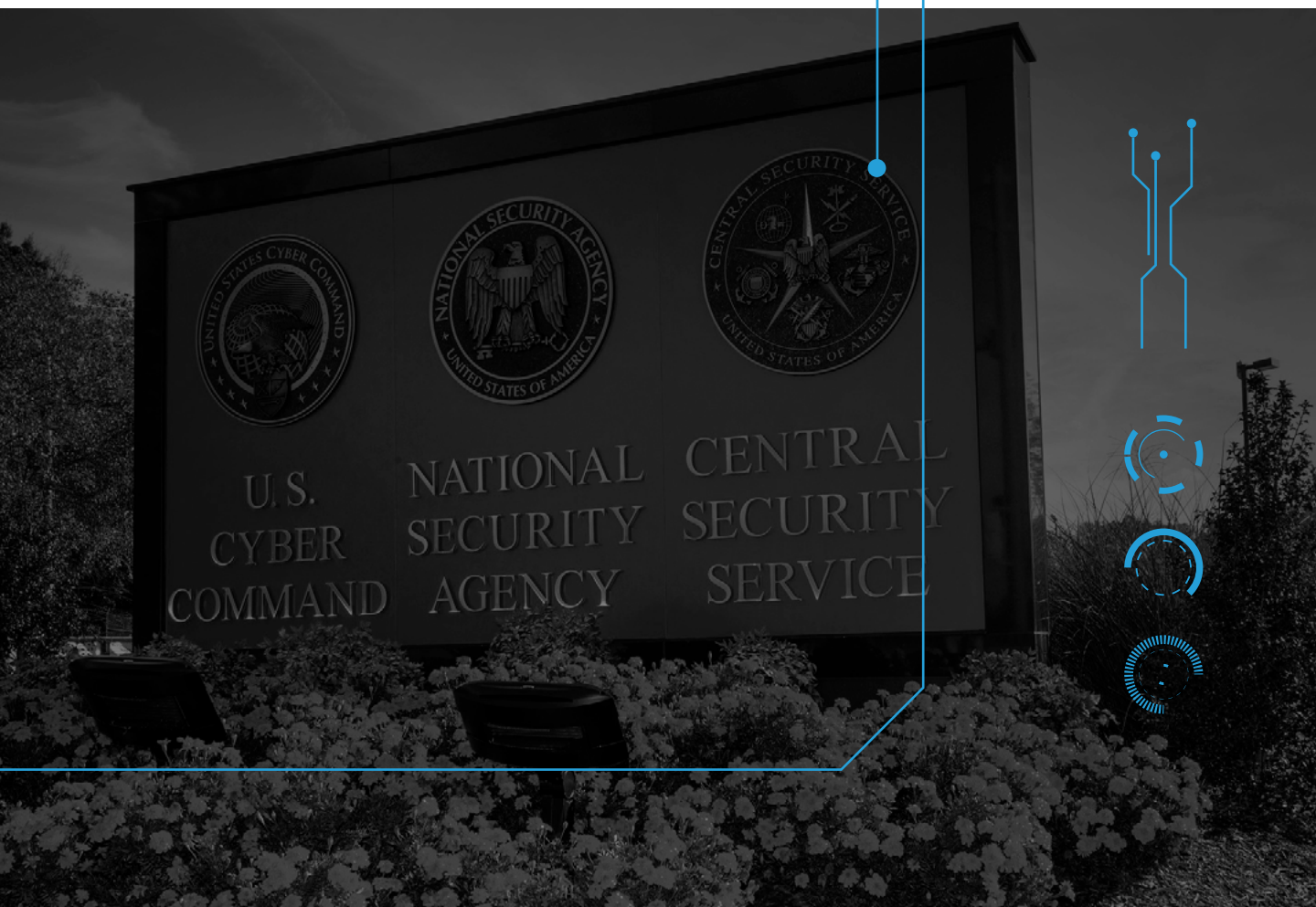
Texas National Security Review: Volume 3, Issue 4 (Fall 2020)
Print: ISSN 2576-1021 Online: ISSN 2576-1153

THE ESCALATION INVERSION AND OTHER ODDITIES OF SITUATIONAL CYBER STABILITY

Jason Healey

Robert Jervis



As the United States shifts to a new military strategy of defending forward against adversaries in cyberspace, research into the role of cyber capabilities in crisis stability is especially relevant. This paper introduces the concept of situational cyber stability, suggesting the key question is not "whether" cyber capabilities are escalatory but rather how they are escalatory under certain geopolitical conditions. We identify four key mechanisms: Pressure Release, Spark, Pull Out the Big Guns, and the Escalation Inversion. Optimists (believing that "No, cyber conflict is not escalatory") and pessimists ("Oh, yes it is") have each touched on parts of these mechanisms. This paper integrates research from both views to better understand crisis stability in cyberspace across the range of geopolitical contexts, from relative peace to impending war. We examine the role of surprise in cyber conflict and introduce policy recommendations to reduce the chances of crises escalating.

