larger context of budget challenges the US military faced in 2013. Due to Congress’s inability to reduce the federal budget by $1.2 trillion that year, in March the Obama administration sequestered “budgetary resources across nonexempt federal government accounts” requiring the Department of Defense take “a 7.8 percent reduction in nonexempt discretionary funding.” That year, this amounted to reductions of approximately $74.4 million in discretionary appropriations and direct spending for the military.

Given these dynamics, Air Force leaders needed to find additional cost savings with minimal impact to combat capability. Then-Air Force Chief of Staff General Mark A. Welsh III scrutinized the Air Force’s five missions—air and space superiority; intelligence, surveillance, and reconnaissance; rapid global mobility; global strike; and command and control—to determine where spending decreases were feasible with the least impact to operations. He concluded none of the Air Force’s cost-cutting options were ideal, but retiring the A-10 had the least operational impact.

From his perspective, the A-10 was built for a specific threat environment, performing a single-mission role as an exclusive air-to-ground platform. The belief persisted among Air Force leaders that the service needed to eliminate entire fleets of aircraft to reach the congressionally mandated budget cut levels. The Air Force estimated a cost savings of $4.2 billion through fiscal year 2019 by divesting itself of the A-10 fleet,
which became its policy recommendation to both its executive branch and legislative branch principals.\textsuperscript{15}

Welsh advocated for this tough decision through reports to Congress and during a hearing to the House Armed Services Committee on September 18, 2013.\textsuperscript{16} By retiring the aging A-10 aircraft, the Air Force intended to modernize its fleet with multirole aircraft that excelled at multiple missions. Reinvesting savings from the A-10 into the F-35 would provide the Air Force with combat capability for a conflict against more advanced adversaries, such as China. Air Force leaders, including Welsh and then-acting Secretary of the Air Force Eric Fanning, consistently advocated for this option, which would allow the service to balance its budget with the least impact to operations overseas.

The proposed A-10 divestiture fell directly within Congress’s purview to approve and appropriate funds for military activities. In response to the Air Force’s proposal, 33 members of Congress from the House and Senate Armed Services committees and the House and Senate appropriations committees drafted and sent correspondence to the secretary of defense and chairman of the Joint Chiefs of Staff arguing against this proposal.\textsuperscript{17} These senators and representatives, most of whom had Air Force bases in their states and districts, asserted the retirement would create a capability gap and endanger service members in future conflicts. The House and Senate subsequently approved an amendment to the fiscal year 2014 National Defense Authorization Act (NDAA) prohibiting the Air Force from retiring the A-10 until the planned replacement was fully operational and flying combat operations.\textsuperscript{18} This outcome contravened the position advocated by the Air Force.

The executive branch held constitutional authority to generate policies regarding the A-10 and supported the Air Force’s position to retire this fleet. President Barack Obama repeatedly announced the executive branch’s preference through statements of administrative policy in May 2014 and June 2015.\textsuperscript{19} These statements strongly objected to the congressional provisions restricting retirement and storage of the A-10.

Throughout the three-year duration of this case, the most vocal members of Congress, those with A-10 bases in their states, ensured congressional authorization and appropriation language, prohibiting the retirement of the A-10, was included in the NDAA. That language, which began in fiscal year 2014, remains in the legislation today.\textsuperscript{20} In the case of the A-10 fleet, the Air Force as the military agent was prohibited from determining which aircraft could best execute its missions and was prevented from autonomy in assigning budget priorities; instead, specific le-
gal restrictions, imposed after the fact by one principal—the legislative branch—constrained the task of reducing its budget.

The prevailing theory suggests the Air Force should have received more autonomy as a function of the divided policy preferences of its principals. Instead, the findings indicate the Air Force did not, in fact, receive more autonomy to choose or implement its preferences.

**Advisory Phase**

In the A-10 case, the Air Force agent and the president possessed an opposing position to that of Congress. The legislative branch, under the constitutional authority of authorizations and appropriations, took action to block the A-10 retirement proposal. Consequently the Air Force received less autonomy, which appears to deviate from the prevailing literature’s hypothesis.

**Implementation Phase**

In order to maintain the status quo of the A-10, Congress limited the Air Force’s ability to retire its fleet. The policy did not allow for flexibility in implementation because it employed all four attributes—specificity, imminence, durability, and enforceability—that restrict trade space for the agent to negotiate how to complete tasks.

First, the language in the fiscal year 2014 NDAA specifically prohibited certain actions associated with the retirement of the A-10, such as aircraft storage and personnel reductions. Second, the law went into effect immediately after a majority in Congress passed it. Third, A-10 fleet restrictions have been written into national security legislation by Congress for eight years running, signifying its durability. Finally, because Congress has responsibility for defense funding, these provisions are inherently enforceable. These four attributes resulted in limited implementation flexibility and less autonomy for the Air Force, in contrast with the preponderance of scholarly literature.

**Outcomes**

Two key variables appear to exert significant influence in this case: the outsized effect of geographic constituent interests for members of Congress and the power of specified authorities given to a single principal. Geographic interest—highly influential in Congress—caused preferences to diverge, an unsurprising and nearly universal finding in the literature. Throughout the duration of this case, the chairman or ranking member of
the relevant Senate committees hailed from states with A-10 units. These committee leaders used their influence to enact legislation blocking the retirement of the A-10.

Congress also maintained sole authority for deciding on the policy preventing retirement of the A-10 and managing its implementation, since both activities revolved around funding. This dynamic effectively removed the theorized maneuver room for an agent in the case of a divided principal. Because Congress held unilateral authority to establish an A-10 policy, a bill prohibiting numerous A-10 activities became law. This legislation ultimately constrained and continues to constrain both the executive branch and the Air Force.

**Case 2: Creation of the Space Force**

A July 2016 Government Accountability Office report highlighted vulnerabilities within the Department of Defense hindering its ability to secure space. The report found lower promotion rates for space professionals, indicating the Department valued space professionals less than other service career fields. Services prioritized funding for space requirements below aircraft requirements in the Air Force, ship requirements in the Navy, tanks in the Army, and amphibious vehicles in the Marine Corps. Ultimately, the report refocused the executive and legislative branches on national space security within the Department.

Representative Mike Rogers, chairman of the House Committee on Armed Services Subcommittee on Strategic Forces, took up the national space security mantle and became its champion. In a 2017 Space Symposium address, Rogers outlined current problems associated with the Department of Defense’s fragmented space organization, disjointed decision making, underprioritized funding requirements, and absence of adequate professional development for those in space career fields. Rogers also introduced legislation calling for the creation of a Space Corps as a new military service responsible for national security programs pertaining to space.

While a vocal congressional minority preferred to establish a separate military branch to focus on the space domain, the Space Corps proposal met resistance from a majority in Congress for two years due to concerns over an expanding defense bureaucracy and budget. The fiscal years 2018 and 2019 NDAAAs did not require the Pentagon to create a space-centric military service, which reflected the preferred policy of the legislative branch—to bolster space functions within the Air Force. Aligning with the majority of Congress, the Air Force preferred to maintain space operations within its service responsibility.
On March 13, 2018, President Donald J. Trump announced his proposal to create a separate military branch, the United States Space Force. Although the executive branch does not possess the authority to establish a separate military branch, the president holds the authority to create combatant commands as the commander in chief of the armed forces. Accordingly, on December 18, 2018, Trump established United States Space Command as a functional unified combatant command. In doing so, he continued the Space Force discussion, building momentum for a new space-focused military branch.

In a move designed in part to overcome Pentagon resistance, Trump signed a Space Policy Directive in February 2019, dictating the Department of Defense develop a plan for Congress establishing United States Space Force as a branch of the United States Armed Forces. Despite widespread opposition from the Air Force, the directive forced the service to craft a plan for a force that would organize, train, and equip military forces to operate in the space domain, similar to the air, land, and sea domain responsibilities of the Air Force, Army, and Navy, respectively.

Throughout 2018 and 2019, Chief of Staff of the Air Force General David L. Goldfein and Secretary of the Air Force Heather A. Wilson continued to advocate for improving space activities while maintaining space functions under the Air Force. In 2019, the Department delivered its proposal for a separate, space-focused military branch according to the presidential directive. The fiscal year 2020 NDAA included language that created a Space Force, and the new military service became law on December 17, 2019.

In seven short pages, Congress outlined the provisions of the United States Space Force including its leadership structure and its position within the Department of the Air Force, mirroring that of the Marine Corps within the Department of the Navy. The legislation prohibited additional authorizations for military personnel and budget increases beyond those outlined in the bill but ultimately created a separate military branch focused on space. The Air Force, as the military agent, did not determine what space functions to perform and how to accomplish space operations better within the Air Force; instead, a new service took over many of these tasks.

The outcome most easily explained by literature indicates the Air Force should have received more autonomy from divided principals during policy decision making and implementation. Instead, the findings reveal the divided principals became united when crafting policy, restricting au-
tonomy for the Air Force. The Air Force did, however, receive more autonomy to execute its preferences.

**Advisory Phase**

Trump and Congress initially opposed one another. Trump strongly advocated for creating the Space Force. Congress’s preliminary preference against this new service was based mainly on fears of budget and bureaucratic expansion. The executive branch leveraged authorities within its scope—the creation of a unified combatant command and a presidential policy directive—to overcome this opposition. Thus divided principals became unified.

**Implementation Phase**

Provisions in the fiscal years 2020 and 2021 NDAAs provided wide latitude for the Air Force to enact the Space Force because they lacked three of the four attributes—specificity, imminence, and enforceability—that restrict autonomy in implementation. First, the legislation did not specify how and what tasks the Air Force agent needed to accomplish to create the new service. Second, while the 2020 bill immediately established a new military branch for space, the transfer of personnel was not a requirement, and numerous deadlines associated with the Space Force extend years into the future.

Third, the laws’ provisions, aside from the existence of a new space-focused military service, were not inherently enforceable due to vague wording or absence of guidance. The Air Force, therefore, gained autonomy to establish the US Space Force in the manner it preferred. Although Congress implemented a course of action that opposed the Air Force’s option, the legislation, as enacted, gave the service (military agent) greater flexibility to build the Space Force in a way it deemed best.

**Outcomes**

The president, the primary sole national security decision maker, garnered support for the US Space Force. He leveraged his authority to institute changes to national space security and overcame congressional resistance. Throughout the duration of this case, the idea of the Space Force gained traction with the president strongly advocating for this new service. More and more supporters, including members of Congress, joined the president’s camp. In the end, the Space Force became law under unified
principals. This policy decision ultimately decreased the Air Force’s autonomy since Congress enacted an opposing position.

Once Congress passed legislation creating the US Space Force, the Air Force—like any agent asked to execute an undesirable policy—preferred an ambiguous, delayed, short-lived, and weakly enforceable one. Congress focused on specifying personnel levels and budgets but gave the Air Force autonomy in all other areas of Space Force implementation. Because Congress possessed a weak preference for the Space Force and maintained sole authority for deciding on the policy, the language in the NDAA was tailored to the authorities Congress retained during the stand-up of the new service.

The remaining details revolving around the establishment of the Space Force overseen by the administration were absent from the law or vague. Policy implementation arrived at a solution satisfactory to the executive branch’s strong preferences and the legislative branch’s concerns—create a Space Force with no manpower or budget increases. The Air Force, as a result, increased its autonomy during execution and was able to dictate how to stand up the new service.

**Case 3: Repeal of Combat Flying Ban for Women**

The repeal of the combat flying ban for women was at the core of a much larger discussion involving the role of women in the military. Throughout the military conflicts of the late 1980s and early 1990s, women in uniform helped achieve the nation’s military objectives. Eight hundred female service members participated in the invasion of Panama during Operation Just Cause. During these operations, female service members engaged in hostile firefights, led forces in battle, commanded assaults on opposing force strongholds, and earned air medals for combat-related missions.

Similarly, 41,000 women deployed to Iraq in 1990 and 1991, which constituted 7 percent of all military personnel involved in the Persian Gulf War. During this conflict, 16 women died, and 2 women became prisoners of war. The notion that Americans would not tolerate women being killed in action or becoming prisoners of war was proven inaccurate by operations in Panama and Iraq and the attendant media. These military conflicts occurring so close together and involving women in hostile-fire situations precipitated the formal discussion about combat roles for women in the military.

The House Armed Services Committee fired the first challenge to the law barring women from combat. In May 1991, Representative Patricia Schroeder, the first woman to serve on the committee, introduced an
amendment to the NDAA repealing the prohibition barring women from flying Air Force combat aircraft. Additionally, Representative Beverly Byron, the first woman to fly on board an Air Force SR-71 aircraft, proposed a similar amendment to repeal the Navy and Marine Corps combat flying bans. These legislative proposals were met with enthusiasm in Congress, which incorporated them into the draft NDAA for fiscal years 1992 and 1993.

The four service chiefs testified in front of the Senate Armed Services Committee Subcommittee on Manpower and Personnel about these proposed amendments. Three of the four service chiefs opposed making women eligible for combat, while one, General Merrill McPeak, then-chief of staff of the Air Force, advocated for allowing equal opportunity to battlefield assignments.

McPeak proved to be an outlier with a majority of the senior military leaders desiring the combat ban remain in place. But with Congress possessing sole authority over legislation, the act repealing the combat aviation exclusion for females became law on December 5, 1991. Nevertheless, while the NDAA allowed the assignment of females to operational units with fighter aircraft, bomber aircraft, or helicopters, it did not mandate such assignments.

President George H. W. Bush, like most of the service chiefs, was not keen on the idea of allowing women to serve in combat flying roles. Realizing the legislative branch wielded the power to repeal the law, Bush avoided making public statements endorsing one stance or another. He could instead leverage his commander-in-chief policy authority during implementation. As a result, Bush asked Congress to establish a presidential commission to study the issue of women in combat roles. The commission recommended women continue to be excluded from air and ground combat.

Due to the commission’s findings and as a matter of policy, the Department of Defense refused to assign females to combat units once the repeal took effect and continued prohibiting women from combat assignments. The Air Force began sending females to pilot training to fly fighter and bomber aircraft at the beginning of 1992, but these women were restricted to teaching at pilot training after they completed training because of this DOD policy.

Since the enacted NDAA covered two years—1992 and 1993—Congress did not have at its disposal a legislative mechanism to force the issue until the fall of 1993. President Bill Clinton took office in January 1993 and decided to arbitrate the different perspectives between the legislative
Marie T. Harnly

branch’s repeal and the recommendation of the presidential commission. Clinton, in his commander-in-chief role, ordered the military branches to open combat aviation to women.

The preponderance of literature predicts the Air Force should have gained more autonomy under divided principals. While the Air Force did in fact receive more autonomy to craft policy and execute its preferences, this case instead represents an instance wherein the Air Force preferred less autonomy in the implementation phase to lock in its policy decision.

**Advisory Phase**

Each principal—the executive branch and legislative branch—used the governing tools at its disposal to pursue its preference. Congress, with the ability to repeal laws, proposed eliminating the combat exclusion for women. The Air Force agent was the only military branch to support the repeal of the ban, which aligned with Congress’s legislative amendments. Bush, on the other hand, preferred to retain the exclusion and leveraged his presidential authority to establish a commission to study the roles of women in combat. Congress, maintaining sole authority for legislation, ultimately repealed the combat flying ban, allowing women to fly fighter and bomber aircraft.

**Implementation Phase**

The NDAA provisions provided wide latitude for policy implementation in the absence of three of four attributes—specificity, imminence, and enforceability. First, the vague language did not specify the military had to assign females to combat flying roles. This flexibility allowed the military services to train female fighter and bomber pilots but did not require them to be assigned to combat units. As a result of this ambiguous language, the president was able to prohibit females from combat assignments.

Second, although the repeal of the combat flying ban for women went into effect immediately, under the law, the presidential commission had a year to study this issue, providing time to renegotiate these terms. Third, aside from the repeal of the combat flying ban for women and the commission details requested by the president, the law’s terms were difficult to enforce due to their absence of guidance.

Yet, instead of receiving expected greater autonomy from the actions of the legislative principal, the Air Force yielded to the executive branch’s decision authority that resulted in a continued ban of females from combat assignments. Unlike Congress, the president did not give the Air Force
room to maneuver with its prohibition of assignments to combat units. The executive branch ultimately constrained the Air Force’s autonomy to execute the repeal of the combat flying ban.

Outcomes

Partisan differences had a strong influence on the outcome because of social issues, in particular conservative versus progressive visions for the military. Congress’s sole authority to decide on a policy also played a role in the outcome. The legislative branch possessed unilateral authority in this case, and the Air Force’s position aligned with Congress’s preference. But in possessing this sole authority, Congress also demonstrated restraint by enacting a law allowing women to fly combat aircraft but not requiring the service to do so.

These dynamics highlight that the Air Force may have desired less autonomy to implement the combat flying ban for women. Since Congress’s repeal of the ban aligned with McPeak’s best military advice, the Air Force would have preferred an unwavering policy requiring female assignment to combat units. While literature assumes the Air Force desires more autonomy, there can be instances where the agent prefers its autonomy to be constrained.

Conclusion and Implications

This research offers three new insights on the dynamics of divided principals and agent autonomy in the field of American civil-military relations. First, the preponderance of principal-agent literature claims that divided principals create a less responsive agent, resulting in more agent autonomy.\(^{39}\) The proposed retirement of the A-10 and the creation of the Space Force, however, yielded a different outcome.

Distinguishing between advisory and implementation dynamics reveals that although these cases had divided principals, Congress enacted a policy opposed to the military agent’s preference during the advisory phases, thus constraining the Air Force’s autonomy during this phase. Similarly, the Air Force’s autonomy decreased during the implementation phase of the proposed retirement of the A-10. The reason for these differences lies in the different authorities of the principals that tend to reside on either side of the policy decision—an insight that leads to the second contribution of this research.

Second, Avant’s baseline model assumes when there are two principals, they share authority over the agent.\(^{40}\) This notion is key to the agent re-
ceiving more autonomy from divided principals because the agent could theoretically play the principals off of one another to gain latitude for its preferred option.

Since the Constitution outlines different authorities for the government principals, in most cases only one principal holds the action authority over the military agent, even though a divided principal situation exists in terms of policy preferences. The A-10 case demonstrates this dynamic: Congress maintained sole authority for both policy decision making and execution. Consequently, Congress was able to prohibit the retirement of the A-10 with legislation and funding. In sole-authority situations, the government principals may use the tools at their disposal to codify their preference in policy or law.

The Space Force case also presents a departure from the literature. Although the legislative branch possessed sole authority to create a new military branch, the president established US Space Command and directed the Air Force to submit a plan to Congress outlining how it would establish the Space Force. As a result, divided-principal scenarios in American civil-military relations once again do not reflect the classic dynamic. Instead, the principals have complementary authorities, where one or both branches of government possess decision-making authority or authority over policy execution.

Third, traditional principal agent theory as expressed in the civil-military model assumes agents always prefer more autonomy. As the cases in this study suggest, however, this assumption does not always hold true in American civil-military relations. Instead, when a principal enacts the military agent’s preferred policy option, that agent may want less implementation autonomy for itself and all others in the policy space. The repeal of the combat flying ban for women illustrates this possibility.

Congress possessed unilateral authority for policy decisions and repealed combat exclusions, which aligned with the Air Force’s preferred policy option, but the language of the enacted legislation did not bind the executive branch to a specific pathway of implementation, offering instead wide flexibility for execution. If Congress constrained agent autonomy in the implementation phase, the Air Force’s option would have been cemented in legislation, eliminating the flexibility to erect such barriers. Therefore, agents do not always desire more autonomy in the implementation phase; there are situations where the military can prefer less autonomy to solidify its preference for the foreseeable future.