It is one of the most important and debated questions for policymakers and scholars of cyber conflict: Are cyber capabilities escalatory?

The pessimists, in whose camp we normally reside, observe a two-decade trend of increasing cyber aggression acting like a ratchet, not a pendulum. Adversary groups aligned with states have caused physical destruction (starting with the U.S.-Israeli Stuxnet attack on Iran)[1]; savaged private sector companies (Iran's attacks on U.S. banks or the North Korean dismembering of Sony)[2]; disrupted national healthcare systems (North Korea's WannaCry, which disrupted the U.K. National Health Service),[3] electrical grids in wintertime (Russia's takedown of the Ukrainian grid),[4] and national elections (Russia again)[5]; and recklessly created global havoc (Russia's NotPetya).[6] If "escalation" means a potentially destabilizing upward spiral in the intensity of cyber hostilities, then cyber conflict may be

1 David Sanger, "Obama Order Sped Up Wave of Cyberattacks Against Iran," *New York Times*, June 1, 2012, https://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html.

2 David E. Sanger and Nicole Perlroth, "U.S. Said to Find North Korea Ordered Cyberattack on Sony," *New York Times*, Dec. 17, 2014, https://www.nytimes.com/2014/12/18/world/asia/us-links-north-korea-to-sony-hacking.html.

3 Lily Hay Newman, "The Ransomware Meltdown Experts Warned About Is Here," *Wired*, May 12, 2017, https://www.wired.com/2017/05/ransomware-meltdown-experts-warned/.

4 Kelly Jackson Higgins, "Lessons From The Ukraine Electric Grid Hack," *Dark Reading*, March 18, 2016, https://www.darkreading.com/vulnerabilities---threats/lessons-from-the-ukraine-electric-grid-hack/d/d-id/1324743.

5 Office of the Director of National Intelligence, "Background to 'Assessing Russian Activities and Intentions in Recent US Elections': The Analytic Process and Cyber Incident Attribution," January 2017, https://www.dni.gov/files/documents/ICA_2017_01.pdf.

6 Ellen Nakashima, "Russian Military Was Behind 'NotPetya' Cyberattack in Ukraine, CIA Concludes," *Washington Post*, Jan. 12, 2018, https://www.washingtonpost.com/world/national-security/russian-military-was-behind-notpetya-cyberattack-in-ukraine-cia-concludes/2018/01/12/048d8506-f7ca-11e7-b34a-b85626af34ef_story.html.

Table 1: Situational Cyber Stability

Mechanism	Cross-Domain Relationship	During Conditions Of...	Then...	Because...	So the Overall Impact Is...
Pressure Release	Cyber conflict *instead* of armed conflict	Relative peace (no major crisis) or both sides *strongly* want to limit conflict	Cyber capabilities are generally not escalatory	Seen by states as less threatening as they do not cause casualties and are temporary and covert	**Stabilizing**
Spark	Cyber conflict *begets* armed conflict	Tension in cyberspace between states	Cyber conflict can cause acute geopolitical crises	Cyber conflict is intensifying over increasingly existential issues	**Destabilizing**
Pull Out the Big Guns	*Threat of armed conflict* begets cyber conflict	Frequent geopolitical crises	Temptation for more aggressive cyber moves	States would be less willing to accept cyber attacks as mere "pressure release"	**Destabilizing**
Escalation Inversion	Cyber conflict *precedes or accelerates* war	Acute geopolitical crisis	Cyber dynamics tempt early use in major crises when war may seem likely anyhow	Cyber capabilities are seen as best used in surprise attacks, with an asymmetric impact	**Destabilizing**

"the most escalatory kind of conflict that humanity has ever come across."[7] States are getting closer to crossing the threshold of death and major destruction outside of wartime. How long until one state, through mistake, miscalculation, or maliciousness crosses that line?

The optimists have equally compelling arguments — including the contention that so far, none of these admittedly worrying cyber attacks has ever warranted an armed attack with kinetic weapons in response.[8] How, they argue, can cyber conflict be escalatory when states have never responded to cyber attacks with traditional violence? Indeed, there is at least as much evidence for cyber capabilities *reducing* rather than causing or intensifying international crises.[9]

This paper will examine this debate. Much of the dispute about the escalatory potential of cyber capabilities comes down to scope conditions. The question is not "whether" cyber capabilities are stabilizing or destabilizing. Rather, the issue is which outcome is more likely under certain geopolitical circumstances. Current literature often assumes the impact to stability to be situation-independent, which we find unlikely. The risks to stability can change, perhaps quite rapidly, depending on prevailing conditions between states. We analyze these conditions in a framework of "situational cyber stability" and see four main mechanisms: Pressure Release, Spark, Pull Out the Big Guns, and Escalation Inversion.

During periods of relative peace and stability — that is, since the end of the Cold War in 1991 — several characteristics drive cyber capabilities to act as a pressure-release valve. Cyber capabilities open up stabilizing, nonlethal options for decision-mak-

7 Jason Healey, "Cyber Warfare in the 21st Century: Threats, Challenges, and Opportunities", Testimony before the House Armed Services Committee, March 1, 2017, https://docs.house.gov/meetings/AS/AS00/20170301/105607/HHRG-115-AS00-Bio-HealeyJ-20170301-U1.pdf.

8 The main publicly known case of a kinetic response to cyber attacks is the Israeli Defense Forces (IDF) attack on a Hamas hacking cell. Hamas is a nonstate group and the IDF provided warning so that the building was empty when hit. See, Lily Hay Newman, "What Israel's Strike on Hamas Hackers Means For Cyberwar," *Wired*, May 6, 2019, https://www.wired.com/story/israel-hamas-cyberattack-air-strike-cyberwar/.

9 Julian E. Barnes and Thomas Gibbons-Neff, "U.S. Carried Out Cyberattacks on Iran," *New York Times*, June 22, 2019, https://www.nytimes.com/2019/06/22/us/politics/us-iran-cyber-attacks.html.

ers, which are less threatening than traditional weapons with kinetic effects. During periods of acute crisis, however, cyber capabilities have other destabilizing characteristics. In these situations, there are greater opportunities for provocation, misperception, mistake, and miscalculation. Dangerous positive feedback loops can amplify cyber conflict so that it takes on a life of its own with diminishing room for strategic choices by policymakers. Table 1 summarizes our findings.

These findings are likely to have general applicability, applying to the relationship between the United States and its major cyber adversaries of Iran, North Korea, Russia, and China, and also to relationships between rivals such as India and Pakistan.

The first section of this article defines key concepts: stability, escalation, and the new cyber strategy of persistent engagement. We then examine the strong evidence supporting the argument that the use of cyber capabilities has generally not been destabilizing or escalatory (in the sense of leading to a larger, traditional conflict), and theories as to why this is so. Next, we explore the circumstances under which this happy situation might change, with cyber capabilities inviting war. There are also sections on feedback loops in cyber conflict and the poorly understood role of surprise. We conclude with implications and recommendations for policymakers.

Concepts Old and New

Situational cyber stability links concepts that are rather old — including stability, escalation, and intensification — with concepts that are quite new, such as persistent engagement. It is worth explaining each concept in detail.

Stability

The technical definition of stability is *negative feedback* in the sense that moving a system in one direction calls up pressures or forces that move it back toward its original position. This contrasts with *positive feedback*, in which movement in one direction leads to greater movement in that direction.[10]

In Cold War security literature, scholars distinguished between *arms-race* or *strategic stability* and *crisis stability*.[11] These concepts can be used quite successfully to analyze cyber conflict. Traditionally, arms-race stability meant that building a weapon or a force posture would lead to negative feedback encouraging the other side to build fewer or less dangerous weapons. This contrasts with a situation of positive feedback, in which more spending or building by one side would lead to more spending or building by the other side. The research here was highly debated, in part because data on Soviet spending were unreliable and arms procurement involved long time lags.

Crisis stability in a Cold War context meant that the moves that one side took in a crisis reduced the incentives for the other side to do something dangerous — in the extreme case, to start the war. The standard argument was that vulnerable weapons systems or force postures invited an attack, thus increasing crisis instability.

Escalation and Intensification

In the Cold War, scholars made the simple distinction between *vertical escalation* (increased intensity of violence) and *horizontal escalation* (geographic spread). The implication was that escalation brought one closer to all-out war. But, as with NATO's then-doctrine of "escalating to deescalate," the reverse could also be the case.[12]

In cyber conflict, horizontal escalation has come to mean intensification *within* cyberspace itself and is generally considered less serious than vertical escalation *out of* cyberspace to the use of lethal, kinetic weapons. Martin Libicki adopts the definition of Morgan et al., and defines escalation as "an increase in the intensity or scope of conflict that crosses threshold(s) considered significant by one or more of the participants." Intensity is both "number of troops committed to the fight" (measuring inputs, which is comparable to sending more infantry and marines to Afghanistan) and cyber operations that have a more significant impact (measuring outputs or effects).[13] Libicki also adds a third element, determining if one incident was in response to another. We fully agree with the first two elements, though, as we explore further below, we believe the third element may be unnecessary.