The three new contributions impact policy makers and military advisers in the practical sense. Divided principals do not always lead to more au-
tonomy for the military. Understanding this allows those in the legislative and executive branches and those in the military to develop various approaches for issues. In a divided-principals situation, levels of autonomy vary on a scale from constraining autonomy to producing autonomy. Policy makers need to develop policy that accounts for this spectrum. Similarly, military advisers need to provide guidance to military leaders that considers this spectrum.

Additionally, military advisers and policy makers need to account for the military agent’s preferred level of autonomy in specific circumstances. Furthermore, even with divided principals, in some cases government principals have complementary authorities and in others principals retain sole authority. The situation itself and authority contexts should drive guidance that military advisers provide military leaders and policy recommendations that policy makers provide the government. Thus, the real-world applications of the three contributions from this research infuse nuance into both policy and advice for military and government leaders.

Marie T. Harnly

Major Harnly, USAF, commander of the 39th Civil Engineering Squadron, Incirlik Air Base, Turkey, holds a master of science in engineering management from the Air Force Institute of Technology, a master in public administration from Harvard University, and a master of philosophy in military strategy from the School of Advanced Air and Space Studies.

Notes


38. Grant, “Quiet Pioneers,” 35.


In today’s threat environment, adversaries can hold the continent hostage unless leaders can bolster its deterrence posture. Rather than deterrence by punishment, however, the focus of NORAD, USNORTHCOM, and the Canadian Joint Operations Command must be on deterrence by denial and increasing the costs of actions by adversaries should they pursue an attack on North America.

To ensure credible deterrence by denial, the North American Aerospace Defense Command (NORAD) and the Canada–US (CA-NUS) defense relationship needs modernizing. Not only do sensors need to be updated and significant expenditures made, but the entire approach to the defense of North America needs to materially change. We must rethink the domains that require defending and how deterrence by denial moves beyond the current outdated Cold War mindset that evolved in an ad hoc manner.

Beginning with General Charles H. Jacoby Jr., USA—dual-hatted as commander of United States Northern Command (USNORTHCOM) and NORAD from 2011 to 2014—and the 2013 NORAD Next study, successive dual-hatted commanders have raised concerns about the vulnerability of North America. A new generation of advanced strike weapons, two peer US competitors, and violent extremists seek to exploit all domains to undermine the credibility of US and allies’ defenses.

Deterrence is fore of mind for security analysts, but rather than punishment and imposing a cost on adversaries in the form of nuclear annihilation, the focus is on denial and raising an adversary’s costs of action. The question is, What does credible deterrence by denial look like for North America in the 2020s?

This analysis briefly examines the strategic logic underpinning the need to modernize North American defense, focusing primarily on NORAD and deterrence by denial. It is vital that structural changes to the North American deterrence posture, including necessary investments, are made to alter adversarial perceptions so that North America cannot be held hostage. Beyond the need to modernize NORAD’s early warning and defense control capabilities to meet the new threat environment, both countries
must modernize NORAD—the organization—and rethink the importance of protecting the North American homeland.

The Strategic Rationale for Modernization

In the immediate post–World War II era, the United States and Canada paid significant attention to and made resource investments in North American air defense cooperation. This focus led to the creation in 1957 of a binational command—the North American Air Defense Command, which centralized operational control of continental air defenses against the threat of Soviet bombers. Attention to NORAD waned, however. The defense of North America and NORAD’s contribution to that mission, especially since the end of the Cold War, have largely taken a backseat to Canadian and American strategic priorities and investments.¹

North America has not been entirely neglected. As evidence, in the 1980s, the 1950s-era Distant Early Warning Line radar system was modernized to create the existing North Warning System (NWS)—a series of uncrewed long- and short-range radars stretching from Alaska, through Canada’s Arctic, and down the East Coast.

Moreover, in the wake of 9/11, internal air radar feeds from the Federal Aviation Administration and NAVCanada were integrated with NORAD’s NWS feeds to warn of approaching threats, creating a more complete air picture for the NORAD and USNORTHCOM Command and Control Center. Now, NORAD monitors the internal air picture and the (usual) air approaches to North America. Nonetheless, continental defense (Canadian parlance)/defense of the homeland (US parlance) has not been a priority. Two factors explain this situation.

First, drawing from the interwar and World II experience and the demands of the Cold War, the strategic priority of both countries continues to be overseas commitments or forward defense (the “away” game).² Second, beginning in the 1960s with the development and deployment of long-range intercontinental and submarine-launched ballistic missiles, with no defense possible at the time, the focus was on deterrence by punishment.

Beyond the need to have early warning of a strategic attack, a mission assigned to NORAD, defense of North America was based on the offensive threat of American strategic nuclear retaliation. Indeed, it was largely assumed that any Soviet Union attack against North America could quickly escalate into a nuclear exchange because of the deterrence by punishment logic—a defense, of sorts, for North America. Air defense was not entirely forgotten or ignored but became a secondary concern to early warning of an attack.³ In the 1960s, the famous hardened Combat Opera-
tions Center in the Cheyenne Mountain Complex in Colorado Springs was completed to withstand a nuclear attack, and a series of radars, radar nets, and other early warning attack systems were brought online.  

Today, the overseas priority has not changed (consider, for example, the pivot to the Indo-Pacific), but the North American threat environment has changed significantly. Successive NORAD and USNORTHCOM commanders have raised concerns about the vulnerability of North America—emanating from Russia and China primarily—linked to a new generation of advanced strike weapons.

Most recently, the former commander of NORAD and USNORTHCOM, General Terrence J. O’Shaughnessy, USAF, and the deputy director of NORAD operations, Major General Peter M. Fesler, USAF, provided the fundamental strategic logic for significant investments in North American and NORAD defense modernization. As the American way of war has focused on large deployments overseas to project overwhelming force, the solution for adversaries “is to prevent deployment in the first place.” North America thus becomes a primary target and will be vulnerable to subversion and coercion as well as conventional and nuclear attacks. The requirement to raise the costs of action by adversaries against North America is paramount.

Emphasizing North America is no longer a sanctuary, O’Shaughnessy and Fesler argued a credible deterrence by denial posture is vital to support the credibility of the American strategic deterrence posture overseas. North American vulnerability may embolden China or Russia to challenge the status quo in the Asia-Pacific or European theaters, generating a major crisis and possibly war. Specifically, new strike capabilities (including hypersonic weapons) enable competitors to threaten, and, in a worst-case scenario, destroy North American military bases and embarkation points vital for reinforcing forward-deployed forces.

With few extant defensive capabilities at home to meet this threat, the willingness of the United States to stand firm in a crisis overseas would be at issue. Ensuring the capacity to detect, deter, defend, and defeat such threats to North America via denial is essential to reduce incentives for Russia and China to challenge the overseas status quo by threatening the homeland.

Of course, issues surrounding deterrence postures and credibility, both globally and for North America, are complicated and contentious in the new world of great power rivalry. Among others, the threats posed by new, dual-capable nuclear and conventional strike systems will be center stage in future debates about North American and NORAD defense modern-
ization. Nonetheless, the vital issue is to recognize and detail North American deterrence-by-denial requirements, including the need to go beyond simple resource investments to modernize the Canada-US defense relationship and NORAD’s place within it.

North American Deterrence Requirements

Any evaluation of deterrence requirements must first recognize the objective is North America, not Canada or the United States separately. A threat to either is a threat to both. From this starting point, the current structure of the defense relationship underpinning a credible North American deterrence-by-denial posture is itself problematic. The relationship, at its strategic and operational levels, is divided in several ways with no overarching true central structure to provide unity of effort and command for North America. Part of the relationship is binational as embodied in NORAD with its functional responsibility for aerospace (air and ballistic missile) and maritime warning and aerospace control (air). The remaining parts are bilateral.

Overall cooperation and coordination are implemented through the tri-command arrangement consisting of NORAD, USNORTHCOM, and Canadian Joint Operations Command (CJOC)—N2+C—established roughly a decade ago. It is at best an informal command arrangement, and whether it will evolve to become a more formal, centralized North American command depends on political will.

Moreover, the N2 legs of the arrangement are devoted strictly to North America, while CJOC is responsible for all Canadian military operations, home or abroad, that do not involve NORAD or special forces. At one time, CJOC devoted most of its attention and limited resources to overseas operations. Today, due to climate change, COVID-19, and the need to provide assistance to Canadian civilian agencies, the split in terms of resources and attention is 50 percent at home and 50 percent overseas.

Seams to Consider

First identified by the now defunct Binational Planning Group stood up after 9/11 to consider how best to defend North America, N2+C, along with the mixed binational and bilateral components of the North American defense relationship, have created North American command “seams” with implications for deterrence credibility. For example, while NORAD can warn of a maritime threat to North America, the US Navy and Royal Canadian Navy operate unilaterally and bilaterally and under US-
NORTHCOM and CJOC commands, respectively, with different areas of responsibility and jurisdiction. An adversary need only find the seams between CJOC and USNORTHCOM areas of responsibility, and precious response time will be lost coordinating an ad hoc, bilateral solution to fortify the command and geographic seams.

Another seam—between denial and punishment or raising versus imposing costs—is directly related to the concept of deterrence. The North American command components (N2+C) operate in the denial sphere. The United States’ punishment authority and capabilities relative to North America are assigned to US Strategic Command (USSTRATCOM), another command within the US Unified Command Plan. Canada has no such capability other than via its Ally status with the United States and via NATO.

Regional commands in the Unified Command Plan, including USNORTHCOM, possess both denial and punishment authority and capabilities; NORAD does not. For example, NORAD warns of an inbound ballistic missile, but the defeat decision and capability rest entirely with USNORTHCOM with no Canadian input. Therefore, Canadian personnel assigned to NORAD on the NORAD and USNORTHCOM Command and Control Center watch floor will see and warn of an attack. But then they will step aside for USNORTHCOM US personnel to decide how best to react.

Certainly, such defeat authority and capabilities could be given to NORAD as they partially once were when US Space Command (USSPACECOM) and NORAD were situated under the same commander with punishment authority. (After 9/11, USSPACECOM was separated and dissolved and its responsibilities folded into United States Strategic Command.) Successive Canadian governments, most notably the Martin government in 2005, have long ceded punishment to the United States for domestic political reasons.

In terms of the US part of the deterrence equation, USNORTHCOM also confronts horizontal, geographic seams as a function of the Unified Command Plan. It shares Alaska with US Indo-Pacific Command (USINDOPACOM), and many of USNORTHCOM’s capabilities are held by USINDOPACOM (fig. 1). There are three geographic combatant command seams in the Arctic approaches to North America—USNORTHCOM, USINDOPACOM, and US European Command (USEUCOM). Three geographic combatant command seams also impact North America as a whole—the Atlantic and USEUCOM, the Pacific and USINDOPACOM, and the south via US Southern Command.
In addition to jurisdictional, authority, and geographic seams, North American deterrence also confronts domain seams. Reflective of the military service structure, the geographic domains of air, land, maritime, and space remain conceptually and structurally separate even though these domains increasingly blur together as a function of technological change and hybrid tactics. Thus, for example, a maritime threat as a function of cruise missile technology can quickly transition into an air-breathing threat.

The United States’ solution is to adopt Joint all-domain command and control (JADC2) to connect sensors from all military services—Air Force, Army, Marine Corps, Navy, and Space Force—into a single network. The implications for the North American deterrence structure remain to be seen, but JADC2 implies the potential merger of punishment and denial. A long list of obstacles remains to achieve this concept within the US military, let alone the challenges involved in including the Canadian military. Ideally, some level of discussion and engagement with Canada in
JADC2 development is necessary versus the traditional approach wherein the US decides on a course of action, and Canada reacts.

Two additional domains require consideration. The first is not a traditional domain per se but involves violent extremists. Violent extremism (formerly terrorism) has significantly receded from defense and security agendas (even at a time when there is a rise of right-wing, national violent extremism as opposed to foreign and mainly radical Islamic forms of terrorism). Moreover, many national security decision makers today question whether terrorists can truly be deterred.

But this domain cannot be ignored, as it resides in the seam between military and civil security agencies. The other domain—cyber—has risen noticeably on the defense and security agenda and with it, the cognitive domain (think misinformation, disinformation and malinformation campaigns). In these worlds, denial and punishment are also separated—punishment in the cyber world appears to be the exclusive domain of US Cyber Command—but denial entails the military, civilian security agencies, and the private sector.

**Capability Gaps**

Beyond structural seams, notable capability deficiencies—gaps—are identified in several reports, including the Heritage Foundation’s 2021 Index of US Military Strength, which graded all services’ capabilities as “marginal.” Further, the Heritage Index, reflective of many studies on the US military, does not consider North America: only suitability for operations in Europe, Asia, and the Middle East are assessed. United States Northern Command and NORAD employ the concepts of detection, denial, defense, and defeat. Although these elements are not necessarily understood to be linear, detection and defense are the key concepts to evaluate capability deficiencies. In this regard, a credible capacity to detect and defend equates to a credible deterrence-by-denial posture.

**Detection**

Detection is the first ingredient of denial credibility and is central to NORAD’s mandate. The North American aerospace warning mission is essential as is its maritime warning mission. Both missions have complicated national and bilateral elements embedded in their processes, especially in the maritime domain. Three key deficiencies stand out. First, NORAD’s air warning component is almost exclusively defined as syn-
onymous with the information provided by the North Warning System rather than a North American warning system.\textsuperscript{15}

The NWS is technically obsolete; as a result and notwithstanding new artificial-intelligence-inspired additions, NORAD’s air warning capability is potentially on the precipice of failing. Because of its 1970s technology and physical location, the NWS is challenged to detect long-range air- and sea-launched cruise missiles, not to mention drones that fly at speeds and altitudes not envisioned for 1970s air threats.

All relevant parties recognize these deficiencies. In response, a bina- tional structure is in place to identify sensor solutions and requirements to move and filter large quantities of sensor data into NORAD for analysis and action (NORAD modernization). Nevertheless, there seems to be no pressing urgency to move forward. In 2017, in the joint statement released after the summit between Canadian Prime Minister Justin Trudeau and US President Donald Trump, and reiterated in the first, virtual summit with US President Joe Biden, the leadership of both countries placed North American defense and NORAD modernization among their pri- orities. To date, too few significant investments have occurred.\textsuperscript{16}

Certainly, as the future North Warning System is likely to entail a complex array of ground-, air-, maritime-, and space-based sensors, technology hurdles do exist, especially in terms of systems integration. The danger lies in waiting for the final, perfect solution rather than building the sys- tem as partial solutions come online. Such a delay will leave a major detection gap for some time to come. Indicative of this trend, the current NWS radars that will reach the end of their life cycle in 2025 are already set to be extended until 2035.

Second, the future NWS/North American Warning System sensor sys- tem remains largely conceptualized as a perimeter system, looking outward from the continent (fig. 2). In the wake of 9/11, NORAD acquired an internal air picture of North America through its link to the US Federal Aviation Administration and NAVCanada radars. But it is unclear whether these internal radars possess a cruise missile detection and tracking capability and/or future improved drone-tracking technology. A pe- rimeter system must be augmented by internal detection capabilities, in the very least as assurance should the defense side of the equation fail at the perimeter.
Third, the detection domains remain largely separate rather than integrated into an all-domain detection and thus analysis structure. While NORAD has air and ballistic missile warning functions, and with the latter, a space-tracking function as well, these appear to be largely independent, reflecting the traditional division between air and outer space. Yet, as the future North American Warning System will likely comprise a significant space-based component, threat detection against these key space-based assets is essential. Moreover, threats to these components also extend to a wide range of space-based assets vital to the military and the economy, especially in low Earth orbit.

Clearly, such threats (especially to space-based assets) are in fact threats to the North American homeland. Moreover, attacks against these assets are not just a physical attack against the territorial homeland but could involve the direct loss of life. Adversary calculations of the repercussions of their attacks on assets alone will be distinctly different from a direct threat or attack against North America. This factor does not imply NORAD should acquire a space defense mission per se; rather, NORAD’s
ballistic missile warning mission should include detecting threats against space-based assets. Detecting these threats should also be part of its integrated tactical warning/attack assessment function.

In addition, the development of hypersonic weapons technology foreshadows the merger of space and air into a true “aerospace” domain. As with the maritime domain, the ballistic threat of maneuverable hypersonics may transition into a maneuverable air threat operating between space and air. That is, the space, aerospace, and air domains need to be integrated into a single detection domain, along with the maritime domain, to generate an integrated, all-domain North American common operating picture. The final geographic domain—land—is less important to include: three oceans effectively mitigate a land-invasion scenario. The cyber domain, however, is vital.

**The Cyber Complication**

Threats emanating from the cyber world have attracted growing attention over the last several decades. For many years, the air forces of the United States and Canada (and to a lesser degree NORAD) have made a claim on the domain, notwithstanding US Cyber Command and its unclear role in the North American deterrence equation. Regardless, central to the detection problem in the cyber domain and distinct from the other domains, attribution of a cyberattack is extremely problematic. Due to the complexities of the internet and the ability of states such as China and Russia to employ—implicitly or explicitly—private actors, it is difficult to ascertain whether any attack has been motivated just for mischief, for criminal purposes, and/or for state purposes.

Moreover, this domain is structurally more complicated than the maritime domain. It involves not only the military relative to its own systems and other government agencies but also private actors within the economic system. The overwhelming majority of cyber critical infrastructure resides in private hands within the integrated North American economy. In this regard, private business interests related to corporate viability act to some degree as disincentives to report cyberattacks.

So long as North American officials continue to emphasize cyber vulnerabilities and fear the consequences, adversaries have incentives to exploit the cyber world. Whether the attempt by Russia, as attributed, to influence the 2016 US presidential election had any real impact on its outcome is a moot question. It is the attempt itself and the fears it generated of other, potentially more devastating attacks that Russia uses to its advantage. At the core of this problem is detection and attribution.
A cyberattack occurs in near real time, usually with no warning or with such obfuscation that targets may not even realize they are under attack. In contrast, the kinetic world provides, to varying degrees, early warning signals due to advanced intelligence and surveillance capabilities. One can expect, for example, that long-standing, normal patterns of military activity will be altered in preparation for employment (e.g., mobilization of personnel and assets).

Such deviations do not necessarily mean a decision to use force has been made. In some cases, preparations may simply be a means of threat signaling to alter adversarial responses, with no intent to escalate to the use of force. Political contexts that suddenly change or evolve over time also provide signals. Regardless, in the kinetic world, the probability or fear of a bolt from the blue is less likely.

Cyberattacks and probing are, however, a world of “bolts from the shadows.” As an element of deterrence, in this case by punishment, state-sponsored or directed deterrence attacks may simply be intended to demonstrate what an adversary can and might do in the future to alter calculations. In other cases, these attacks are meant to disrupt a state’s ability to track and react at a later point in the decision-making process or to obfuscate an adversary’s actions.

Operating at a low level of effect and thus having only a temporary, limited, and marginal impact—shutting down a website or a pipeline—the act is meant to indicate the potential to do more damage. Moreover, at least to date, these attacks are calculated as insufficient to generate a kinetic response. Furthermore, the problem of attribution and thus plausible deniability also adds complexity to the detection side of the equation. This complexity is compounded further with the potential for embedded computer viruses, such as the case of Stuxnet in Iran, that may remain undetected until triggered under certain conditions—a potential attack in the making.