10 Robert Jervis, *System Effects: Complexity in Political and Social Life* (Princeton, NJ: Princeton University Press, 1999).

11 Thomas C. Schelling, *Arms and Influence* (New Haven, CT: Yale University Press, 1966).

12 Robert McNamara, "Address by Secretary of Defense McNamara at the Ministerial Meeting of the North Atlantic Council," December 14, 1962, in David W. Mabon ed., *Foreign Relations of the United States, 1961-1963, Vol. VIII: National Security Policy* (Washington DC: U.S. Government Printing Office, 1996), Document No. 120, https://history.state.gov/historicaldocuments/frus1961-63v08/d120.

13 Martin C. Libicki, "Correlations Between Cyberspace Attacks and Kinetic Attacks," in *20/20 Vision: The Next Decade*, eds. T. Jančárková, L. Lindström, I. Signoretti, and G. Visky Tolga (Tallinn, Estonia: NATO CCDCOE Publications, 2020), 201, https://ccdcoe.org/uploads/2020/05/CyCon_2020_11_Libicki.pdf.

Persistent Engagement

Within the U.S. military over the last two decades, the predominant image for what defined cyber success was rooted in Cold War traditions of deterrence: Stability is achieved by having fearsome cyber capabilities and an understood willingness to use them if pressed. Since early 2018, thanks in large part to the work of several international relations scholars, this has shifted to a different assessment: To achieve stability, the military must not only possess capabilities but also routinely use them to counter adversaries.

The U.S. Cyber Command Vision in 2018 insisted on the need for fewer operational constraints. This would allow them to "defend forward" and "pursue attackers across networks and systems." With this agility, they can take the initiative to introduce "tactical friction … compelling [adversaries] to shift resources to defense and reduce attacks."[14] In addition, persistent engagement is expected to enable "tacit bargaining," as each side develops "more stable expectations of acceptable and unacceptable behavior" through repeated engagements.[15] Deterrence is expected to play a role as well, especially through the cumulative frustration of adversary operations.[16]

Though persistent engagement is still in some sense an escalation, as it involves a more intense U.S. response to cyber aggression, proponents argue it can "improve security and stability," because U.S. adversaries will back off through friction, tacit bargaining, and deterrence.[17] The argument that persistent engagement leads to stability requires the assumption that a more forward defense introduces negative feedback to bring activity back toward historical (or agreed-to) levels. It is also possible, of course, that a more engaged forward defense might have the opposite effect — it could create positive feedback where adversaries see the new, more active U.S. position as a challenge to meet rather than back away from.[18]

Many academics have cast doubt on whether cyber capabilities are an effective means of coercion,[19] are effective on the battlefield,[20] or provide asymmetric and substantial advantage to attackers over defenders.[21] This paper will argue throughout, with evidence, that policymakers and militaries are generally acting as if cyber capabilities do give a substantial advantage against other states, before and during crises, and on the battlefield.

Pressure Release: Cyber Capabilities Are Generally Not Escalatory During "Peacetime"

Cyber conflict has not escalated into more traditional kinetic conflict. In 2013, one of us looked back at the history of cyber conflict and wrote that "Nations have not sought to cause massive damage … outside of larger geopolitical conflicts" and "have stayed well under the threshold of conducting full-scale strategic cyber warfare and have thus created a *de facto* norm."[22] Newer research has significantly expanded such assessments.

During times of general peace and stability, or when all particpants *strongly* want to limit their conflict, cyber capabilities have been dampening, providing negative feedback to geopolitical crises. States have not responded kinetically to cyber attacks from other states. Even the responses to the most provocative incidents — those which came closest to the level of an armed attack — have been non-kinetic and mild (or perhaps covert and not yet known). As summarized by Martin Libicki: "Rarely do events in cyberspace – much less escalation in

14 U.S. Cyber Command, *Achieve and Maintain Cyberspace Superiority: Command Vision for US Cyber Command*, March 2018, 6, https://assets.documentcloud.org/documents/4419681/Command-Vision-for-USCYBERCOM-23-Mar-18.pdf.

15 Michael P. Fischerkeller and Richard P. Harknett, "What Is Agreed Competition in Cyberspace?," *Lawfare*, Feb. 19, 2019, https://www.lawfareblog.com/what-agreed-competition-cyberspace.

16 Jason Healey, "The Implications of Persistent (and Permanent) Engagement in Cyberspace," *Journal of Cybersecurity* 5, no. 1 (Aug. 26, 2019), https://doi.org/10.1093/cybsec/tyz008.

17 U.S. Cyber Command, *Achieve and Maintain Cyberspace Superiority*, 2.

18 Healey, "The Implications of Persistent (and Permanent) Engagement in Cyberspace."

19 Erica D. Borghard and Shawn W. Lonergan, "The Logic of Coercion in Cyberspace," *Security Studies* 26, no. 3 (July 3, 2017): 452–81, https://doi.org/10.1080/09636412.2017.1306396.

20 Aaron F. Brantly, Nerea M. Cal, and Devlin P. Winkelstein, *Defending the Borderland Ukrainian Military Experiences with IO, Cyber, and EW*, Army Cyber Institute at West Point, 2017, https://vtechworks.lib.vt.edu/handle/10919/81979; Nadiya Kostyuk and Yuri M. Zhukov, "Invisible Digital Front: Can Cyber Attacks Shape Battlefield Events?," *Journal of Conflict Resolution* 63, no. 2 (February 2019): 317–47, https://doi.org/10.1177/0022002717737138.

21 Rebecca Slayton, "What Is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment," *International Security* 41, no. 3 (January 2017): 72–109, https://doi.org/10.1162/ISEC_a_00267; Jon R. Lindsay and Erik Gartzke, "Coercion through Cyberspace: The Stability-Instability Paradox Revisited," in *The Power to Hurt: Coercion in Theory and in Practice*, ed. Peter Krause (New York: Oxford University Press, n.d.). In a forthcoming article, we will engage with these debates – though largely with the same result as in this article, finding that disagreement often stems from different scope conditions.

22 Jason Healey, ed., *A Fierce Domain: Conflict in Cyberspace, 1986 to 2012* (Washington DC: Cyber Conflict Studies Association, 2013).

cyberspace – lead to serious responses."[23]

Perhaps the most comprehensive quantitative analysis on cyber incidents, by Brandon Valeriano, Benjamin Jensen, and Ryan Maness, found that: "Rivals tend to respond only to lower-level incidents and the response tends to check the intrusion as opposed to seek escalation dominance ... These incidents are usually 'tit-for-tat' type responses."[24]

Why do cyber capabilities act as a pressure release? Josh Rovner has compellingly argued that states see cyber competition largely as an intelligence contest, which operates under different rules than a military one: "[C]yber operations may provide a non-kinetic option for leaders who feel pressure to act in a crisis, but who are wary of using force."[25] The U.S. conflict with Iran offers a clear example. After Iran attacked oil tankers and downed a U.S. drone in June 2019, President Donald Trump cancelled punitive U.S. airstrikes at the last minute out of concern for the escalatory impact of causing perhaps 150 casualties.[26] However, he allowed nonlethal cyber disruption of Iranian computer systems, correctly anticipating Iran would not respond violently.[27] Likewise, according to the *New York Times*, "Iran's supreme leader has blocked any large, direct retaliation to the United States, at least for now, allowing only cyberactivity to flourish, according to American and allied officials briefed on new intelligence reporting."[28]

Valeriano and Jensen argue this is partly because cyber capabilities "offer great powers escalatory offramps [and] signaling mechanisms" and can "shape an adversary's behavior without engaging military forces and risking escalation."[29] Michael Fischerkeller and Richard Harknett likewise describe the "cyber strategic competitive space short of armed conflict" where states "design operations to generate a range of damage ... short of internationally agreed upon definitions of use of force and armed attack."[30] Adversaries have "tacitly agreed on lower and upper bounds" and accordingly "have mutual interests in avoiding escalation to violent conflict."[31]

CYBER CONFLICT HAS NOT ESCALATED INTO MORE TRADITIONAL KINETIC CONFLICT.

Erica Borghard and Shawn Lonergan root their explanation less in the motivations of states than in the specific characteristics of cyber capabilities, which render them "imperfect tools of escalation." Capabilities may not be ready in time for a sudden crisis and have uncertain and often limited effects; their use creates important trade-offs (such as revealing specific, closable vulnerability); and there are few appropriate kinetic response options.[32]

Through survey data, Sarah Kreps and Jacquelyn Schneider found that "for the American public, cyber attacks are qualitatively different from those of similar magnitude from other domains," so that individual Americans "are far more reluctant to escalate in the cyber domain than for ... conven-

23 Libicki, "Correlations Between Cyberspace Attacks and Kinetic Attacks", 211.

24 Brandon Valeriano, Benjamin Jensen, and Ryan Maness, *Cyber Strategy: The Evolving Character of Power and Coercion* (New York: Oxford University Press, 2018), 76.