Political warning signals, too, can emerge to challenge the status quo and can be generated and transmitted across the complicated North American cyber world, requiring greater vigilance. Additionally, ongoing analysis to discern potential patterns of cyberattacks over time and space may provide some modicum of prediction and thus detection. Ultimately, however, detection is exclusively in the hands of the owners of the private, public, and military networks. As a result, detection capabilities, and thus vulnerabilities, vary widely across the North American cyber world.

While one cannot expect every network in the North American cyber world to implement a common standard, and apart from the problem of
determining what critical infrastructure is and is not, critical infrastructure across North America needs to adopt a common detection standard in terms of detection software. In addition, intelligence or information sharing must be formalized across the private, public, and military divides following cyberattacks.

The state of the cyber domain in North America is reminiscent of the state of the intelligence world prior to 9/11 and of the maritime domain prior to the undertaking of significant steps in the years following those attacks. Improvements to threat detection in the maritime domain included NORAD’s acquisition of a maritime warning mission, the creation of the US National Maritime Integration Intelligence Office, and the establishment of Canada’s Marine Security Operations Centres.

In this regard, a NORAD or perhaps N2 cyber detection mission for North America might be conceptualized based on maritime warning and its protocols. Designed not to duplicate existing and evolving private/public actors and processes, this mission would provide a centralized analytical function based upon its integrated tactical warning/attack assessment function. This mission would serve as the only North American eyes at the end of the intelligence collection process as it currently exists nationally and bilaterally. As NORAD was a key promoter and supporter of greater interagency cooperation to enable its maritime warning mission, it may also act to spur greater intelligence cooperation and information sharing across North America as a whole.17

**Defense**

Alongside detection, defense is the second capability component of a credible North American denial deterrent. As with detection, existing gaps may affect adversary and North American (Canada and the United States) deterrence calculations. Several stand out in the traditional defense domains. Assuming Canada agrees on a CF-18 replacement and given the presence of US anti-cruise missile interceptors, the question becomes whether intercept density relative to NORAD’s assigned assets is sufficient to defend against cruise missile threats.

NORAD is also looking at existing northern forward operating locations and other possible locations farther south to meet maritime threats and potentially provide some form of layered defense. Additionally, there is a recognized requirement for in-flight refueling capabilities, and the deployment of anti-cruise missile point defenses must be considered. These factors strongly suggest more resources need to be dedicated to the
air defense component of North American deterrence and then integrated into the detection side of the equation.

Related to air defense requirements, the aforementioned merger of air and space into a true aerospace domain raises the subject of combining air and missile defense capabilities. This process is already underway with the US Army developing the Integrated Air and Missile Defense Battle Command System.\(^\text{18}\)

Merging these capabilities raises the thorny issue of Canadian participation and with it, concerns related to intercept priorities and centralized command and control, which in part derailed Canada’s participation in ballistic missile defense (BMD) in 2005.\(^\text{19}\) A reversal of Canada’s “not yes” to missile defense is likely to entail assigning command and control to NORAD. Doing so ensures Canada’s direct participation in decision making per the binational agreement and potentially clears the way for the merger of the J-3 position in the NORAD-USNORTHCOM command center—the only position currently not combined. Otherwise, the credibility of the North American denial posture is undermined, with Canadian vulnerability providing a venue for an adversary to exploit.

Relatedly, assuming the United States proceeds with a third continental missile defense site in the Northeast, its requirements may entail an advanced tracking and cueing radar deployed to Canada. Such a radar, in turn, would also likely serve other valuable detection functions related to North American defense.

**Maritime Complexity**

Turning to the maritime domain, beyond the logic of evolving the current bilateral structure of the Canada–United States (CANUS) naval relationship into a binational one, the defense equation is problematic. Naval preferences are currently forward-defense oriented against cruise-missile-capable surface and subsurface ships (Archer class) rather than homeland-defense oriented against sea-launched cruise missiles (the Arrows). While not ignoring the defense value of this preference, the Archers are located outside the Royal Canadian Navy and USNORTHCOM’s areas of responsibility. Defense against the Arrows is secondary when it should be primary for North American deterrence.

In this regard, major surface combatants (including the future Royal Canadian Navy combat vessel) need to deploy sufficient anti-cruise missile air defenses, and these defenses need to be integrated into NORAD’s air defense assets. At a minimum, the role of maritime assets must be fully
integrated into NORAD exercises to bolster North American deterrence requirements.

Other Domains and Resilience

In the terrorism and cyber domains, defense has long been outside the military mandate. The military has been assigned the role of second responder to deal with the consequences of an attack. Defense is in the hands of police forces and bilateral cooperation between Canada and the United States. There appears to be no reason to change the military’s role except to ensure protocols governing the provision of mutual support are fully developed in response to a major incident. In this regard, the concept that has recently emerged is deterrence by resilience. Simply stated, capabilities are developed to mitigate the consequences of a major terrorist or cyber event quickly and effectively, thereby reinforcing deterrence credibility.

In many ways, deterrence by resilience is not a denial posture. Rather, it is a recognition that denial is not possible. In traditional military jargon, it is a damage limitation posture that serves to enhance credibility, demonstrating to an adversary that its attack will unlikely reap expected benefits. Canada and the United States need to enhance their ability to assist civil agencies. Furthermore, this assistance should not be constrained by the border, and, at a minimum, such requirements should be a priority for the tri-command structure.

Conclusion

From the perspective of North American homeland defense and security, the current CANUS command structure and capabilities are locked into an exclusive deterrence-by-denial posture. Punishment as an alternative is not an option, which does not mean that an adversary does not confront a credible punishment threat. Rather, the punishment threat and thus punishment capabilities reside elsewhere and are exclusively American. The question then is whether the CANUS part of the equation is adequately structured and resourced to present a credible denial threat to an adversary. Arguably, an adversary could be dissuaded from directly threatening or attacking independent of a punishment threat conceived of as a last resort.

Importantly, any adversary, regardless of perceptions of denial credibility, cannot ignore or simply discount punishment given the reality of US strategic conventional and military capabilities. Of course, as a psychological theory designed to alter adversarial thinking and calculations, it is
extremely difficult to know or predict how an adversary thinks and responds to a deterrence posture. Perhaps, then, what is more significant is how North American decision makers think about their own credibility. It is here that the North American conundrum resides.

The North American component of the US-led Western global deterrence posture should exist as the central deterrence hub such that an adversary does not perceive it as a vulnerability that could be exploited to deter US-led responses to regional challenges. Yet it is questionable whether US and Canadian decision makers even think in these terms about the homeland.

Both arguably remain fixated on the overseas components, with North America as an afterthought despite the rhetoric.²⁰ Moreover, beyond NORAD and USNORTHCOM and to a much lesser degree CJOC, two different viewpoints exist. The American view is that neither Russia nor China would dare strike North America due primarily to its overarching military superiority and last-resort strategic punishment capabilities. The Canadian view is really a nonview. Essentially, Canada does not really think in deterrence terms because it lacks the capabilities to deter credibly and because deterrence is an American responsibility, with Canada helping and warning where it can.

The net result may be a (vicious) feedback loop. An adversary comes to believe it can exploit homeland vulnerability, thus emboldening it to undertake a regional challenge by threatening actions short of war to deter a regional overseas response by North America. The United States and, to a lesser degree, Canada quickly recognize their vulnerability (and that of vital overseas Allies and partners) and are unwilling to respond effectively, being forced to fall back on a strategic punishment threat to deter. This approach, in turn, emboldens the adversary to initiate further challenges, raising doubts among overseas Allies and partners that the United States will defend them.

The basic answer is to alter deterrence thinking in North America. Structural changes, including necessary investments, to the North American deterrence posture must be made to alter adversary perceptions so that North America cannot be held hostage. In fact, the current North American deterrence-by-denial posture remains embedded in an outdated Cold War mindset that has largely evolved in an ad hoc manner.

These changes are obviously easier said than done. Despite the best efforts by senior NORAD and USNORTHCOM officials to communicate this message, it may take an unexpected overseas regional challenge resulting in a major crisis in which the lack of North American denial credibil-
North America's Imperative: Strengthening Deterrence by Denial

Andrea Charron
Dr. Andrea Charron is an associate professor of international relations and director of the Centre for Defence and Security Studies at the University of Manitoba in Winnipeg, Manitoba, the home of Canada’s NORAD headquarters.

James Fergusson
Dr. James Fergusson is a professor of international relations and deputy director of the Centre for Defence and Security Studies at the University of Manitoba. His latest book is Beyond Afghanistan: An International Security Agenda for Canada (2016).

Notes
5. O'Shaughnessy and Fesler, Hardening the Shield, 3.


17. Charron, Fergusson, and Allarie, Left of Bang, 42–44.


Will Emerging Technology Cause Nuclear War?: Bringing Geopolitics Back In

MATTHEW KROENIG

In order to fully understand the link between nuclear stability and emerging technology, the current geopolitical situation must be accounted for. Incorporating emerging technologies into US, Ally, and partner militaries will likely reinforce the prevailing global strategic stability.

Will emerging technology cause nuclear war? For more than 70 years, the world has avoided major-power conflict, and many attribute this era of peace to nuclear weapons. In situations of mutually assured destruction, neither side has an incentive to launch a nuclear first strike because doing so will only result in self-annihilation. Maintaining secure, second-strike capabilities—the ability to absorb an enemy nuclear attack and respond with a devastating counterattack—is the key to deterrence. Recently, analysts have begun to worry, however, that new military technologies may call into question this model of global strategic stability.

The world is experiencing a fourth industrial revolution (4IR) in which a wave of new and transformative technologies is being developed, including artificial intelligence (AI), additive manufacturing, quantum information technology, hypersonic missiles, biotechnology, and directed energy. While these technologies are expected to have profound implications for societies and economies, most are dual use and will also affect national security, including nuclear strategic stability.

According to an emerging conventional wisdom, new technology may upset nuclear strategic stability by calling into question the survivability of nuclear forces. The solution, according to some analysts, is for nuclear-armed states to eschew military applications of at least some of these technologies and lead an international effort to control their spread. But these studies too often consider new technology and nuclear strategy in the abstract without adequately considering the prevailing geopolitical context into which these new technologies have been introduced.

This article argues understanding the link between new technology and nuclear stability must consider the prevailing geopolitical context. For the past several decades, the United States, its Allies, and like-minded part-
ners have formed the core of the existing international order. They have benefited from this system and would like to see it strengthened, revitalized, and defended. If the new technologies of the 4IR are incorporated into US, Ally, and partner militaries, then any advantages they provide will likely reinforce the prevailing distribution of power and existing sources of strategic stability.

In contrast, China and Russia are revisionist powers intent on disrupting or displacing the US-led system, and they would likely employ new technological advantages to pursue revisionist aims. The greatest danger from emerging technology for nuclear stability, therefore, may result from the possibility that new technology provides Russia or China an enhanced military advantage over vulnerable US Allies and partners, leading to a regional conflict with a significant risk of nuclear escalation.

This article contributes to the growing literature on new technology and nuclear stability by emphasizing politics take precedence over technology. Technology rarely transforms states. More commonly, states employ technologies to achieve preexisting ends. It is not simply the technologies themselves that are destabilizing but the geopolitical ambitions of the states that possess them.

In emphasizing the divergent positions of the United States of America and its nuclear-armed rivals in the international system, this article also contributes to a growing body of literature that takes seriously hierarchy in international relations theory. The United States, the international system’s leader for the past several decades, is likely to use new technology to reinforce its advantageous position within the existing international order. China and Russia will most likely employ new technology in bids to erode America’s privileged position. Analyses not grounded in an understanding of these states’ different positions in the prevailing international order risk overlooking this important source of variation in conflict behavior and nuclear-escalation dynamics.

This framing of the problem leads to a different set of policy implications. The United States and its Allies and partners must retain second-strike capabilities, preserve current power distributions, maintain an innovation edge, and prevent the proliferation of destabilizing military technologies to revisionist powers.

**Emerging Technology and Nuclear Stability**

A growing body of literature expresses concern that emerging technology could undermine nuclear strategic stability through its effect on nuclear second-strike capabilities or on dual-use systems, including nuclear com-
mand, control, and communications. For example, cyberattacks and conventional hypersonic missiles might be combined to provide credible first-strike capabilities against an adversary’s nuclear forces, and advanced, directed-energy missile defenses could be employed to absorb enemy nuclear retaliation.

By threatening an enemy’s secure, second-strike capabilities, these new technologies might undermine nuclear strategic stability. If leaders believe they can disarm their opponents, they may be motivated to use nuclear weapons first in a crisis. Alternatively, leaders fearing a disarming attack may choose to use their nuclear weapons first before they lose them.

New technology may also contribute to accidental or inadvertent nuclear escalation by threatening dual-use command and control assets in space and cyberspace or by compressing time for leadership decisions. Leaders may choose to initiate a nuclear war under the mistaken belief that an enemy nuclear attack has already begun or is imminent.

There are several limitations, however, to the existing analysis. First, the underlying theory of nuclear conflict this body of thought advances is debatable. It rests heavily on the “use it or lose it” cause of nuclear war, but use it or lose it is rooted in the logical fallacy of the false dilemma. States have many options in a crisis other than suffering a disarming nuclear attack or launching one. Moreover, faced with a range of choices, the use-it-or-lose-it logic assumes a state will intentionally choose to initiate a nuclear war—the most risky and costly available option. The use-it-or-lose-it pathway to nuclear war, therefore, is in tension with mainstream nuclear deterrence theory that maintains states will be reluctant to conduct a deliberate attack on another nuclear-armed state.

A second limitation of this approach is that theories of nuclear instability developed in the early days of the Cold War are in tension with current understandings of the causes of war in contemporary international relations theory. The nuclear stability framework rests on the notion that parity in the balance of power is associated with peace. The prevailing bargaining model of war, however, maintains that parity contributes to uncertainty about the balances of power and resolve, which hinders efforts to reach negotiated settlements short of armed conflict. The empirical record supports this theory and demonstrates parity in the balance of power is associated with conflict, and uneven balances of power are associated with peace. Situations of obvious strategic nuclear superiority, therefore, may be more stable than situations of strategic parity.

Perhaps the most important limitation of the existing debate is its tendency to theorize in the abstract, divorced from real-world geopolitical
conditions. Proper nouns are rarely used. States, in these analyses, are treated as black boxes endowed with nuclear weapons and new technology, facing off against a mirror-image rival. The question of interest to scholars is whether the new technology could incentivize a generic nuclear-armed state to launch a nuclear first strike. The varying geopolitical positions, foreign policy ambitions, or ongoing political conflicts of interests among the major nuclear powers in the world today—the United States, China, and Russia—are not of immediate interest.

New technology is not acquired by black boxes, however. Emerging technology is diffusing into an international system in which the United States has been the world’s leading power for the past several decades. This system is increasingly being challenged by nuclear-armed competitors, including China and Russia. States will likely use the advantages provided by new technology in a bid to advance preexisting foreign policy objectives.

In short, scholars have devoted excessive attention to abstract conjectures about interactions among technologies and weapons systems. Missing from the literature is an examination of how the diffusion of new technology might affect the behavior of today’s principal nuclear-armed powers and, in turn, the strategic stability of the contemporary international system.

The next section provides a novel framework, grounded in the prevailing geopolitical context, for understanding how new technology might affect nuclear strategic stability. Namely, the spread of new technology to the United States and its Allies and partners—status quo powers at the core of the existing international system—will tend to shore up sources of strategic stability. Conversely, the spread of new technology to revisionist powers China and Russia presents the greatest risk of conventional conflict that might escalate and threaten nuclear strategic stability.

**Geopolitical Context**

The United States has been the most powerful country in the international system by almost any measure since 1945. In the aftermath of World War II, the United States and its Allies constructed the outlines of the world we inhabit to this day. They attempted to construct a security order that would prevent the recurrence of major conflict. Alliances in Europe and Asia—and the extension of the US nuclear umbrella—deterred conflict in those geopolitical regions and contributed to peace and stability. American nuclear security guarantees also dissuaded Allies from pursuing independent nuclear arsenals. The Treaty on the Non-Proliferation of Nuclear Weapons slowed the spread of nuclear weapons to additional states.
Throughout the Cold War, the United States and its Allies and partners in the free world competed with the Soviet Union for global preeminence. Through much of this competition, Moscow was a revisionist power with the explicit goal of exporting its Marxist-Leninist revolutionary model abroad.\textsuperscript{19} It challenged the status quo repeatedly, including by initiating crises in Berlin, Cuba, and elsewhere.\textsuperscript{20} Indeed, according to the International Crisis Behavior Project data set, Moscow initiated 13 of the 17 crises between the Soviet Union and its nuclear-armed opponents during the Cold War.\textsuperscript{21} Many of these crises entailed a significant risk of nuclear escalation, and fears of nuclear instability featured prominently in the discussions of defense analysts during this time.\textsuperscript{22}

Following the collapse of the Soviet Union, the United States remained a unipolar power.\textsuperscript{23} Alongside its Allies and partners, Washington used this position to deepen and expand the US-led, rules-based international system. Countries previously locked behind the Iron Curtain in Europe rushed to join the West, adopting democratic forms of government and market-based economies and entering NATO and the European Union (EU). Some analysts predicted the “end of history,” as there was no obvious competitor to the Western model of free politics and open markets.\textsuperscript{24}

As great power competition receded into the background, so too did fears of nuclear instability. The 2010 US \textit{National Security Strategy} mentions Russia and China more often as partners for cooperation to address shared challenges than as military threats. Indeed, the document states the risk of nuclear war with these powers was extremely remote, as “the specter of nuclear war has lifted.”\textsuperscript{25} Reducing the role and number of nuclear weapons became a central objective of US nuclear policy.

In recent years, however, great power competition has reemerged.\textsuperscript{26} Russia is dissatisfied with the spread of the US alliance system to its sovereign borders. It has invaded its neighbors—Georgia in 2008 and Ukraine in 2014—with the goal of preventing them from joining Western institutions such as NATO and the EU. Putin has declared the collapse of the Soviet Union the “greatest geopolitical catastrophe of the 20th century” and aspires to recreate a greater Russia.\textsuperscript{27} Russia desires the dismantlement of America’s alliance architecture in Eastern Europe.\textsuperscript{28} As Putin has said, he wants “new rules or no rules.” US strategists fear Putin may use military power in a bid to force that objective.\textsuperscript{29}

China has also become more assertive in recent years.\textsuperscript{30} For decades, Chinese leaders followed former premier Deng Xiaoping’s dictum that China should “hide its capabilities and bide its time.” In recent years, however, Chinese president Xi Jinping has abandoned that doctrine. China
has expressed its dissatisfaction with the territorial status quo in Asia. Beijing regards Taiwan as a renegade province that will eventually be re-incorporated into China. It has not ruled out the use of force to achieve this objective and has increased its military activities in the Taiwan Strait.31

Further, China has ongoing territorial and maritime disputes with several of its neighbors, including the land border with India, the Senkaku/Diaoyu Islands with Japan, and throughout the South China Sea with several claimants in Southeast Asia. These disputes have been sources of increased military contestation in recent years, and all are potential flashpoints for great power military conflict.

The unclassified Summary of the 2018 National Defense Strategy of the United States of America: Sharpening the American Military’s Competitive Edge identifies the return of great power competition with Russia and China as the principal threat to US national security.32 Of particular concern is the risk that Russia or China could attack a vulnerable US Ally in Eastern Europe or East Asia, respectively, presenting the United States with a fait accompli.33

Washington would be faced with a difficult decision. It could fight a major war with a nuclear-armed rival to liberate a beleaguered Ally. Alternatively, the United States could back down to avoid conflict, but at the risk of failing to protect a treaty ally. This path would allow Russian and Chinese aggression to stand, undermine US credibility, and call into question America’s other formal alliance commitments. A congressionally mandated National Defense Strategy Commission report warns of the possibility of a major war with Russia or China—one that the United States and its Allies and partners might lose.34

This understanding of the geopolitical context provides the necessary baseline against which to assess the likely impact of new technology on global stability.

**New Tech Arms Race**

Many analysts believe the emerging technology of the 4IR could profoundly affect military capabilities and operational concepts.35 New technology has had revolutionary effects on warfare and international politics throughout history from the Bronze Age to the gunpowder and nuclear revolutions.36

New technologies with direct military application are in development, including AI, quantum information technology, hypersonic missiles, directed energy, additive manufacturing, and biotechnology. How exactly these technologies will affect the future of warfare is still uncertain. The
National Defense Strategy Commission report charges that the United States lacks clear operational concepts for combat with Russia and China. Still, there is reason to believe these new technologies could have meaningful military applications but perhaps not to the advantage of the United States and its Allies and partners. At present, Russia and especially China might transcend the United States and its Allies and partners in some key 4IR technologies.

Indeed, AI could transform the future of warfare, including through the development of lethal autonomous systems. These “killer robots” may lower the threshold of conflict by allowing political leaders to take a country to war without risking the lives of human soldiers. When produced in large numbers, these drones could operate in swarms that overwhelm enemy military platforms and targets.

Artificial intelligence could also be employed to rapidly sort through vast quantities of data, improving intelligence, surveillance, and reconnaissance and making it easier to track and target enemy forces. The United States retains important advantages in AI, including through its world-leading university system. But China, with its large population and surveillance tactics, has access to more data to train its AI algorithms. Beijing is also less constrained by ethical and moral concerns and has the lead in some applications of AI, including facial-recognition technology.

Quantum computing promises information advantages including the ability to have secure, encrypted communications and to decode enemy communications. In its 2021 Military Balance report, the International Institute for Strategic Studies states, “the integration of quantum technologies currently represents one of the most anticipated advances for armed forces . . . There is little doubt that they will have disruptive effect when they are employed at scale.” China may have the edge in this area, as it was the first country to conduct a successful test of a quantum satellite.

Space and cyber are increasingly important military domains. Space-based weapons, sensors, defensive interceptors, and the diffusion of counterspace capabilities will make space an increasingly contested military environment. The United States is relatively more dependent on space-based assets and computers than its rivals, and the US Department of Defense warns Russia and China will likely employ cyber and counterspace attacks in the early stage of any conflict with the United States in a bid to disrupt US command, control, communications, computers, intelligence, surveillance, and reconnaissance (C4ISR).

Hypersonic missiles, maneuverable and able to travel at over five times the speed of sound, could allow states to conduct low- or no-warning at-
tacks and to evade missile defenses. These weapons could also execute large-scale, nonnuclear strategic attacks, the rate of speed compressing the decision-making time leaders have to respond to such attacks. Although the United States developed the initial concepts for these weapons, Russia and China have prioritized their production, testing, and deployment. China has conducted more hypersonic tests than any other nation, and Moscow and Beijing have deployed hypersonic weapons.

Many other emerging technologies have military applications. Directed-energy microwaves and lasers could allow states to develop more effective integrated air and missile defense systems or to degrade an enemy’s command and control. Additive manufacturing could greatly reduce the cost of producing component parts of military platforms and creates the potential for large and rapid quantitative increases in weapons systems, from drones and tanks to submarines and nuclear weapons.

Biotechnology could be exploited to produce “super soldiers.” China has genetically engineered beagles with three times the muscle mass of a typical canine, a technology that could possibly be applied to humans. Exoskeletons could provide soldiers with superhuman strength, and brain implants promise superior cognitive performance. China employed exoskeletons in combat in its 2020 border conflict with India.

It is not yet clear how these new technologies, when combined with novel operational concepts, will affect the future of warfare, but it is likely they will. A future state may, for example, be able to use additive manufacturing to produce masses of inexpensive drones directed by new AI algorithms to swarm and overwhelm adversaries. The attack might be preceded by cyber and counterspace attacks that blind an adversary and disrupt its command and control.

Following a successful advance, the country could then employ directed-energy weapons, autonomous mines, and other advanced defenses to lock in territorial gains and thwart enemy attempts to roll back its aggression. It is possible that the first state to hone these technologies and devise effective operational concepts will have a military edge over its opponents.

Novel Applications

How will states use such a newfound advantage? Technology rarely fundamentally changes the nature or objectives of states. More often, states use technology to advance preexisting geopolitical aims. Moreover, enhanced power can result in greater ambition. Given the geopolitical landscape described, it is likely the United States and its Allies and partners at the core
of the international system will behave differently with new military technologies than will revisionist powers, such as Russia and China.

The spread of new technology to the United States and its Allies and partners would likely serve, on balance, to reinforce the existing sources of stability in the prevailing international system. At the end of the Cold War, the United States and its Allies and partners achieved a technological-military advantage over its great power rivals, with the US using its unipolar position to deepen and expand a rules-based system. They also employed their military dominance to counter perceived threats from rogue states and terrorist networks. The United States, its Allies, and partners did not, however, engage in military aggression against great power, nuclear-armed rivals or their allies.

In the future, these status quo powers are apt to use military advantages to reinforce their position in the international system and to deter attacks against Allies and partners in Europe and the Indo-Pacific. These states might also employ military power to deal with threats posed by terrorist networks or by regional revisionist powers such as Iran and North Korea. But it is extremely difficult to imagine scenarios in which Washington or its Allies or partners would use newfound military advantages provided by emerging technology to conduct an armed attack against Russia or China.

Similarly, Moscow and Beijing would likely use any newfound military strength to advance their preexisting geopolitical aims. Given their very different positions in the international system, however, these states are likely to employ new military technologies in ways that are destabilizing. These states have made clear their dissatisfaction with the existing international system and their desire to revise it. Both countries have ongoing border disputes with multiple neighboring countries.

If Moscow developed new military technologies and operational concepts that shifted the balance of power in its favor, it would likely use this advantage to pursue revisionist aims. If Moscow acquired a newfound ability to more easily invade and occupy territory in Eastern Europe, for example (or if Putin believed Russia had such a capability), it is more likely Russia would be tempted to engage in aggression.

Likewise, if China acquired an enhanced ability through new technology to invade and occupy Taiwan or contested islands in the East or South China Seas, Beijing’s leaders might also find this opportunity tempting. If new technology enhances either power’s anti-access, area-denial network, then its leaders may be more confident in their ability to achieve a fait accompli attack against a neighbor and then block a US-led liberation.
These are precisely the types of shifts in the balance of power that can lead to war. As mentioned previously, the predominant scholarly theory on the causes of war—the bargaining model—maintains that imperfect information on the balance of power and the balance of resolve and credible commitment problems result in international conflict. New technology can exacerbate these causal mechanisms by increasing uncertainty about, or causing rapid shifts in, the balance of power. Indeed as noted above, new military technology and the development of new operational concepts have shifted the balance of power and resulted in military conflict throughout history.

Some may argue emerging military technology is more likely to result in a new tech arms race than in conflict. This is possible. But Moscow and Beijing may come to believe (correctly or not) that new technology provides them a usable military advantage over the United States and its Allies and partners. In so doing, they may underestimate Washington.

If Moscow or Beijing attacked a vulnerable US Ally or partner in their near abroad, therefore, there would be a risk of major war with the potential for nuclear escalation. The United States has formal treaty commitments with several frontline states as well as an ambiguous defense obligation to Taiwan. If Russia or China were to attack these states, it is likely, or at least possible, that the United States would come to the defense of the victims. While many question the wisdom or credibility of America's global commitments, it would be difficult for the United States to simply back down. Abandoning a treaty ally could cause fears that America's global commitments would unravel. Any US president, therefore, would feel great pressure to come to an Ally's defense and expel Russian or Chinese forces.

Once the United States and Russia or China are at war, there would be a risk of nuclear escalation. As noted previously, experts assess the greatest risk of nuclear war today does not come from a bolt-out-of-the-blue strike but from nuclear escalation in a regional, conventional conflict. Russian leaders may believe it is in their interest to use nuclear weapons early in a conflict with the United States and NATO. Russia possesses a large and diverse arsenal, including thousands of nonstrategic nuclear weapons, to support this nuclear strategy.

In the 2018 Nuclear Posture Review, Washington indicates it could retaliate against any Russian nuclear “de-escalation” strikes with limited nuclear strikes of its own using low-yield nuclear weapons. The purpose of US strategy is to deter Russian strikes. If deterrence fails, however, there is a clear pathway to nuclear war between the United States and Russia.
As Henry Kissinger pointed out decades ago, there is no guarantee that, once begun, a limited nuclear war stays limited.\textsuperscript{56} 
There are similar risks of nuclear escalation in the event of a US-China conflict. China has traditionally possessed a relaxed nuclear posture with a small “lean and effective” deterrent and a formal “no first use” policy. But China is relying more on its strategic forces. It is projected to double—if not triple or quadruple—the size of its nuclear arsenal in the coming decade.\textsuperscript{57}

Chinese experts have acknowledged there is a narrow range of contingencies in which China might use nuclear weapons first.\textsuperscript{58} As in the case of Russia, the US Nuclear Posture Review recognizes the possibility of limited Chinese nuclear attacks and also holds out the potential of a limited US reprisal with low-yield nuclear weapons as a deterrent.\textsuperscript{59} If the nuclear threshold is breached in a conflict between the United States and China, the risk of nuclear exchange is real.

In short, if a coming revolution in military affairs provides a real or perceived battlefield advantage for Russia or China, such a development raises the likelihood of armed aggression against US regional allies, major power war, and an increased risk of nuclear escalation.

**Implications**

Future scholarship should incorporate geopolitical conditions and the related foreign policy goals of the states in question when theorizing the effects of technology on international politics. Often scholars attempt to conceptualize the effects of weapons systems in isolation from the political context in which they are embedded.

Studies treat technology as disembodied from geopolitics and as exerting independent effects on the international system. But technology does not float freely. Technology is a tool different actors can use in different ways. Bakers and arsonists employ fire in their crafts to strikingly different ends. In the current international environment, Russia and China would tend to employ technology toward advancing revisionist aims. Technological advances in these countries are therefore much more likely to disrupt the prevailing international order and nuclear strategic stability.

This approach also suggests the potential threat new technology poses to nuclear strategic stability is more pervasive than previously understood. To undermine strategic stability, new technology need not directly impact strategic capabilities. Rather, any technology that promises to shift the local balance of power in Eastern Europe or the Indo-Pacific has the potential to threaten nuclear strategic stability.
This understanding of this issue leads to different policy prescriptions. If the technology itself is the problem, then it must be controlled and should not be allowed to spread to any states. In contrast, the framework outlined here suggests a different recommendation: preserve the prevailing balance of power in Europe and Asia. Technological change that, on balance, reinforces the prevailing international system should strengthen stability.

Leading democracies, therefore, should increase investments in emerging technology to maintain a technological edge over their adversaries. Export control and nonproliferation measures should be designed to deny emerging military technology to Russia and China. Arms control should be negotiated with the primary objective of sustaining the current international distribution of power. Making progress in these areas will be difficult. But the consequences of failure could be shifts in the international balance of power, conflict among great powers, and an increased risk of nuclear war.

Matthew Kroenig

Dr. Matthew Kroenig is a professor of government and foreign service at Georgetown University and the director of the Scowcroft Strategy Initiative at the Atlantic Council. His most recent book is The Return of Great Power Rivalry: Democracy versus Autocracy from the Ancient World to the US and China (2020).

Notes

17. Ikenberry, *After Victory*.

28. Russian experts, Track 1.5 dialogue, July 2019, Helsinki, Finland.


44. US National Air and Space Intelligence Center (NASIC), Competing in Space (Wright-Patterson AFB, OH: NASIC, December 2018), https://media.defense.gov/.


52. Fearon, “Explanations for War.”


58. Chinese experts, Track 1.5 and Track 2 dialogues, 2011–present, Washington, DC, and Beijing, China.

59. OSD, *Nuclear Posture Review*. 
Sophons, Wallfacers, Swordholders, and the Cosmic Safety Notice: Strategic Thought in Chinese Science Fiction

WENDY N. WHITMAN COBB

Science fiction, due to its ability to encourage creative thinking in environments estranged from our own, can be employed in strategy development and inspiration. Liu Cixin’s *The Three-Body Problem* trilogy focuses on galactic relations and defense in a rich and nuanced way and provides insight into Chinese thought. An analysis of four strategies—sophons, wallfacers, swordholders, and the cosmic safety notice—highlights the ways in which Chinese strategic thinking systematically differs from Western modes of thinking.

Science fiction, like other types of literature and entertainment, is used for a multitude of purposes including entertainment, explanation, exploration, and persuasion. The genre explores topics from the technical to the emotional, examining technological and scientific futures while considering the nature of humanity and relationships between ourselves and the more alien among us. Because of their (often) futuristic settings, science fiction stories are frequently framed allegorically, highlighting contemporary social and moral problems in a way that is more detached. Given the popularity of science fiction entertainment and its tendency to comment and reflect upon serious questions of the day, it is no wonder the genre has drawn the attention of and proven instructive for military thinkers and scholars alike.

Political scientists, particularly those studying international relations, have identified ways in which science fiction serves as a mirror for political reality but also helps to constitute it. As a mirror, scholars have examined the ways in which science fiction franchises like *Doctor Who*, *Star Trek*, and *Battlestar Galactica* reflect major ideas in international relations and political science.¹

In addition to simply reflecting politics, others, in a constructivist vein, argue “popular culture makes world politics what it currently is.”² Jutta Weldes writes, “popular culture, then, helps to construct the reality of world politics for elites and the public alike and, to the extent that it reproduces the content and structure of dominant foreign-policy discourses, it helps to produce consent to foreign policy and state action.”³ J. Furman
Sophons, Wallfacers, Swordholders, and the Cosmic Safety Notice

Daniel and Paul Musgrave take this one step further in proposing a model for how this influence occurs. In their theory of “synthetic experiences,” popular culture creates a sort of “pseudo-recollection” that informs and influences the way people perceive and think about the world around them. In terms of military thinking, scholars argue there is a long history of science fiction influence in America and a strong science fiction influence in late twentieth-century nuclear and military policy. As the Industrial Revolution arrived and the speed of technological innovation increased, militaristic fantasies not only predicted military developments but informed them as well. Stories from the likes of H. G. Wells influenced Robert Goddard, the father of American rocketry, and later military space policy. Wells even coined the term “atomic bomb” in his 1914 book, The World Set Free. Later in the century, science fiction author Robert Heinlein’s Solution Unsatisfactory predicted to a degree the ramifications of nuclear weapons.

This influence continues today as military analysts turn to science fiction in the development of strategic fictions. These narratives “further military funding and development”—after all, if this is the possible future, military leaders must prepare for it today. Much as the strategic narrative of the second Iraq War is an example of such a fiction, the 2001 Rumsfeld Report that predicted scenarios of a “space Pearl Harbor” can also be considered a strategic fiction. More recently, the 2015 book Ghost Fleet features a Chinese attack on the United States that begins with an assault on the International Space Station.

Another role for science fiction is as a source for creating or inspiring current strategy. This article explores possible strategies derived from a non-Western science fiction source, Chinese author Liu Cixin’s trilogy The Three-Body Problem. The strategies adopted in this trilogy serve two main functions. First, they present a nontraditional, creative source for Western strategic thinking, providing military strategists “with a mental framework that allows [them] to think outside the current paradigm.”

Second, the trilogy provides Western readers with insights into Chinese thinking. In an era of strategic competition, understanding the culture and intentions of possible adversaries is vital, particularly when the Chinese government has shown interest in leveraging culture as an element of soft power. Chinese government calculations and motivations can often be difficult to understand and analyze; accordingly, cultural artifacts such as The Three-Body Problem offer increased insights into the people and culture.
The Three-Body Problem

The first book in the series, the eponymous *The Three-Body Problem*, was published in China in 2008 and became a best seller in a country unfamiliar with reading and consuming science fiction on a broad scale.\(^{14}\) Two sequels—*The Dark Forest* and *Death’s End*—followed, and the first English translations were published in the United States in 2014.\(^{15}\) The series quickly garnered attention in the Western world, even making it onto then-US President Barack Obama’s year-end reading list. The trilogy, along with Liu’s other books and short stories, have evolved into a cultural phenomenon, stoking considerable internet fan debate and even additional books set in *The Three-Body Problem* universe. A television series is now under development at Netflix.

In addition to its pop-culture impact, *The Three-Body Problem* is significant in other respects. First, it is a major science fiction work coming from a country that has not often engaged in such literature.\(^{16}\) Though science fiction briefly flourished in the years following the Cultural Revolution, a 1983 editorial in a Communist Party newspaper claimed science fiction was “spiritual pollution,” thus stymieing the growth of the genre.\(^{17}\) Mingwei Song argues a “new wave” of Chinese science fiction authors appeared around 1989 with Liu among them.\(^{18}\) This generation of writers has focused on themes of a rising China, the myth of development, and posthuman developments, with Liu falling most squarely in the last of these.\(^{19}\) Although Liu does not focus on politics directly, the influence of Mao Zedong, the Cultural Revolution, and China’s scientific and technological capabilities are palpable in his work.\(^{20}\) Yet, even with Liu’s success in both the global and Chinese markets, there has still been some doubt about the cultural impact of science fiction in general in China.\(^{21}\)

Second and more important, Liu’s work represents a particular Chinese way of thinking about strategy. Though Liu has denied that *The Three-Body Problem* speaks directly to Chinese-American relations, to the extent it is different or emphasizes different concepts, the series provides insights into how the Chinese approach strategy.\(^{22}\) While one series of books by one author cannot possibly be representative of an entire culture, if fiction is indeed reflective of the society and culture from which it emerges and helps shape the external environment, *The Three-Body Problem* is an important object of study.

While larger elements of the story will be discussed below, in brief, *The Three-Body Problem* begins in China during the Cultural Revolution where university professors are beaten and killed for not being sufficiently communist in their thoughts and teachings. Like her father, Ye Wenjie, a young
astrophysics student, is deemed a traitor, and she is sent to a work camp in rural China. There, she sets in motion an event of galactic magnitude when she uses a radar research project to blast a message to the universe and any potential extraterrestrial life.

Eight years later, a message from the planet Trisolaris arrives, warning against further communication lest the Trisolarans attack and destroy Earth. Disillusioned by the Chinese regime, her life, and the state of humanity, Ye Wenjie ignores the warning and once again broadcasts a message inviting the invaders to do as they please.

Given the immense distances involved, the consequences of Ye Wenjie’s actions are not fully known for some decades. As the novel transitions to the present, the books detail a series of events that begins with Earth working to protect and defend against a Trisolaran invasion, explains how they come to an uneasy truce, and concludes with how the future Earth attempts to survive in what is revealed to be a very dangerous universe indeed.

It is impossible in this format to discuss every strategic concept Liu introduces. Accordingly, this article will focus on four strategies that speak most clearly to current strategic debates including the use of technology, deterrence and credibility, and signaling: (1) the use of “sophons” to prevent Earth from making technological progress that might threaten the Trisolaran invasion; (2) the “wallfacers” whose job it is to secretly devise strategy to defeat the Trisolarans; (3) the “swordholder,” the one person responsible for unleashing a cosmic attack on Trisolaris as a form of deterrence; and (4) a cosmic safety notice that clearly broadcasts a civilization’s peaceful intentions. Each strategy speaks to potential, employable real-world strategies and provides insights into Chinese cultural influences, the two purposes of this article.