25 Josh Rovner, "Cyber War as an Intelligence Contest," *War on the Rocks*, Sept. 16, 2019, https://warontherocks.com/2019/09/cyber-war-as-an-intelligence-contest/.

26 Peter Baker, Eric Schmitt, and Michael Crowley, "An Abrupt Move That Stunned Aides: Inside Trump's Aborted Attack on Iran," *New York Times*, Sept. 21 2019, https://www.nytimes.com/2019/09/21/us/politics/trump-iran-decision.html.

27 Julian E. Barnes and Thomas Gibbons-Neff, "U.S. Carried Out Cyberattacks on Iran," *New York Times*, June 22, 2019, https://www.nytimes.com/2019/06/22/us/politics/us-iran-cyber-attacks.html.

28 Julian Barnes, David Sanger, Ronen Bergman, and Lara Jakes, "As U.S. Increases Pressure, Iran Adheres to Toned-Down Approach," *New York Times*, Sept. 19, 2020, https://www.nytimes.com/2020/09/19/us/politics/us-iran-election.html.

29 Benjamin Jensen and Brandon Valeriano, "What Do We Know About Cyber Escalation? Observations from Simulations and Surveys," *Atlantic Council, Scowcroft Center for Strategy and Security*, November 2019, 2, https://www.atlanticcouncil.org/wp-content/uploads/2019/11/What_do_we_know_about_cyber_escalation_.pdf.

30 Michael P. Fischerkeller and Richard J. Harknett, "Deterrence Is Not a Credible Strategy for Cyberspace," *Orbis* 61, no. 3 (2017): 382, https://doi.org/10.1016/j.orbis.2017.05.003.

31 Fischerkeller and Harknett, "What Is Agreed Competition in Cyberspace?"

32 Erica D. Borghard and Shawn W. Lonergan, "Cyber Operations as Imperfect Tools of Escalation," *Strategic Studies Quarterly* (Fall 2019): 122–145, https://www.airuniversity.af.edu/Portals/10/SSQ/documents/Volume-13_Issue-3/Borghard.pdf.

tional or nuclear attack" with the same impact.[33] This, they argued, reinforces a firebreak — a sharp discontinuity — between cyber and kinetic conflict.

Situational Cyber Stability: When Cyber Capabilities Can Be Destabilizing

To sum up: Cyber conflict has not escalated and there are strong, theory-backed reasons why it provides negative feedback, acting as a pressure release pushing back against geopolitical crises. We agree with these conclusions, which explain why cyber conflict *has not yet* escalated and *may not* in the future. However, we believe they hold *only if* the next few decades generally resemble the past few.

A cyber incident might cross the threshold into armed conflict either through a sense of impunity or through miscalculation or mistake.

This stability is situational and we see three major, interrelated mechanisms by which it may change. Cyber conflicts and competition are intensifying over increasing stakes and might inadvertently or intentionally spark a larger conflict; there is a higher likelihood of acute crises, far worse than the relatively bland geopolitical conditions of the past decades; and in times of acute crisis, the dynamics go through an inversion, encouraging rather than suppressing escalation.

Spark: Cyber Conflict Can Cause Acute Geopolitical Crises

As cyberspace becomes increasingly existential for economies and societies, states compete more aggressively over the same cyber terrain and treasure. In such circumstances, cyber capabilities add positive feedback, intensifying conflict within cyberspace. Ben Buchanan has featured some of these dynamics in his book, *The Cybersecurity Dilemma*. If a "potential adversary bolsters its own security by increasing its methods of secrecy and ratcheting up intrusive collection of its own — or by shooting back at the collectors — the first state will often feel a need to respond" with "still more intrusive collection."[34] This situation is one which can easily notch upward but only with great difficulty be reversed. This section will summarize the relevant dynamics of cyber conflict, establish that conflict is escalating in cyberspace, and discuss how this dangerous mix of factors can spark war.

Escalation in Cyberspace

Cyber conflict and competition are intensifying. A cyber incident might cross the threshold into armed conflict either through a sense of impunity or through miscalculation or mistake. Alternatively, the cyber attack might be brazen or reckless enough to demand a muscular response from the target state. Libicki's framework of cyber escalation requires three elements: an increase in intensity, the crossing of significant thresholds, and causal links between cyber incidents (i.e., "one attack is in response to another").[35]

We believe the first two elements are important and it is not necessary to balance each incident with its tit-for-tat response. Cyber conflict can be escalatory even if there is not a direct retaliation ("you did A, so we will do X") but rather a trend over time ("we caught you doing A and B, and suspect you of C … so we'll do X and Y and for good measure see no reason to further hold off on Z"). It is through this larger picture, the series of campaigns and capabilities, that the escalatory mechanics become obvious. Despite no provable chain of causation from A to Z, the series can show evidence of intensification and ignored thresholds, if the direction and magnitude of the vector are consistent over a long period of time. A full analysis of escalation requires its own paper, but as an initial analysis we have selected four points each separated by a decade over forty years in order to illustrate this trend:

In 1988, nations did not have major cyber organizations. Within the U.S. Department of Defense, there were small groups planning and conducting offensive operations, but there was no dedicated civilian defensive team in the United States until the creation of the Computer Emergency Response Team, funded by the Defense Department,

33 Sarah Kreps and Jacquelyn Schneider, "Escalation Firebreaks in the Cyber, Conventional, and Nuclear Domains: Moving beyond Effects-Based Logics," *Journal of Cybersecurity* 5, no. 1 (January 1, 2019): 2, https://doi.org/10.1093/cybsec/tyz007.

34 Ben Buchanan, *The Cybersecurity Dilemma: Hacking, Trust and Fear Between Nations* (Oxford: Oxford University Press, 2017), 29.

35 Martin Libicki, "Correlations Between Cyberspace Attacks and Kinetic Attacks."

in November 1988. There were significant incidents — such as the Morris Worm of 1988 and a case known as the Cuckoo's Egg of 1986 which involved German hackers who searched for information on U.S. ballistic missile defense technologies and then passed their findings along to the Soviet KGB. However shocking at the time, those incidents still had quite modest scope, duration, and intensity.[36]

Ten years later in 1998, the world's first combat cyber unit — established in the U.S. Air Force — had already been in existence for three years, with 93 officers and enlisted.[37] The first major cyber bank heist was in 1995 against Citibank, while the U.S. military created the first cyber command in 1998 in response to the internal Eligible Receiver exercise and Solar Sunrise incident.[38] This command was staffed by about two dozen defenders (including one of the authors) and worked with the larger Computer Emergency Response Team and similar teams in the military services to defend against and trace the major Moonlight Maze espionage case to Russia.[39] Within two years, the command expanded and took on responsibilities to coordinate offensive operations, growing to 122 personnel with a $26 million budget.[40]

Only 10 years after that, in 2008, Estonia suffered a debilitating cyber attack from Russia. Espionage against the United States from Russia became increasingly worrisome, including a case known as Buckshot Yankee, where Russian spies breached classified networks. Chinese theft of intellectual property would be known as the "greatest transfer of wealth in history" by 2012.[41] In direct response to these incidents, the Department of Defense combined their dedicated offensive and defensive task forces into a single U.S. Cyber Command in 2010.[42] What had been a defensive-only command with 25 people in 1998 grew to cover both offense and defense with a staff of over 900 by 2011.[43]

In the decade leading up to 2018, the United States launched a sophisticated cyber assault on Iranian uranium enrichment facilities; Iran conducted sustained denial of service attacks on the U.S. financial system; North Korea attacked Sony; and Russia disrupted the Ukrainian power grid in winter (twice) and the opening ceremony of the Olympics.[44] U.S. Cyber Command grew to 6,200 personnel just in the operational element.[45] Iran and China created their own cyber commands as did the Netherlands,[46] the United Kingdom,[47] France,[48] Singapore,[49] Vietnam,[50]

36 Healey, *A Fierce Domain,* 89–119.

37 Sarah Payne White, "Subcultural Influence on Military Innovation: The Development of U.S. Military Cyber Doctrine" (Ph.D. diss, Harvard University, July 2019), https://dash.harvard.edu/bitstream/handle/1/42013038/WHITE-DISSERTATION-2019.pdf?sequence=1&isAllowed=y.

38 Eligible Receiver was a major no-notice cyber exercise run by the Joint Staff in the autumn of 1997 which alarmed the U.S military to the possibilities of a debilitating attack on the nation. Only a few months after that exercise, in February 1998, the U.S Air Force detected a string of intrustions possibly emanating from Saddam Hussein's Iraq, an incident named Solar Sunrise. See, Healey, *A Fierce Domain,* 42–43.

39 Moonlight Maze was a campaign of espionage against the U.S. Department of Defense in the late 1990s and early 2000s, eventually traced back to Russia. The response, led by the FBI, was the first early test of the new military cyber commands. See, Healey, *A Fierce Domain*, 152–163 .

40 U.S. Strategic Command Public Affairs, "Joint Task Force - Computer Network Operations," February 2003, http://www.iwar.org.uk/iwar/resources/JIOC/computer-network-operations.htm.

41 Josh Rogin, "NSA Chief: Cybercrime Constitutes the 'Greatest Transfer of Wealth in History,'" *Foreign Policy*, July 9, 2012, https://foreignpolicy.com/2012/07/09/nsa-chief-cybercrime-constitutes-the-greatest-transfer-of-wealth-in-history/.