**Sophons**

As the Trisolarans plot their invasion of Earth, they quickly realize Earth could soon be on the cusp of scientific and technological developments that would enable it to respond to their attack. To prevent this and to provide a means of communication between Trisolaris and its allies on Earth, Trisolaris creates sophons, advanced 11-dimensional supercomputers that when folded down into three dimensions are the size of a proton. Because of their size, Trisolaris launches them toward Earth well in advance of the larger invasion force. Once on Earth, they stop scientific advances by interfering in experiments like those being performed at the Large Hadron Collider. Additionally, because the sophons are quantum entangled with other particles that remain on Trisolaris, the sophons can
communicate with their home world in real time and facilitate cooperation between the Trisolarans and their Earth allies.

The sophons represent a strategy of technological strangulation—if Earth cannot develop advanced technology, then they cannot threaten the technologically sophisticated Trisolaris. Today, this approach is evident in regulations restricting the proliferation of certain types of technologies including missile and nuclear.

There are two differences, however, between sophons and regimes such as the Missile Technology Regime and International Traffic in Arms Regulations. One, systems preventing the proliferation of technology do not prevent the knowledge of such technology from proliferating. Indeed, using the internet or academic papers, a would-be nuclear bomb builder can quickly learn the basic construction of such a weapon. Two, nonproliferation efforts are widely known and understood; sophons act in stealth to subvert scientific discovery.

While 11-dimension supercomputers are likely not possible, sophons do suggest a particular strategy in terms of the nonproliferation of knowledge: rather than restrict the acquisition of machinery and capabilities, restrict the availability of the knowledge. Admittedly, this approach is difficult, as the failure of the United States after World War II to keep secret its nuclear know-how confirmed. But the cyber domain offers a new opportunity to attempt a similar endeavor.

Utilizing cyber tools to enact a sophon-like strategy is not simply akin to the Stuxnet attack where a destructive worm inserted into Iranian computer systems destroyed the nuclear centrifuges necessary to create weapons-grade uranium. Instead, cyber weapons could be used to infiltrate scientific and research laboratories and disrupt experiments or spoof results. In this way, cyber weapons act like the sophons to interrupt the gaining of knowledge rather than the application of that knowledge.

Of course, any strategy aimed at disrupting knowledge acquisition would be difficult to undertake for several reasons. First, a country undertaking such a strategy would need to know what knowledge it is trying to suppress. This assumes it knows what is dangerous and potentially dangerous, which in turn presumes the knowledge in question is already understood. Unfortunately, the dangerous effects of knowledge are not always known in advance. In *The Three-Body Problem*, the Trisolarans were advanced enough to know what type of research would ultimately threaten their attack.

Second, civilian-based research is difficult to suppress. While the ultimate goal of scientists is different (e.g., money, publication of results), most goals are premised on sharing acquired knowledge with others. Pre-
venting such sharing might be easier to do with government-sponsored research, but even then, suppression is not guaranteed.²³

Third, scientific research is often conducted by multiple teams at multiple locations. For a sophon strategy to be successful, all researchers undertaking a particular line of inquiry would need to be monitored and suppressed, not just those in a target country. Because of their unique, quantum nature, sophons were able to do this on a large scale. Using cyber technology, this may not be impossible, but it may not be probable either.

Finally, this strategy certainly conflicts with the openness of a democratic society. While countries around the world have often committed acts that conflict with their core national values in the name of national security, a sophon strategy, if revealed, could damage the international image and consequent soft power of the United States. Moreover, this damage would almost certainly not be limited to those outside of the United States but extend to US citizens who value science, understanding, and the discovery of knowledge.

Wallfacers

While sophons might not be possible, wallfacers hew a bit closer to reality. The omnipresent sophons allow the Trisolarans to know exactly what Earth’s leaders are doing to prepare for their invasion, giving them a strategic advantage. To overcome this, the United Nations creates the Planetary Defense Council which, in turn, creates the position of wallfacer, an individual responsible for developing a strategy to counter the Trisolarans.

Liu writes that the name “wallfacer” is meant to evoke “that ancient Eastern name for meditators [which] mirrors the unique characteristics of their work,” as they present a false narrative to the public to conceal the true strategy from the sophons.²⁴ Wallfacers are given any resource or power they wish to enact their strategy, but they are not to divulge what that strategy is. In this way, their strategy can be kept secret from the Trisolarans.

Though a real-world country might not face such an existential threat, intelligence gathering methods keep a state’s strategy, or at least the parts that should remain opaque, secret. In a world where cyber espionage and space-based data gathering make keeping state secrets hard, a wallfacer strategy might be a viable option. In this sense, a wallfacer might be charged with protecting a country’s national security, empowered with creating and enacting strategies that promote a country’s survival. They would not have to tell anyone what the ultimate strategy consists of but would have to direct the allocation of resources to enact it.
The wallfacer strategy encompasses a number of advantages. For one, it retains the element of surprise. And even if a wallfacer does discuss a particular strategy, there is no way to know whether that strategy is the true one—disclosing a strategy might be a feint in one direction preserving freedom of maneuver in another. Indeed, one of the wallfacers in *The Dark Forest* encounters just this problem. Luo Ji, an otherwise unremarkable Chinese sociologist, is selected as a wallfacer for unknown reasons. Not understanding why he was chosen, he refuses to work, lives a profli- gate and expensive lifestyle, and uses the resources offered to him to build a luxurious home and find a wife. When he explicitly tells officials that this is not part of his strategy, they refuse to believe him. Ultimately, they become frustrated with his lack of work and put his wife and child into hibernation as a means of forcing him to work.

A wallfacer strategy also centralizes authority and planning in one individual with explicit responsibility. One of the critiques of the current American national security establishment is that either there is no one in charge of thinking holistically about national strategy or that very few are. Further, senior leadership changes from election to election—a president may be in office between four and eight years with no guarantee a successor will continue a particular strategy. This situation yields either a short-term strategy that can be executed or a long-term strategy that might not come to fruition. A wallfacer whose term of office is unlimited can rise above short-term electoral politics to enact long-term strategies that might better advance the position of a given country.

The dangers of secrecy, however, outweigh the advantages. Given that wallfacers are endowed with whatever resources they could possibly ask for without question, accountability is lacking, a severe disadvantage in democratic societies. One wallfacer attempts to design a method in which Mercury crashes into the sun, causing it to explode and destroy the entire solar system. Yet another devises an elaborate plot to be seen as cooperating with the Trisolarans only to double-cross them. While a form of ac- countability was ultimately imposed on Luo Ji, a wallfacer strategy lacking such a mechanism is ultimately risky unless the individual merits high levels of trust.

The individual becomes the second problem—who exactly could be made a wallfacer? How would that individual be chosen? What characteristics would such a person need to have? In the United States, citizens presume the president is ultimately in charge of national security strategy; would it be constitutional to “subcontract” such a responsibility? Would the wallfacer need to change after each presidential administration? Would
the wallfacer be able to enact domestic changes to ensure national security? How might a wallfacer be removed from their position? Clearly, in an open, democratic society, a wallfacer strategy would face almost insurmountable barriers.

**Swordholders**

Luo Ji, though initially perturbed by his selection as a wallfacer and careless in his approach, does eventually discover a possible mechanism by which Trisolaris could be thwarted. In the second book, it emerges that Luo Ji had a brief encounter with Ye Wenjie in his younger years in which she imparted some general principles of a cosmic sociology. “First: Survival is the primary need of civilization. Second: Civilization continuously grows and expands, but the total matter in the universe remains constant.”

Additionally, Ye Wenjie tells Luo Ji of two other strategic concepts he will need: chains of suspicion and the technological explosion. In describing a chain of suspicion, Liu writes that even if two galactic civilizations believe each other to be benevolent, it is impossible to “know what you think about what I think about what you’re thinking about me.” Technological explosions are defined as the capability of a civilization to undergo significant technological change when faced with an immediate threat.

From this, Luo Ji finally discovers the universe is essentially a “dark forest” and describes realism on a universal scale. If one civilization encounters another and communication is possible, then the technology of both is on a scale where one can threaten the other and vice versa. Because of the tyranny of distance and time on a universal scale, there is no way to allay mutual fears and no way to know the other civilization’s ultimate goals and aims. Not knowing the other’s intent, it is in the interest of any given civilization to immediately attack and destroy the other lest they give away their own civilization’s position to others in the dark forest. Thus, all occupants of the dark forest have an inherent interest in remaining undetected.

Based on this theory, Luo Ji recognizes one way to defeat Trisolaris is to reveal its position in the universe to the other members of the forest. To test his theory, Luo Ji utilizes his wallfacer power to cast a magic spell: using the sun as an amplifier like Ye Wenjie did, he broadcasts the location of an innocuous star. If he is correct, that star will be destroyed several hundred years in the future.

While the Trisolarans clearly understand the intent of the spell and place a hold on their invasion force, Luo Ji uses deception on Earth to hide the true intent of his strategy. Understanding that his spell will take
time, Luo Ji chooses to hibernate for 200 years. When he emerges, Earth technology has greatly advanced despite the sophon lock. Defensive forces from Earth have amassed a fleet of spaceships and are preparing to intercept the inbound Trisolaran fleet.

While Earth’s fleet ultimately fails to slow the advancing Trisolaran fleet, it is soon revealed that Luo Ji’s magic spell indeed resulted in the destruction of the star. To overcome this new threat, the Trisolarans respond by preventing Luo Ji from transmitting their planet’s location and continue their advance. In response, Luo Ji and Earth’s defense organizations plot to use nuclear weapons to create vast fields of dust and debris allowing Earth to monitor the progress of the enemy fleet. Moreover, Luo Ji realizes he can use the bombs, if detonated, to broadcast Trisolaris’s location. Trisolaris and Earth again find themselves mutually deterred: if Trisolaris attacks, Earth will broadcast its location. Alternatively, if Earth attacks Trisolaris, the Trisolarans’ advanced technology not only could broadcast Earth’s location but destroy the Earth on its own.

As a result of the course of events, the system of precisely placed nuclear bombs remains under Luo Ji’s sole control. Abandoned by his wife and daughter and chastened by his experiences, he assumes the position of swordholder—the one individual empowered to detonate the nuclear bombs at the first indication of a Trisolaran attack. Further, Luo Ji, the swordholder, does not have the luxury of time in responding to any Trisolaran action as their capabilities allow for the immediate destruction of the Earth and its inhabitants.

**Credibility**

Of the traditional concepts of deterrence, nuclear strategy, and credibility enshrined in the idea of a swordholder, credibility is the most developed. The system of deterrence described in the series is based on the belief that the other side is capable of an attack. While Trisolaris demonstrates this in action—the destruction of the Earth’s fleet—Earth’s credibility lies solely in Trisolaris’s belief that Luo Ji would initiate the nuclear detonations. Because Trisolaris knows Luo Ji grasps the dark-forest nature of the universe, they believe he will initiate the broadcast, via the detonations, if Earth is threatened.

As is often the case, credibility is not easily transmitted from one officeholder to the next. As it comes time for Luo Ji to retire from the position of swordholder, a new candidate is chosen, Cheng Xin. Through a series of events that also involve enchanted time-lapse sleep, Cheng Xin soon becomes a leading candidate for the next swordholder because of qualities
that make her the antithesis of Luo Ji. She is feminine; she represents an earlier period; and she reveals herself to be someone who is softer and lacks Luo Ji’s jadedness. Yet those characteristics that make people believe she will be a kinder, gentler alternative to the warrior-monk Luo Ji are precisely those which make her a less credible threat to Trisolaris.

As Cheng Xin enters the swordholder bunker and assumes the mantle from Luo Ji, Trisolaris initiates a devastating attack on Earth and its population. This attack reveals Trisolaris did not judge Cheng Xin to be a credible swordholder—they did not believe she would initiate the broadcast, which is exactly the case. In the moment of ultimate threat, Cheng Xin declined to initiate the nuclear broadcast that would not only destroy Trisolaris but Earth as well.

Cheng Xin’s failure demonstrates credibility is not easily conferred and is not conferred on an individual simply because they possess a position of power. In contrasting these two swordholders, Liu suggests credibility is to be judged based on a depth of true understanding, though how that understanding is to be acquired is entirely unclear.

Examining the wallfacer and swordholder strategies together, one element they hold in common is the centrality of individual power. While this might seem anathema to those in the West with a strong democratic heritage, Chinese notions of central leadership and authority are consistent with these approaches. While an entire machinery of people and effort support both wallfacers and swordholders, the ultimate power and authority is vested in one individual, reflective of Chinese cultural notions of governance.

And while these approaches are attractive for several reasons (e.g., attribution of responsibility, clear decision making, ease of decision making, lack of need for compromise), Liu appears sharply critical based on his depictions of both wallfacers and swordholders. In the case of the wallfacers, three of the four are ultimately heralded as turncoats and war criminals for the strategies they propose that would kill many, if not all, people.

In the case of the swordholder, Liu’s narrative suggests this might be a productive strategy if credibility can be established and maintained. In the absence of credibility, it too fails because of the centrality of one person. Research on the credibility of deterrent threats in recent years has similarly identified this problem. Thus while the swordholder and wallfacer strategies appear to be superior solutions to the problems Earth faces, Liu is certainly aware and ultimately wary of the points of failure that exist.
Personal Cost

Luo Ji’s story is also a warning of how the individual chosen for such a task may indeed suffer. The Earth prefers to center its existential decision-making power in one individual rather than instill this power in a distributed command and control system. In his time first as a wallfacer and then as swordholder, Luo Ji is worshiped alternately as a savior and a monster. Luo Ji suffers from the burden: initially his wife and child are taken away, and when they ultimately leave on their own volition, he suffers from alcoholism and depression.

When he finally passes on the duty of the swordholder, Luo Ji lives in a monk-like state in isolation in a bunker deep underground. By the end of the trilogy, he is the lone human on Pluto and the last caretaker of Earth’s most precious treasures. When considering strategies that place inordinate power in the hands of a single individual, societies must contemplate the question of extreme personal cost.

Cosmic Safety Notice

Following their attack on Earth, the Trisolarans forcibly move all remaining people to Australia to prepare for the arrival of the main attack force. Despite the severe decimation of Earth’s population, the continent cannot support the remaining people. Consequently, the move to Australia also becomes a means of culling the population to a more manageable size. Owing to the escape of two of Earth’s war ships from Trisolaris’s earlier attack, it is soon discovered that Trisolaris itself is destroyed by another civilization because of a location broadcast from one of the escaped ships.

Knowing it is only a matter of time before Earth, too, is caught in the galactic crosshairs, the Trisolarans abandon their invasion and seek to impart the notion of a “cosmic safety notice” to those remaining on earth. This notice is intended to signal peaceful intentions to the rest of the universe. Once decoded, the message suggests one way to do this is to slow down the speed of light via a series of black holes around the solar system so that no light escapes.

In this way, not only does a civilization declare its intentions, but it closes its cosmic neighborhood off to escape and to other potential civilizations. This approach is reflective of proposals involving “costly signaling.” In this instance, foreclosing all future opportunities to engage with the universe is not only the cost but ensures Earth’s future protection from galactic threats.
The notion of a cosmic safety notice raises interesting parallels to strategy today: How does a state signal peaceful intentions despite actions that seem potentially aggressive? While this happens in all domains, consider the use of near Earth space in recent years. With increasing movement towards the overt weaponization of space, Russia, China, and the United States have all said they are taking actions they believe are required to protect and defend their space-based assets and that their actions are not offensive in nature.

But because defensive weapons can also be used offensively and space technology is inherently dual-use, it is difficult for each country to trust the other’s statements particularly with more aggressively hostile rhetoric being used on all sides. In such a situation, how might a country legitimately and credibly signal their intentions?

Liu’s answer suggests signaling cannot necessarily be undertaken without also hobbling one’s own capabilities. While the solar system, subsequently destroyed by another galactic civilization, is never able to implement the cosmic safety notice, if it had been, humans would have been doomed to remain within the solar system, never able to explore outside its boundaries.

This is not, however, the same as unilateral disarmament. If Russia, China, or the United States were to forego weapons in space, that would leave that respective country open to attack by another actor. With the cosmic safety notice, the method that declares peace is also the method that protects it. For this type of strategy to work in the real world, whatever action the message consists of must both defang and defend at the same time.

**Conclusion: Science Fiction Strategies**

Leaning too far into science fiction’s influence can yield certain dangers. Due to its relative conceptual accessibility, it appears to be somewhat realistic or achievable. More often than not, however, science fiction presents a far-flung future that is decades if not centuries ahead of us (if indeed it ever comes to pass). Consequently, its technological optimism may unduly influence our thinking about contemporary international relations challenges. Instead of focusing on the realities of strategic challenges, our attention is drawn more to the problems we might face in the far future. Further, if future scenarios painted by science fiction never come to pass, the strategies writers describe in them may never be feasible or possible.

Still, and with these limitations in mind, science fiction can be a source of inspiration in overcoming strategic challenges. After years immersed in
traditional doctrine and strategy, military practitioners may find thinking imaginatively a difficult task. Exercising that imagination through regular reading of (science) fiction helps keep the mind limber and able to make conceptual leaps and bounds not otherwise readily apparent.

Science fiction can create new linkages and pathways and new ways of thinking about common problems, as evidenced by The Three-Body Problem trilogy. In thinking about how global international relations may be played out on a galactic scale, Liu provides the reader a sense of estrangement from the real world that allows for a clearer-eyed assessment of that world. Estrangement is often used in science fiction to allow writers to explore topics relevant to their audience but from a place apart. The Three-Body Problem embodies this principle—in providing another arena in which to examine international relations theory at play, we can stand apart and above the situation and see it a bit clearer than we often do from within it.

The strategies outlined here—sophons, wallfacers, swordholders, and cosmic safety notices—though perhaps not entirely applicable in an American context—can still cause us to think more on what is possible or how these may be adapted to our own concerns. Further, the series is valuable in its role as Chinese science fiction. Improved understanding of Chinese perspectives and thinking is important as policymakers seek to counter Beijing’s more bellicose rhetoric and actions—culture and literature are important components to deepening this knowledge. Certainly, the individualistic and centralized strategies depicted in the trilogy are not only keys to understanding Chinese thinking but also contain implicit critiques of them from a Chinese author.

Wendy N. Whitman Cobb

Dr. Whitman Cobb is associate professor of strategy and security studies at the School of Advanced Air and Space Studies. Her most recent book is Privatizing Peace: How Commerce Can Reduce Conflict in Space (2020).

Notes

Sophons, Wallfacers, Swordholders, and the Cosmic Safety Notice


Six Steps to the Effective Use of Airpower

On “The Drawdown Asymmetry: Why Ground Forces Will Depart Iraq but Air Forces Will Stay”

JOSEPH B. PIROCH
DANIEL A. CONNELLY

“War can . . . engrave lessons like no other human endeavor.”
—Thomas Hughes, Over Lord

Then-Lieutenant Colonel Clinton S. Hinote’s 2008 analysis of the Iraq drawdown and the continued role of airpower in that conflict serves as a foundation for six steps to the effective use of airpower today.

In Summer 2008, Strategic Studies Quarterly published an article by then-Lieutenant Colonel Clinton S. Hinote that anticipated six steps of airpower. Written shortly before the withdrawal of forces from Iraq began, the article defended the drawdown but argued a large contingent of US Air Force personnel should remain in place to support the coalition with airpower and protect Iraqi airspace.

Hinote’s analysis reflected the operational experience of a combat pilot and former combined air operations center war planner. Moreover, it demonstrated wisdom and sound reasoning in underscoring warfare’s moral dimension, an imperative consideration for all military engagements. Thirteen years later, his analysis remains compelling and supports the six-step framework described below.

Airpower

From the time airplanes first took to the skies, war theorists and military commanders have marveled at their battlefield potential. Early on, military thinkers and leaders saw how the far-ranging mobility and tremendous speed of these new warfighting machines gave them “complete freedom of action and direction,” allowing them to shift the point of attack rapidly and at will.

Engagements in World War II, the Korean War, and the Vietnam War confirmed airpower is indeed a singular and formidable capability in the hands of a battle commander. Notably, however, these campaigns also demonstrated that effective and ethical use of airpower depends on several prerequisites. Airpower is but a tool; it does not within itself contain an
explanation of how to employ it properly in a war effort. Using it well and avoiding its misuse requires military leaders to follow several foundational principles, core tenets of airpower that transcend the platform itself. With Hinote’s analysis in mind, this article identifies six principles for today, briefly illustrating them with examples from World War II to Vietnam.

**Carefully Establish Military Objectives**

The effective use of airpower begins with the communication of clear military objectives. These objectives serve as the foundation of the wartime effort, driving strategic decisions and rules of engagement; an air operation without them will suffer from “all Mach and no direction.” Military objectives include the positive—what the military intends to achieve—and negative—outcomes the military hopes to avoid. Examples include US President Lyndon B. Johnson’s insistence during the Vietnam War that South Vietnam remain free from communism (positive), and that China not enter the war (negative).

These objectives were not clearly set forth during Vietnam, however. As Mark Clodfelter explains, discordant objectives regularly emerged from the president’s weekly meetings with Secretary of Defense Robert S. McNamara, and military leaders interpreted them in contradictory ways, muddling the war effort.4

Developing good military objectives is hard for several reasons. First, they may be unrealistic—a point that may escape leaders in times of crises. Historians have questioned whether, for instance, it was possible for the US military to win the war in Vietnam without striking the North, even at the risk of Chinese intervention. Second, military objectives may be unachievable due to ancillary factors such as political battles, budgetary restrictions, or waning domestic or international support, all of which complicated Johnson’s decisions during the Vietnam conflict.

Third, military objectives may be constrained by other military conflicts around the globe. Fourth, they are nearly always complicated, as they confront the complexities and friction of war and attempt to predict an unknown future. The United States successfully navigated such complexities in the Korean War when it fought a limited war for the first time and found a way to achieve victory without using nuclear weapons or getting into a larger war with China or the USSR.5 Fifth and finally, military objectives may be disjointed from political ends. Without strong political-military integration, the military may spend blood and treasure achieving battlefield victories that fail to accomplish goals of the state.6
Hinote consistently referred to these elements in his article. First, he raised questions about the US military strategy in Iraq, noting the surge (alternatively viewed as the ways or means) was not clearly tied to objectives (the ends)—that is to say, political–military integration was incomplete or missing in the Iraq campaign. To this point, he cited General Anthony Zinni, who had complained there was, in fact, no strategy at all.7 Other issues were also limiting the military’s ability to succeed, according to Hinote. Many years of fighting had exhausted soldiers and strained families as well as the air fleet itself—items that retrospectively call into question whether military objectives were achievable. Additionally, the large commitment of military forces to the Middle East had degraded America’s capacity to project power in other parts of the world—in other words, the Iraq campaign was constraining.

Hinote believed a reorientation of US military strategy in Iraq was needed. But he also believed a complete US withdrawal would shake the country, leading to a failed state that terrorists would exploit. He felt the troop level should be militarily sustainable abroad and politically acceptable at home (i.e., both realistic and achievable).8 Following the surge, the US military did draw down its forces in Iraq gradually over the next four years, ending its presence in 2012. The country still was unstable at that point, leading to the rise of the Islamic State of Iraq and the Levant (ISIL). The United States defeated ISIL in Operation Inherent Resolve using small ground force teams supported by powerful airstrikes—a balanced approach concordant with Hinote’s ideas.9

**Determine the Role of Airpower**

Once leaders have established and clearly conveyed their objectives, they must determine airpower’s role in achieving them. The USAF does not by itself accomplish all objectives in a military conflict and win the war but works in conjunction with other services, agencies, and partners in Joint warfare, usually operating as a combined force that leverages a wide variety of capabilities. As Hinote knew, fighting Jointly is never easy. In the Korean War, for example, interservice disputes and a lack of Joint doctrine created substantial friction, although in Vietnam, new plans such as the Concept for Improved Joint Air-Ground Coordination advanced Joint warfare.10

Although airpower’s role may vary extensively from campaign to campaign, its signature function in any engagement is to maintain air superiority. Military theorists J. C. Slessor and Giulio Douhet emphasized the need to command the air in conflict, and World War II general officers,
including General George Kenney, reiterated the importance of air superiority in all military operations, as did commanders in Korea and Vietnam. Without command of the air, armies and navies always face a far higher risk of defeat.\textsuperscript{11}

Fundamentally, air forces support a war effort by providing strategic airpower, tactical airpower, air support, or a combination thereof. Strategic airpower places air commanders in a supported role, which they typically execute in bombing sorties that strike the enemy’s warfighting capacity.

While various approaches to bombing have been proposed, the US strategy of precision bombing employed in World War II, Korea, and Vietnam was highly effective. The targets of precision strikes may be the “industrial fabric” of a nation such as petroleum depots, power plants, or factories for military parts; lines of communication including conduits for transportation, supply, and information; or forces themselves consisting of command-and-control nodes, military bases, or soldiers on the ground.

Tami Davis Biddle argues that in the World War II European theater, although bombing methods were still being refined, the United States had such success with precision bombing that the Nazis were compelled to pull forces from the front, disperse their factories, and invest in defensive weaponry.\textsuperscript{12} This approach also yielded success in the Korean War, with attacks on North Korean industry, and in the Vietnam War, especially during the devastating Linebacker I and II campaigns that forced peace negotiations.\textsuperscript{13}

Tactical airpower puts air capabilities in a supporting role and generally focuses on providing close air support and air interdiction of enemy forces, which may consist of standard military units, insurgents, or guerrilla fighters. Targeting plans should be developed in close collaboration with ground force commanders.

To the detriment of operations and responsible use of resources, the Korean and Vietnam Wars featured discord between US Army and Marine Corps leaders and the Air Force as they debated whether strategic bombing or close air support should be the primary air mission. This discord is an issue that clear objectives from Washington could have perhaps solved. If forces cannot obtain the air support they need, they may begin developing their own capabilities—the US Army’s equipping of helicopters with close air support armaments in Vietnam being but one example—often resulting in unnecessary redundancies.\textsuperscript{14}

Air support includes critical functions such as transportation and airlift, intelligence collection, and communications facilitation. In World War II, the Germans conducted impressive airlift operations during the Blitz-
krieg, and the United States was similarly successful in airlifting supplies to the Chinese to keep them in the war.

Shortly after World War II, the Berlin Airlift of 1948–49 further highlighted this role of airpower. Air support also extends to intelligence gathering, a fundamental function that ensures commanders have situational awareness of the battlefield and real-time tactical reporting to aid their operational planning and execution. Similarly, air support ensures dependable communications between units, allowing them to be in lockstep even when the operational tempo is fast. Throughout the Vietnam War, intelligence, reconnaissance, and communications support were highly successful. Each of these air support functions has vastly different requirements that commanders must prioritize and synchronize with due attention.\textsuperscript{15}

Hinote considered these realities in his analysis and advocated supporting the remaining troops with robust airpower. He correctly stated that just as the Air Force cannot succeed without firepower on the ground, armies cannot succeed without help from the air. The Air Force needed to, therefore, continue Joint air operations in Iraq: attack the enemy, prevent it from massing forces, and protect military bases (all elements of tactical airpower). At the same time, the Air Force needed to provide airlift, intelligence, search and rescue, and communications services—all elements of air support.

Throughout that fight, the Air Force needed to maintain air superiority over the country, giving ground forces top cover and deterring conventional attacks on Iraq by adversaries.\textsuperscript{16}

**Recognize Capabilities and Limitations of Airpower**

As planners determine airpower’s precise role in the war, they must consider its capabilities and limitations. Their assessment begins with choosing the right platform and weaponry. Well-defined objectives help commanders make good decisions in this regard, although questions forever remain about what capabilities are best since each aircraft has distinct strengths and weaknesses. Strategic platforms often emphasize firepower and defensive armaments, while tactical platforms provide faster response and longer loiter times. Regardless, selecting the right aircraft for complex battle campaigns is never easy.

The US Air Force has occasionally tried to use a single airframe to fulfill multiple roles—utilizing strategic bombers for ground support during World War II or modifying fighters for strategic bombing in Vietnam—but these dual-use aircraft often are jack-of-all-trades, master-of-none platforms that do no job overly well.\textsuperscript{17}
In addition to selecting their aircraft, planners must consider tactics, which often morph over the course of a war as pilots attempt to take advantage of new capabilities, correct flying deficiencies, and exploit enemy vulnerabilities. World War II aviator Pete Quesada’s use of dive-bombing, radar, and enhanced radio communications; Kenney’s development of parafrag and skip-bombing; and fighter pilot John Boyd’s creation of ingenious air maneuvers in the 1950s are examples of ingenuity in action.\(^{18}\)

Airpower’s effectiveness can be hampered by the absence of clear command and control processes which can be difficult to establish in Joint and combined warfare. The institution of a single air commander for theater operations in the Korean War was an illustration of leaders overcoming this shortfall.\(^{19}\)

Finally, campaign planners look for stable doctrine that captures the most critical principles of warfare and encourages unity of effort within the force. Good doctrine on precision bombing, for instance, emerged during Vietnam, providing theater commanders with a common standard for the first time.\(^{20}\)

Airpower has other inherent limitations. Pilots and crew members implicitly understand restrictions imposed by weather and terrain that can make mission accomplishment difficult or impossible. Tyranny of distance may geographically separate commanders from the battlefield, making operational decisions harder and slower. Intelligence is always incomplete and sometimes wrong, complicating planning and operations. Technology is persistently advancing, requiring aircraft and weaponry upgrades to stay ahead of the adversary. Training is frequently abridged or skipped during crises, expanding the number of untrained or undertrained airmen in the force.

On the adversary’s side, defenses are constantly adapting, forcing planners to adjust for success. Simply stated, decision makers must resist the temptation to focus solely on airpower’s impressive capabilities without giving due consideration to its extensive limitations.\(^{21}\) Hinote was keenly aware of the need to consider both the capabilities and limitations of airpower. He discussed people, equipment, airframes, and weapon systems, examining opportunities and risks. And he acknowledged commanders would need to overcome tyranny of distance, battlespace complexity, logistical barriers, and other challenges to succeed in Iraq.\(^{22}\)

**Adopt a Whole-Force Approach**

Having prepared an operational plan that articulates airpower’s role, leaders must then empower its enabling functions. General Henry “Hap”
Arnold may have expressed this idea best when he proclaimed air wars are won through “total aviation activity,” not just by bombers and fighter jets. The total air effort includes functions across the force that enable airpower: logistics, acquisitions, maintenance, training, and myriad other activities.23

For this reason, Arnold restructured the air forces during World War II, evoking Napoleon’s maxim that organizing the military properly is the most critical step in attaining victory. Kenney, too, understood this reality; he oversaw bombing sorties but also emphasized intelligence, maintenance, airlift, and other functions. Commanders during the Korean and Vietnam Wars extended the whole-force approach to the entire war machine and codified it in doctrine. While results were initially mixed, these leaders did lay the groundwork for a post-Vietnam focus on Joint doctrine that ultimately improved warfighting.24

As a pilot and planner, Hinote understood the importance of the whole force in executing air operations effectively. He discussed at length the logistics, ground support, systems, and intelligence activities that must be integrated across the campaign for the commander to succeed.25

Adapt and Overcome

As war proceeds, leaders must be prepared to adjust their operational plans to accommodate changes in the strategic environment and solve new problems that arise. Battle conditions are unpredictable—a phenomenon Clausewitz alluded to as “a fog of greater or lesser uncertainty”—and no objectives or plans survive unscathed once fighting begins.26 Airmen should be poised to adapt and ready to experiment, seeing each crisis as an opportunity rather than a threat. Moreover, military leaders should prepare their organizations to embrace this eventuality.

Arnold, Quesada, and Kenney embodied these principles in the opening decades of airpower, truly a time of trial and error. Arnold was particularly keen on adaptation; he challenged his Airmen to innovate, and he forged relationships with industry, putting the greatest minds to work on airpower’s hardest problems. By contrast, Nazi Germany was rather poor at adaptation; historians have suggested its overbearing, top-down control and aversion to trusting field commanders were key contributors in the Axis loss of the war.

After World War II, US modernization continued with the enhancement of night-mission capabilities and airburst weapons in the Korean War and the introduction of new gunships, laser-guided bombs, and better radar in Vietnam.27 New capabilities and novel methods are very often the keys to victory. Hinote discussed adopting adaptive and innovative
methods to transition the Air Force from an active fighter force into a smaller contingent charged with partnering and overwatch roles. He also considered its ability to support in-country humanitarian missions and messaging campaigns, all the while managing deployments effectively.  

**Act Ethically Always**

For modern just war thinkers of high credibility such as James Turner Johnson, the demands of ethics and prudence align seamlessly. This alignment can be seen in his call for the maintenance of noncombatant immunity, which in common speech is a special admonishment to the military to protect and defend all innocent life always.  

As leaders execute an air campaign, they must ensure ethics guide their decisions, even, or perhaps especially, in the heat of battle. They have not always done so. Prior to World War II, for example, the United States and its Allies unanimously adopted the Hague Rules, preventing the indiscriminate killing of civilians, a practice they called the “hallmark of barbarism.” Yet during the war as commanders encountered difficulties in bombing campaigns, they gradually disregarded this precept and began to engage in relentless area bombing in Europe and firebombing in Japan, methods that indiscriminately killed thousands. During the Korean War, the US Air Force continued to utilize firebombing as a tactic.  

Some contend the World War II raids were essential, a driving factor in the Axis surrender, but Richard Overy has shown that in Europe they did not crush the enemy’s resolve as intended, and Biddle has argued that the impact of mass bombing on morale is simply not measurable. Regardless, even if these methods did work, efficacy should never trump righteousness—the ends can never justify the means, and leaders are always called to a higher standard.  

Importantly, Hinote also made an ethical claim: the United States was morally obligated to stay in Iraq until the government could stand on its own. To withdraw fully and leave 28 million Iraqis to face chaos and turmoil would have been wrong, he believed. Hinote agreed with diplomat James Dobbins who said the United States had assumed responsibility for Iraq when it invaded the country and removed Saddam Hussein from power—in short, we broke Iraq, and we had the obligation to fix it.  

On ethics, the article could have made an even stronger argument by appealing to the just war tradition, particularly its discussion of *jus post bellum* (“justice after war”). Over the centuries, just war thinkers including Saint Augustine, Thomas Moore, Martin Luther, Francisco de Vitoria,
Francisco Suárez and others noted the importance of not harming the innocent and being merciful to the vanquished.\textsuperscript{33}

The article also may have profited from a discussion on right action, which Aristotle and Thomas Aquinas said requires the moral virtue of practical wisdom, or prudence.\textsuperscript{34} Answers to questions such as whether to stay in Iraq and for how long were ultimately ethical prudential judgments, needing to respect the common good of both the US public and the Iraqi people.

Personal integrity and the pursuit of virtue are the core of ethics for moral action begins with the self. The Air Force embodies this principle in one of its core values: Integrity First. In a timeless but oft-neglected passage of \textit{On War}, Carl von Clausewitz warns commanders to develop their sense of war’s ethical dimension in his remarks on the nobility of moral considerations on the battlefield.

In their efforts to act morally, commanders must be on their guard: war is brutal and violent, and as Conrad Crane has noted, the constant presence of death during war can harm decision-making on ethical matters.\textsuperscript{35} At a minimum, Airmen should be armed with an understanding of the just war tradition, a time-tested guide rooted in centuries of sound philosophical thought that sets forth the ethical principles of warfighting—before, during, and after conflict.

Just war doctrine is an invaluable resource to help guide commanders when decisions are less clear. Officially enshrined in the \textit{Department of Defense Law of War Manual}, this tradition explains we are moral creatures before we initiate combat and tells what obligations we must not ignore, lest we become the evil we are fighting.\textsuperscript{36}

**Conclusion: Airpower Cannot Determine Its Own Utility**

Airpower is a potent asset in the military arsenal—“the offensive weapon par excellence”—but whether it succeeds or fails in achieving wartime objectives depends on how leaders employ it.\textsuperscript{37} Clausewitz stated military leaders need to lean on warfighting principles in times of conflict, and the six steps above, drawn from the annals of history, are proven tenets that should serve commanders well.\textsuperscript{38} Building on Hinote’s analysis, these ideas should both embolden and caution us to remember that the correct understanding of airpower stems not from its platforms and its capabilities alone, but from the deepest possible appreciation of the purpose of war and our moral obligations.
Joseph B. Piroch
Mr. Joseph Piroch is an Air Force Fellow at RAND Corporation, Santa Monica, California.

Daniel A. Connelly
Dr. Daniel Connelly is an assistant professor in the Department of Leadership at Air Command and Staff College, Maxwell AFB, Alabama.

Notes
14. Crane, Airpower Strategy in Korea, 7, 62, 176; Clodfelter, Airpower, 2–3, 35; James S. Corum and Wray R. Johnson, Airpower in Small Wars: Fighting Insurgents and Terror-

15. Hughes, Over Lord, 102; Griffith, MacArthur’s Airmen, 245; and Clodfelter, Airpower, 8.


27. Daso, Hap Arnold, 106, 125, 130, 132, 161, 186–87; Hughes, Over Lord, 18, 128–33, 183–88; Griffith, MacArthur’s Airmen, 65; and Crane, Airpower Strategy in Korea, 42, 68.


30. Overy, Bombers and the Bombed, 35.

31. Crane, Strategy in World War II, 184; Crane, Airpower Strategy in Korea, 46, 65; Biddle, Rhetoric and Reality, 64, 80, 130, 158, 278, 290; and Overy, Bombers and the Bombed, 41–48, 168–69, 209.


35. Crane, Strategy in World War II, 86.

Cultivating Future Airpower Strategists

On “Developing Twenty-First-Century Airpower Strategists”

JOHN G. TERINO JR.

In 2008, Major General R. Michael Worden forecast specific challenges for airpower strategists including emerging technology, transnational terrorist organizations, an explosion of information power, budgets, and resourcing. His predictions have borne out in what the Air Force faces today, and Air University is responding, providing the next generation of airpower strategists.

Connecting operations to strategy seems easy, but the last 20 years of conflict for the United States shows that is not necessarily the case. Developing effective operations to achieve national strategic outcomes in an era of renewed and heightened strategic competition, especially in the air, space, and cyber domains, will be even more challenging than what was attempted over the past two decades. But the process for developing airpower strategists has proven sound for over a century; continuing it should serve the US Air Force and US Space Force well for decades to come.

Major General R. Michael Worden’s 2008 article in Strategic Studies Quarterly forecast many challenges facing future American airpower strategists because of technology, transnational terror organizations, rapid and unfettered information power, and fiscal and recapitalization challenges for the force. Many of the issues and the rate of change he postulated are spot on and, more importantly, his prescription for human capitalization and the development of new generations of airpower strategists is still quite sound.

Worden channeled the wisdom of many icons of military thought—Sun Tzu, Carl von Clausewitz, and John Warden. These thinkers advocated the serious study of military, political, and cultural history; honing and sharpening one’s capacity for critical thought and communication through written expression; and developing a sound, professional tactical acumen through training and field experience.

Worden specifically mentioned familiarity with the technologies that undergird current capabilities and those expected to transform air, space, and cyber power. He further refined the traditionally accepted notion of
developing strategists by adding that the Air Force must inculcate a “winner’s creed” consisting of a commitment to innovation, integration, and incorporation of the results of the application of this process regularly and rapidly.²

From the vantage point of more than a dozen years later, Worden's analysis and recommendations are still useful. When it comes to preparing for the future, there is no substitute for a broad education that instills habits of mind and patterns of inquiry coupled with rational and thoughtful analysis. For the US Air Force and the US Space Force, there have been several changes in organization and focus that reflect the wisdom of Worden’s reasoning. Many of these changes can be found throughout the services, particularly at the intellectual and leadership center of the Air Force, Air University. This is only fitting because Worden's article centered on developing and teaching the study of war, history, politics, and culture as the bedrock of cultivating future airpower strategists.

There is nothing like failure to inspire introspection and change. Much of Clausewitz’s motivation for writing *On War* stemmed from Prussian defeat at the hands of Napoleon. The transformation of the US military in the late twentieth century was animated by failure stemming from the Vietnam War. And today, it is apparent that elements of the push for rapid change in the US military are driven by a response to strategic competitors and frustration with the outcome of the campaigns of the last two decades. Worden’s article highlights rapid changes in technology, awareness of changes in the domains of conflict, and the need for introspection and hard thinking based on intellectual development to respond to the challenges he presents and ones that will emerge.

Worden begins with a discussion on innovation. He describes the value of training and testing at the tactical level and then challenges both Airmen and the institutional Air Force to invest in intellectual and academic pursuits such as advanced degrees. Worden also advocates for simulations or war games, conferences and roundtables, and sponsored research initiatives all intended to develop innovations and agility at the strategic level. Whether in direct response to Worden or not, the Air and Space Forces have changed education and training accordingly over the better part of the last decade. The services are leveraging virtual reality and other twenty-first-century models of learning in essential areas such as pilot training and technical training.³

Air University has been the vanguard of professional military education reform. The institution has increased its intellectual rigor, enhancing the professional standing of faculty. Air University faculty today publish more
relevant pieces in traditional venues—books, scholarly journals, and newspapers—and emergent publishing fora including blogs, online journals, and social media.

Worden also detailed the challenges regarding the recognition of space as a contested domain well before the US Space Force was established. A decade later, the Air Command and Staff College collected a premier faculty and developed an innovative and rigorous space curriculum concentration in the Schriever Space Scholars program that has already produced significant research and served as the basis for space professional military education at the intermediate and senior levels and beyond.

Knowledge generation is often associated with the ivory tower syndrome where scholars are isolated from populations that could benefit from their research and discoveries. Worden understood integration of ideas and best practices was the key to agility and adaptability. If ideas and procedures did not make it to the field or did not get incorporated in operational units as normal practice, they were not useful toward preparing for future challenges.

Development of research task forces at Air University has helped bring relevant operators and subject matter experts together to examine pressing tactical and operational problems. Online education through the Global College of Professional Military Education tailors learning modules for all ranks and programs of the Air and Space Forces. This allows for more education at home stations, fewer expensive TDYs away from operational bases and units, and potentially more responsive course content.

For resident education, concentrations such as the Joint All-Domain concentration at Air Command and Staff College and the Grand Strategy program at Air War College help develop officers who are educated to implement tactics, operations, strategy, and whole-of-government solutions at the theater level and higher. Graduates of these programs look beyond military solutions and seek to integrate operational art with political and economic frameworks of analyses.

Worden discusses incorporation—taking the results of innovation and integration of the ideas and practices stemming from new thinking—as the final step in the process of developing the twenty-first-century strategist. It is not enough to produce ideas or develop lessons from mistakes. To be meaningful, solutions must be integrated into organizations; tactics, techniques, and procedures; and operating instructions.

Those who understand the technology and practices that differentiate and define their profession also need to understand the broader national security and policy implications of their profession. Those who do can
then incorporate their understanding and expertise correctly with other military forces into a whole-of-government approach and with Allies and partners to achieve greater and more significant results.

Some sources have pointed to a moribund or ineffective professional military education structure as an impediment to the future effectiveness of the US military and the cause of defeat in the last two decades of conflict. Critiques such as these, even if not universal across the military services or largely unfounded, rest on the same foundation as Worden’s article. A truly effective airpower strategist is not merely a tactician or an operator. Such individuals make intellectual study and curiosity significant aspects of their personal and professional development because these pursuits are integral to effective implementation of national strategy.

Ultimately, generations of military professionals have understood the value of education, study, reflection, publication, and forums for discussion to the development of the best strategists. Technology has not changed the necessity of such education and of professional and personal development. American Airmen understood this in the 1920s and 1930s when, critical of what they saw as outmoded Army professional military education, they established the Air Corps Tactical School to provide an institution to promote these elements. More modern airpower strategists, John Warden and John Boyd among them, also embraced the ideas promulgated in Worden’s essay as essential. The commitment of Air Force professional military education to these ideas today solidifies the service’s dedication to training and educating tomorrow’s leading senior strategists.

John G. Terino

Dr. Terino is the associate dean for policy and strategy at the Air Command and Staff College.

Notes
In 2010, Daryl Morini argued that a diplomatic surge was the only way to achieve a lasting peace and stability in Afghanistan. A decade later, it is unknown whether Morini’s surge would have worked. What is clear is that a lack of strategic design contributed to US and Coalition failures in Afghanistan.

The recent collapse of the Afghan government, the fall of Kabul, and the final chaotic period of American/NATO withdrawal have once again focused attention on what became America’s longest war. In testimony before Congress, General Mark Milley, chairman of the Joint Chiefs of Staff, called the withdrawal a “logistical success, but a strategic failure.” The impact of such a strategic failure, often equated with the American withdrawal from Vietnam, will doubtlessly resonate into the future, challenging alliances and relationships around the globe. For the United States, the fall of Kabul signals a need to evaluate how we assess and link strategy to operations and how we focus the elements of national power toward our strategic goals.

In 2010, Australian Daryl Morini penned an intriguing article looking at the Afghanistan problem at a strategic level from an international relations perspective. Prophetically, he noted, “without a commensurate multilateral diplomatic surge, efforts towards lasting peace and stability in Afghanistan will most likely fail.”

Looking at the Afghanistan problem through the lens of the major actors—Russia, Pakistan, India, Iran, and China—Morini grappled with historical interests and risk calculi, noting their convergences, then proposed diplomatic actions for policy-maker consideration. These actions, he hoped, would create an environment in which stability within Afghanistan could emerge, freed from divisive interventionism.

In Search of a Strategy

As Morini suggested, most Afghanistan War analyses then and since have been operational or tactical. These analyses have often noted that the United States beautifully and effectively executes operations but seems to fail at reaching its strategic goals. While the United States and NATO struggled with Afghanistan’s daunting complexity, Morini set those details aside and instead looked at Afghanistan from a regional point of view, noting “the systematic context of each intervention, including that by...
NATO-ISAF [International Security Assistance Force], is arguably as important as military facts on the ground in explaining changes in the Afghan political scene.”^4

By emphasizing the critical external system links, Morini contributed what could have been an extremely useful set of decisive points. By developing a regional diplomatic strategy, Morini argued, tensions over the precariously balanced Afghan state would be reduced, leading to the potential for stability and a negotiated outcome.

In many respects, this was a road not taken. While the United States articulated strategy or strategies with respect to the region, these strategies took a back seat to a focus within the country on internal influences. The role of Pakistan, for example, was only begrudgingly admitted as successive administrations struggled to maintain rocky relationships with the sixth most populous nation in the world, especially in the wake of the operation that killed Osama bin Laden in May 2011.^5

But successive ISAF plans did not articulate Pakistan’s influence, focusing instead on Hezb-i-Islami, Haqqani Network, and Taliban insurgents. While these groups operated with tacit and sometimes complicit Pakistani support, working from safe havens within the Pakistani state, the planning focus was on the insurgent groups themselves rather than on Pakistan’s regional power calculations.

Similarly, Russian and Iranian influence in Afghanistan was, at least by 2018, a captivating concern for the United States and NATO. In the author’s experience, Russian influence in the north and Iranian influence in the west were seen as eroding potential support for the government in Kabul, while Russia, Iran, and China seemed to be carving Afghanistan into spheres of influence. Still, this influence was of operational interest and did not seem to affect the campaign strategy.

We will never know if Morini’s regional diplomatic strategy would have worked, but it does seem clear, especially with the release of “The Afghanistan Papers” (detailed below) that the strategies under which US/NATO/ISAF involvement were cast were insufficient.^6 This led not to Morini’s fear of an extended civil war but, worse, the complete collapse of the Afghan government and the seizure of the entire state by the Taliban and their insurgent allies, in particular the Haqqani Network. Milley noted it as a strategic failure. Retired General David C. Petraeus said, “there is a specter here, the US withdrawal from Afghanistan, that is going to haunt the US-EU relationship and our other relationships around the world.”^7

In 2019, the *Washington Post* ran a series of articles based on Freedom of Information Act requests for raw interview transcripts conducted by the
Special Inspector General for Afghanistan Reconstruction (SIGAR)—“The Afghanistan Papers.” Established in 2008, SIGAR’s mandate was to focus on reconstruction rather than strategy, but the lack of an overall strategy by successive administrations was one of the key lessons learned by parties analyzing the papers.

While Inspector General John F. Sopko observed a strategy by the Trump administration, he also noted a disconnect. “There was an overarching strategy that was announced by President Trump early, and I think it is about a year or two ago [in 2020]. The difficulty we have, and have been asking, is how do our individual programs support that strategy?”

Even during periods when there was a strategy articulated by the White House, SIGAR found the agency and department implementation was disconnected and disjointed.

“The Afghanistan Papers” reveal a series of planning and strategic assumptions too quickly taken as facts, namely, the Afghan government could overcome corruption; development would result in political legitimacy for the government in Kabul; and the Afghan national security forces would be capable of supporting their American-style army and air force so long as they had funding. These assumptions seem to have been taken as gospel by higher authorities, becoming so sacred they could not be assailed.

The lack of a cohesive strategy and the inability to coordinate across the elements of national power will probably emerge as key lessons learned as the war is studied. From the perspective of Morini’s article, the pressing need to incorporate regional powers into a comprehensive set of strategic-level talks may have gone a long way toward reducing tensions, though the contradictions and divergent interests would likely not have resulted in overall stability. But Morini’s observations reflected those by Sopko, as proclaimed in his testimony before the Senate in 2020.

[A] comprehensive political outcome requires compromises among the Afghan political elite to reform and govern inclusively. It requires compromises in the ongoing talks between the Taliban and the United States. It requires sustained diplomacy to secure support from Afghanistan’s neighbors, especially Pakistan and Iran, and others including Russia, China, India, and the Gulf States. This is a major political diplomatic effort, a campaign that needs to be undertaken.

Undoubtedly, a sense of strategy within the American diplomatic community existed. But the governance, development, and security lines of effort that framed ISAF’s long-term approach may not have mapped directly to long-term US goals, nor did they appear to account for the interests of those key actors surrounding Afghanistan.
A Case for Strategic Design

Whether or not Pakistan, Russia, or Iran would have worked with the United States as Morini suggested will remain an unanswerable question, but failure or inability to steer their influences made the counterinsurgency fight harder. The open secret of Pakistan’s direct involvement through the Inter-Services Intelligence organization has been often commented upon, while Iranian and Iranian-styled improvised explosive devices transformed the conflict. Russian information campaigns sought to degrade Afghan perceptions of the American intervention through the Russia Today television network and other operations.

The Need for Strategic Design

In 2010, the Government Accountability Office concluded the changing security environment required an evolution of how strategy was formulated.

National security threats have evolved and require involvement beyond the traditional agencies of DOD, the Department of State, and USAID. What has not yet evolved are the mechanisms that agencies use to coordinate national security activities such as developing overarching strategies to guide planning and execution of missions, or sharing and integrating national security information across agencies. The absence of effective mechanisms can be a hindrance to achieving national security objectives.

As a concentration within the Air Force’s Air Command and Staff College, the Joint All-Domain Strategist program teaches a method of connecting strategy to operational-level plans through what Jeffrey Reilly calls “strategic design.” This process seeks to reduce or eliminate precisely the kind of disconnect identified by Morini and the Government Accountability Office: connecting clearly articulated national interests of key players to a design at the strategic level, which in turn informs and drives the more familiar design process at the operational level. As Morini suggested, it seeks to identify global nodes and connections that allow planners to devise symmetrical or asymmetrical strategic response options leveraging all elements of national power.

Design is familiar to military planners at the operational level and is taught as a fundamental component of the Joint planning process. The Joint All-Domain Strategist concentration adds design at the strategic level to assess strategy as received through higher-level orders and documents or fill in when it is not. It considers the interests of major players by articulating the observed rather than desired operating environment system and identifies tensions expressed through convergences and divergences.
Morini, citing Kaveh Afrasiabi, states, “there is a convergence of interests in Afghanistan.” And, as Reilly notes, “the natural tendency is to concentrate on divergence in an effort to prevent an adversary from reaching their goals. . . . However, using convergence to influence actors provides a mechanism that can advantageously shape current and future strategic environments.” Perhaps a focus on convergence in Afghanistan would have yielded better tools at the strategic level and better outcomes.

Reilly’s strategic design compares the desired system with barriers to implementation, setting expectations and laying out both capabilities and limitations. This was clearly never well understood in the Afghan case. If it had been, strategists would have set political and military expectations allowing successive administrations to manage the narrative toward more sustainable and desirable outcomes, even if those outcomes were not as rosy as the success–is–just–around–the–corner assessments seemed to constantly suggest.

Flowing from this analysis, strategic design considers problem sets and linkages that enable the multivariate articulation of problems to be addressed through strategic lines of engagement, which in turn enables operational planners to embed their plans within a nested strategy that begins with national interests. While interests may shift from one administration to the next, at the macro level, they tend to be remarkably consistent. The United States Objectives and Programs for National Security—NSC-68—is one of the best examples of an effective, durable strategic design that survived political oscillations throughout the Cold War.

While not a replacement for national-level strategy, strategic design is an excellent way of either validating strategy as received or proposing a strategic framework in which operational planning can commence, embedding it within a schema of national interests. The process parallels the operational design process, making it a familiar approach for military planners. Hopefully, strategic design will arm them with the awareness to press for strategic guidance from across the federal government when it is vague or missing.

Conclusion

Morini’s essay identifies one of the most compelling tensions underlying the Afghanistan War: the lack of a clear strategy at the regional or international level. As we continue to conduct autopsies on the loss of the war, writers such as Morini emerge as luminaries, highlighting flaws with our strategic approach 9 years prior to the release of “The Afghanistan Papers” and 11 years prior to the frenetic withdrawal from Kabul.
One positive outcome from this might be a recognition of the need for connections between how the United States conducts its wars and its strategic and grand strategic interests. Using tools like strategic design, perhaps military planners will have the clarity and courage to press for strategic goals embedded within America’s national interests for any campaign.

Brian R. Price
Dr. Price is an associate professor in the Department of Warfighting at the Air Command and Staff College.

Notes

5. Assessment based on author’s personal experience in the field.
12. Chairman of the Joint Chiefs of Staff (CJCS), Joint Publication 5.0, Joint Planning (Washington, DC: CJCS, December 1, 2020), chap. 4; and Jeffrey M. Reilly, Operational Design: Distilling Clarity from Complexity for Decisive Action (Maxwell AFB, AL: Air University Press, 2012).
Leveraging Regional Partners


SEAN P. BRANIFF

Writing in 2013, Robert Ross correctly predicted the rise of China as the leading focus of US grand strategy. But Ross’s opposition to strategic partnerships with mainland Asian nation ignores the essential benefits these nations could provide in US grand strategy vis-à-vis China.

Grand strategy “is a nation-state’s theory about how to produce security for itself. Grand strategy focuses on military threats, because these are the most dangerous, and military remedies because these are the most costly.”1 Writing early in President Barack Obama’s second term, Robert R. Ross argued in the pages of this journal that the United States was charting an unwise course in its grand strategy to balance effectively against a rising China.2

Today the strategic setting differs somewhat, and crucial questions must be asked: What are the critical pieces in assessing a US grand strategy toward a rising China? To what extent does Ross’s analysis reflect the current strategic environment? He correctly identified China’s rise as the chief focus of US grand strategy. But from the vantage point of the start of the third decade of the twenty-first century, Ross’s argument against a focus on mainland partnerships has lost a degree of efficacy. The military certainly is the primary tool in preparing for the possibility of conflict between great powers. Considering the military instrument in isolation, however, could prompt the United States to overlook the strategic advantages of mainland Asian partners.

Ross’s central contention in “US Grand Strategy, the Rise of China, and US National Security Strategy for East Asia” was that balance-of-power politics have been the mainstay of US grand strategy for nearly two and half centuries. Further, he argued the Obama administration’s efforts to balance against China were misguided in their focus on mainland strategic partnerships rather than regional maritime partners. The primary “strategic imperative” for the United States, he wrote, has consistently been “a divided Europe and a divided East Asia, lest a regional hegemon develop the capability and the ambition to reach across the oceans and challenge US security.”3
For much of our history, he added, we managed to balance regional competitors against each other. But following World War II, the United States “could no longer rely on balance-of-power politics to maintain its security by dividing its flanking regions. Instead, it would have to directly involve itself in European and East Asian politics.” Moving the analysis to the present, Ross noted a need for the United States to act intentionally in East Asia. He claimed this effort centered on military modernization and strategic partnerships with maritime nations.

Grand strategy rises above the debate and discord of the foreign policy concern du jour and takes a long view in considering how best to protect a state’s interests. This *longue durée* approach to strategic thinking is exemplified in George Kennan’s assessment of American prospects vis-à-vis the Soviet Union.

Writing under a pseudonym—one that failed to shield his identity—in the pages of *Foreign Affairs* in 1947, Kennan assessed, “it is clear that the main element of any United States policy toward the Soviet Union must be that of a long-term, patient but firm and vigilant containment of Russian expansive tendencies. . . . The possibility remains (and in the opinion of this writer it is a strong one) that Soviet power, like the capitalist world of its conception, bears within it the seeds of its own decay, and that the sprouting of these seeds is well advanced.”

The basic logic of the Cold War was thus set: contain Soviet expansion and wait out its self-inflicted implosion. Although Kennan disagreed with aspects of the Truman Doctrine that emerged contemporaneously with his “X article” and that set the foundations for US strategy in the Cold War, in the words of John Lewis Gaddis, Kennan’s “ideas, more than those of anyone else, did provide the intellectual rationale upon which it was based.” Kennan identified the threat and recognized the need for a long-term and concentrated response.

The first step in grand strategy, therefore, is to assess the key threats to a state’s security. While states face myriad challenges, for a leading state in the international system, the prospect of great power war and the repercussions of being on the losing side stand preeminent. Ross assessed this strategic imperative, arguing “unless balanced, China could achieve regional hegemony,” the very outcome Ross suggests the United States, buffered by two great protective moats, has sought to avoid since the country’s founding.

Through the lens of grand strategy, this evaluation is not only correct but has not changed in the years since Ross wrote his analysis. Russia re-
mains essentially a disturber rather than a disrupter, while China has the potential to upset America’s international standing.

Russia’s actions, including the annexation of Crimea and active attempts to influence foreign elections, demonstrate a desire to increase its footprint in the international system. But the system itself, or more specifically the rules that govern it and determine the distribution of gains within it, face little threat of upheaval from Moscow. Russia is and will remain a strategic concern for the United States—it retains a permanent seat on the UN Security Council and has a robust nuclear arsenal.

Nevertheless, Russia’s capacity to outpace the United States in any meaningful metric and ultimately disrupt the international order from which the United States benefits is doubtful. China, alternatively, has the economic potential to be a true disrupter, placing the United States and China in what Graham Allison calls “the Thucydides Trap.”

Once the greatest strategic threats are established, the next step in grand strategy development is addressing how to best meet those threats. Ross identified China as the chief, long-term threat to US security. He stressed the importance of continued modernization of US military capabilities and called for the management of strategic partnerships in the region. His emphasis on modernization rings true today, especially considering the growth of Chinese military capabilities since his writing, including an expanding nuclear arsenal, blue-water naval capabilities, and advancements in hypersonic weapons. Ross’s treatment of managing strategic partnerships, however, can benefit from insights from the last decade.

Ross argued the Obama administration’s pivot to Asia was misguided, not in its key strategic emphasis on the region, but in its strategic priorities within the region. Ross critiqued the administration’s focus on mainland partners—including Vietnam, Cambodia, and South Korea—as opposed to prioritizing maritime partners in the region. US actions with mainland partners, Ross argued, were “neither necessary nor effective . . . [and would] ultimately be costly to US interests because they [would] destabilize US-China cooperation.” Further, they would “elicit increased Chinese suspicion of US intentions and greater Chinese resistance to US interests in East Asia and elsewhere.”

Ross’s critique is prudent for its recognition that strategy entails tradeoffs. Shedding the strategic light on one area, by necessity, leaves another area in the dark, and any strategic emphasis comes with an opportunity cost. Ross cites Walter Lippmann’s admonition that “a comfortable surplus of power” need be held in reserve when considering a state’s commitments. Through this logic, Ross contends strategic partnerships with Vietnam,
Cambodia, and South Korea “cannot enhance US security. Because both [Indochina and South Korea] are on China’s immediate periphery, US naval power cannot effectively challenge Chinese coercive power. The coercive capability of China’s contiguous ground force capability . . . cannot be adequately mitigated by US offshore presence.”

But Ross’s critique of partnerships with South Korea and the countries of Indochina overemphasizes the military facets of the current stage of US competition with China. The United States must unite all appropriate instruments of national power to counter China’s ability to rewrite the rules of the international system to its advantage. In this sense, strategic partnerships with regional mainland states, in addition to the maritime powers Ross identified, can limit China’s freedom to maneuver diplomatically and economically.

Economics is, after all, at the heart of the matter. The miracle of the Chinese economy has been the driving force—to paraphrase Thucydides—behind the growth of China and the fear this growth sparks in the United States. As Paul Kennedy observed more than 30 years ago, “the relative strengths of the leading nations in world affairs never remain constant, principally because of the uneven rate of growth among different societies and of the technological and organizational breakthroughs which bring a greater advantage to one society than to another.”

Add to this Robert Gilpin’s connection between power transitions and war and the assessment China’s rise relative to the United States warrants concern. “The law of uneven growth continues to redistribute power. . . . Disequilibrium replaces equilibrium, and the world moves toward a new round of hegemonic conflict. It has always been thus and always will be, until men either destroy themselves or learn to develop an effective mechanism of peaceful change.” Power transition theorists tell us to be concerned, and they are correct, but grand strategy need not wait for the transition to occur. In strategy, the adversary gets a vote; in this case, the United States gets a vote in responding to China’s rise.

Writing more than 15 years before his 2013 SSQ analysis, Ross and coauthor Andrew Nathan noted a dualistic icon of Chinese power. “The Great Wall is a symbol of weakness,” they observed, “because it signals susceptibility to invasion, and of strength, because it represents economic and cultural superiority and an ability to ward off invasion with feats of engineering and vigilance.” The Chinese economy—in some ways the modern Great Wall—likewise portends both strengths and weaknesses to China.

China’s annual economic growth was the envy of the world for the last 30 years, and China’s role within the global economy has only expanded
since Ross’s 2013 analysis. Yet Chinese economic growth has slowed considerably in the last eight years. Emerging market debt crises led to the crash of the Shanghai stock market and the flight of $1 trillion in foreign reserves from China in 2015. Moreover, the complete impact of COVID-19 on the Chinese economy will remain unknown for some time.16

According to the World Bank, by 2019 economic growth had slowed to just 5.95 percent—an enviable number for most major states in the international system, but one that bodes challenges for a nation that saw years of growth above 10 percent.17 The economic realities of COVID-19 drove Chinese economic growth to just 2.3 percent in 2020, China’s worst annual growth since 1976. In face of this economic downturn, China is experiencing growing domestic demographic concerns and the burden of financing the vast Belt and Road Initiative and increased military modernization.

The current strategic setting is not a replay of the Cold War, and the United States should not use Cold War strategy as a playbook. Nevertheless, Kennan’s insights about the long-term implications of a bankrupt political or economic system offer useful tools for today’s strategist. An effective US grand strategy would exploit weaknesses such as China’s declining growth but would best be done through a true pivot to Asia. The United States is in a better position to do so today than it was when Ross wrote eight years ago.

In leaving behind the wars of the Middle East and Central Asia and better leveraging Western Europe to balance itself against Russia, the United States can approach a rising China with the surplus of power Lippman suggested. This Asia-focused strategy would actively engage Asian mainland and maritime partners, recognizing the United States has considerable advantages in the region and acknowledging partnerships offer much beyond the strategic setting of a potential military clash.

The key strategic tradeoffs for a grand strategy toward China thus move beyond Ross’s concern for tradeoffs between potential Asian mainland versus maritime partners and toward a discussion of tradeoffs between different instruments of national power. This discussion emphasizes economic and diplomatic outreach to mainland partners and the costs and benefits of a regional focus, shifting away from the United States Central Command and United States European Command areas of responsibility to give pride of strategic place to United States Indo-Pacific Command. [SSQ]

Sean P. Braniff
Dr. Sean P. Braniff is an assistant professor at the Air War College, Maxwell AFB, Alabama.
Notes

For three decades, the vast majority of major military engagements, notably those of France, have been marked from their opening hours by the employment of airpower. From Afghanistan to the Levant, by way of Libya and Mali, aerial intervention characterized the opening moments of these politically motivated military actions so much so that this modus operandi seems to have become the norm. More generally, it seems unfathomable that countries with substantial aerial assets will nowadays intervene in a theater without controlling the airspace, even in a temporary manner. This tendency is even more significant as a strong aversion to risk and loss of life has taken hold. Airpower, with its relatively small footprint on the ground, is the tool of choice for political engagement.

This inclination is reinforced by 20 years of irregular operations against adversaries with little to no capability to challenge this air superiority. And even when the presence of ground forces was proven indispensable, as was the case in Bosnia in the mid-1990s and in Afghanistan in the early 2000s, air forces still constituted an indispensable component and a necessary prerequisite for military engagement.\(^1\)

The advantages of air forces for political decision makers are known. They permit reversibility of action in the sense that an airborne operation can be canceled or delayed at any moment. They offer a phenomenal precision of effects, even from distances of several thousand kilometers. Finally, air forces allow rapid execution within the period of time—occasionally short—between the political decision to launch an operation and its execution.

Due to these characteristics, air forces seem to have become the alpha and omega of every military engagement in the minds of political decision makers. Does this mean, however, that the systematic employment of aerial assets in every military engagement is prescriptive in nature? Are we trending toward a scheme where the ability to apply airpower over a theater of operations influences the decision of whether to launch a military operation?

These questions highlight the existence of a common perception among political and military authorities regarding the use of air forces, and simi-

\(^{*}\) An earlier version of this article appeared in Think on French Wings 34 (October 2016).
larly, the emergence of an aerial strategic culture. This article will outline the key features of this idea.

**Defining Elements**

The concept of strategic culture is a topic of much debate, and there is still no commonly accepted definition. Colin Gray noted in 2006 that strategic culture remains “a notoriously opaque and vague concept.” In the field of military strategy, the notion of strategic culture is, after all, fairly recent. It was coined in the United States at the end of the 1970s in the context of analyzing Soviet strategic thinking. Although the concept remains relatively new as a research topic, the process of studying the political, economic, or cultural characteristics of nations to explain their national strategies is much older. In fact, the study of particular styles of military strategy among populations was even addressed in the writings of classical authors such as Xenophon, Tacitus, and Machiavelli.

Several definitions of strategic culture have been proposed. Hervé Coutau-Bégarie and Bruno Colson hold to the definition offered in 1991 by Yitzhak Klein, who considers strategic culture “the attitudes and beliefs held close within the heart of a military institution regarding the political objective of war and the strategic and operational method most effective at attaining it.” This definition echoes another proposed a decade earlier, in 1977, by Jack Snyder. Snyder, a researcher at RAND, is the author of the study that formalized the term. He defined strategic culture as “the sum total of the ideas, conditioned emotional responses, and patterns of habitual behavior that members of a national strategic community have acquired through instruction or imitation and share with each other with regard to nuclear strategy.”

The term “nuclear strategy,” relevant as it was to Snyder’s subject of study, need not be understood as restrictive. While Klein’s notion of strategic culture speaks of the military institution, Snyder expands the notion to include members not belonging to a military establishment by speaking of a “national strategic community.” Carnes Lord further considers society as a whole and proposes a slightly less restrictive definition from the point of view of the actors who share this strategic culture. “It is the sum of traditional practices and ways of thinking that, within a society, governs the organization and the use of military force in the service of political objectives.”

Three principal elements of strategic culture emerge from these definitions. First, a strategic culture can be understood as a shared set of technical preferences, moral and ethical values, and specific practices. Further-
more, strategic culture directly influences the choices made in military operations through its aim to achieve national political objectives. Third, this strategic culture is shared by a defined group of actors.

The preceding definitions do not completely agree on the composition of this group. On one hand, a limited circle of military authorities can be clearly identified (Klein); on the other hand, a more inclusive body of decision makers, “members of a national strategic community,” governs the organization and the use of military force (Snyder and Lord). This article will proceed from the idea that the group of actors who share a strategic culture is, by definition, situated at the level of political and military decision makers. This group is distinguished by the fact that it is responsible for defining national and organizational objectives and for planning and executing the commitment of military forces that are tasked to follow their orders.

A framework emerges from the analysis of these definitions, one that allows for reflection on the notion of an aerial strategic culture.

**Preferences, Values, and Practices**

Culture develops over time. Technical preferences, moral values, ethics, and specific practices of decision makers influence the creation of policy and foreign strategy. In the context of this article, aerial strategic culture reflects the role of air forces in national history, in the evolution of a nation’s geopolitics, and, more broadly, in the consistency of public attitudes toward the employment of airpower to achieve national goals.

In France, the common refrain among five white papers on defense or strategic review published since 1972 has been that of a defense and security policy which allows France to maintain her autonomous decision making, affirm her sovereignty, defend her areas of interest, and continue to weigh in on the international stage. It speaks to a constant attitude of decision makers—for the most part Gaullist—when it comes to state security policy and defense strategy. Since the 1960s, this has manifested as three primary missions: nuclear dissuasion, protection of national territory, and intervention beyond the national borders.

The initial face of French nuclear dissuasion was presented through the Air Force with the Mirage IV bomber and the C-135F refueling aircraft. After this, the protection of national territory manifested itself as an execution of aerial means of defense, guaranteeing both national sovereignty in the airspace and aerial defense of the land. Finally, show-of-force missions from the aerial domain illustrate the “Intervention” pillar of France’s defense policy.
To this point, it makes sense to include here the strategic function “knowledge and anticipation,” since the actions of political and military decision-making entities are also informed by intelligence obtained through airborne sensors. From the signals intelligence of the DC-8 SARIGuE and the C-160 Gabriel to the imagery intelligence of the Mirage IV and the new-generation reconnaissance pod on the Rafale, as well as the remotely piloted Reaper, the real-time acquisition of intelligence through aerial means contributes to an optimization of the decision-making process.

Thus, at the heart of each of these primary missions, the air forces take a leading role; it is this input to the primary strategic functions that contributes to decision-maker development. It guides the perception of the decision-making body with respect to the third dimension and its employment potential. Through this process, an aerial strategic culture develops over time according to the contributions of air assets to the primary strategic missions.

**Employment Potential**

If strategic culture influences the choices made about the employment of military assets, it appears that, in return, the capabilities of air assets and their possible uses fuel aerial strategic culture.

The ability of aviation to operate quickly and at a distance has considerably disrupted the notions of time and scale of those who make arms-employment decisions. The combination of combat and refueling aircraft enables strikes at several thousand kilometers several hours after the political order to strike has been given. Additionally, weapons precision and communication capabilities should not be overlooked. Technical progress on strike precision and target acquisition has played and continues to play an essential role in the constant expansion of the strategic role of aviation.

One operation in particular illustrates this range of capabilities. Recall the 0930 hours raid by Rafale fighter aircraft in January 2013, which flew almost 6,000 kilometers before striking targets in northern Mali and only 48 hours after the head of the French armed forces had agreed to a request for aid from Mali. The speed of this intervention should also be linked to advancements in information management. We could also recall the strikes in Syria against ISIS few days after the terrorist attacks in Paris in November 2015. The acquisition of intelligence, in certain cases real time, and the means of command and control reinforce the authorities’ desire to use air assets and contribute to the visibility of airpower.
Consequently, these elements are the motors of an evolution in aerial strategy towards increasingly advanced forms that expand the intervention options for policy decision makers. In writing policy and foreign strategy, the role of air forces is also understood through the capabilities of aerial transport during emerging crises within countries hosting national citizens. Therefore, whether it be armed confrontations resulting from a civil war (Libya 2011), a natural disaster (Haiti 2010), or an industrial catastrophe (Fukushima 2011), it may be important to be able to protect nationals through rapid action and at some distance.

With regard to the diplomatic and human stakes involved, the decision to evacuate French citizens is often maximally delayed by the political authority, which explains why these types of operations are most often begun and conducted in a state of emergency. It also reemphasizes that air transportation remains one of the ways to conduct a rapid evacuation, the Kabul airlift, being the most recent example.

The decision-making body’s understanding of the competencies and capabilities of air forces is both an expression and a result of aerial strategic culture.

With regard to the variety of aerial missions, which may or may not be coercive in nature, it seems that the employment possibilities of airpower are instruments within the framework of international relations under the logic of hard power and soft power.

A Diplomatic Factor

Naturally, the notion of aerial diplomacy shines through from the moment the topic turns to understanding the use of aerial assets in order to further foreign policy goals. By extending the definition of military diplomacy proposed by Coutau-Bégarie, aerial diplomacy can be understood as the use of air forces in service of foreign policy outside of a traditional war. Aerial diplomacy combines the cooperative and coercive use of aerial assets in international relations every time the resolution to a diplomatic issue is sought via negotiation rather than a confrontation of military forces.

A nation’s aerial strategic culture is a reflection of that nation’s conception of the role of air forces in foreign policy. Nowadays in France and the West generally, the inclination of political authorities to use aerial assets in managing crises is telling. In 2008, looking back on two decades of aerial operations, France’s Secretary General of Defense and National Security, Louis Gautier, estimated that “in managing multiple international crises
after the Cold War, one also notices a particular enthusiasm for airpower as a political tool, a tool to gesticulate, to pressure, and to coerce. From now on, it is at least as much the long-ranged destructive capacity as the flexibility of airpower that interests the political actor.”

Regarding the surge of crises around the world since this date, this analysis remains relevant. From coercive to humanitarian operations, evacuations to intelligence gathering, aerial asset plasticity, to borrow the expression of Jérôme de Lespinois, was assimilated by decision makers and contributes to the aerial strategic culture. In an interview concerning the situation in Syria, the former president of the French Republic, François Hollande, observed that diplomacy was not possible without military credibility, and he voiced support for the ability of France to conduct aerial strikes at a safe distance.

Hollande also declared that “the threat of strikes, the efficiency of strikes, because they would certainly be pertinent and measured, proportional, and we would not have needed to fly over Syrian territory, that tells you the quality of our army; but the fact that this threat existed facilitated the arrival of a political solution. Therefore, there is no diplomacy possible if there is not also military credibility.”

By way of conclusion, it seems important to note that if the aerial strategic culture of the decision-making body favors the use of airpower, limitations on using it—for operational or diplomatic reasons—also influence the political decision whether to lead an intervention. Naturally, notions to deny access and contest airspace happen. If, among the various types of confrontation, the Air Force becomes a weapon of uncertainty and raises uncertainty in the sense that “it introduces political and military hypotheticals even if it cannot necessarily respond to them,” what happens when the Air Force cannot act? Does this limitation constitute a red line that could result in a political decision of noninterference?

Furthermore, another idea to take into consideration is the aerial strategic culture of public opinion beyond that of the decision-making circles. The perception that public opinion has of national military aerial assets, and the resulting image of their uses, influences to varying degrees the manner in which these assets can be used in a crisis. Finally, the aerial strategic culture is also influenced by the imagination of public opinion and by the manner in which the media handles such questions.

In summary, the question of a shared aerial strategic culture is relevant. French, American, British, and Russian decision makers can identify common views regarding airpower and its use.
Mickaël Aubout

Captain Aubout, French Air Forces, holds a PhD in geography from Sorbonne-Université.

Notes

Mission Statement

Strategic Studies Quarterly (SSQ) is the strategic journal of the Department of the Air Force, fostering intellectual enrichment for national and international security professionals. SSQ provides a forum for critically examining, informing, and debating national and international security matters. Contributions to SSQ will explore strategic issues of current and continuing interest to the larger defense community, and our international partners.

Disclaimer

The views and opinions expressed or implied in SSQ are those of the authors and should not be construed as carrying the official sanction of the Department of the Air Force, the Department of Defense, Air Education and Training Command, Air University, or other agencies or departments of the US government.

Comments

We encourage you to e-mail your comments, suggestions, or address change to StrategicStudiesQuarterly@au.af.edu

(Emails to SSQ will be forwarded to the new email address for the journal, and the web information will be updated in early 2022.)

Article Submission

Æther: A Journal of Strategy and Airpower considers scholarly articles between 4,000 and 6,500 words from US and international authors. Please send your submission in Microsoft Word format via e-mail to

StrategicStudiesQuarterly@au.af.edu

Strategic Studies Quarterly (SSQ)

600 Chennault Circle, Building 1405
Maxwell AFB, AL 36112–6026
Tel (334) 953–7311

View and Subscribe to Strategic Studies Quarterly at https://www.airuniversity.af.edu/SSQ/

Free Electronic Subscription

Like SSQ on Facebook at https://www.facebook.com/StrategicStudiesQuarterly

Strategic Studies Quarterly (SSQ) (ISSN 1936-1815) is published by Air University Press, Maxwell AFB, AL. This document and trademark(s) contained herein are protected by law and provided for noncommercial use only. Reproduction and printing are subject to the Copyright Act of 1976 and applicable treaties of the United States. The authors retain all rights granted under 17 U.S.C. §106. Any reproduction requires author permission and a standard source credit line. Contact the SSQ editor for assistance.