42 Healey, *A Fierce Domain*, 72-72.

43 Keith B. Alexander, "Building a New Command in Cyberspace," *Strategic Studies Quarterly* 5, no. 2 (Summer 2011): 3–12, https://www.jstor.org/stable/26270554.

44 For Iranian denial of service attacks, see, Nicole Perlroth and Quentin Hardy, "Bank Hacking Was the Work of Iranians, Officials Say," *New York Times*, Jan. 8, 2013, https://www.nytimes.com/2013/01/09/technology/online-banking-attacks-were-work-of-iran-us-officials-say.html; for the North Korea Sony attack, see, Sanger and Perlroth, "U.S. Said to Find North Korea Ordered Cyberattack on Sony"; for Ukraine, see, Kelly Jackson Higgins, "Lessons From The Ukraine Electric Grid Hack," *Dark Reading*, March 18, 2016, https://www.darkreading.com/vulnerabilities---threats/lessons-from-the-ukraine-electric-grid-hack/d/d-id/1324743; for the Olympic Ceremony hack, see, Andy Greenberg, "Inside Olympic Destroyer, the Most Deceptive Hack in History," *Wired*, Oct. 17, 2019, https://www.wired.com/story/untold-story-2018-olympics-destroyer-cyberattack/.

45 Mark Pomerleau, "Here's How DoD Organizes Its Cyber Warriors," *Fifth Domain*, Sept. 13, 2018, https://www.fifthdomain.com/workforce/career/2017/07/25/heres-how-dod-organizes-its-cyber-warriors/.

46 Dutch Ministry of Defense, "Defence Cyber Command," March 30, 2017, https://english.defensie.nl/topics/cyber-security/cyber-command.

47 Cyber Security Intelligence, "The British Cyber Command," Jan. 22, 2020, https://www.cybersecurityintelligence.com/blog/the-british-cyber-command-4748.html.

48 Christina Mackenzie, "France's New Cyber Defense 'Conductor' Talks Retaliation, Protecting Industry," *Fifth Domain*, Sept. 30, 2019, https://www.fifthdomain.com/international/2019/09/30/frances-new-cyber-defense-conductor-talks-retaliation-protecting-industry/.

49 Prashanth Parameswaran, "What's Behind Singapore's New Integrated Military Cyber Command Objective?," *The Diplomat*, March 10, 2020, https://thediplomat.com/2020/03/whats-behind-singapores-new-integrated-military-cyber-command-objective/.

50 Prashanth Parameswaran, "What's Behind Vietnam's New Military Cyber Command?," *The Diplomat*, Jan. 12, 2018, https://thediplomat.com/2018/01/whats-behind-vietnams-new-military-cyber-command/.

Germany,[51] and others. If intensification is measured as worsening levels of violence, then cyber conflict has intensified across all periods. By 2018, the problems faced in 2008 seemed minor and the organizations small and limited, while the cyber incidents from 1998 and 1988 appeared positively trivial. Operations that had appeared risky 20 years beforehand were now routine.

The intensification trend is also clear according to the measurement of Libicki's "number of troops committed to the fight." The Defense Department expanded the central cyber warfighting force from zero troops in 1988 to 25 in 1998, 900 in 2011, and at least 6,200 in 2018. The first commander of the U.S. Cyber Command noted in 2011 that its creation "garnered a great deal of attention from other militaries," which he hoped was not a sign of militarization but rather "a reflection of the level of the concern with which"[52] nations must indeed be concerned, as there are now dozens of copycats. Jensen, Valeriano, and Maness, using more quantified methods, have similar findings to this qualitative assessment, tracking a strong growth of latent cyber power by Russia and China from 2001 through 2014.[53]

There is no obvious evidence pointing to a decrease or even a plateau in the intensity of cyber conflict, or that fewer thresholds are being passed now than 10, 20, or 30 years ago. The direction and magnitude of the change over four decades has marched in only one direction: a relentless increase as nations build their organizations and employ them in more frequent and more dangerous incidents.

There are three potential criticisms of this assessment. First, few if any of these incidents can be proven to have been direct retaliation. The trend line is clear enough, however, and incidents have driven the creation of new organizations and more assertive strategies. Three generations of U.S. cyber defense organizations were in direct response to incidents and Gen. Paul Nakasone of U.S. Cyber Command directly links his strategy of persistent engagement to the intransigence of others. Because adversaries have had "strategic impact" with their cyber operations, U.S. Cyber Command evolved "from a response force to a persistence force."[54] Likewise, Stuxnet "generated [a] reaction" from Iran, according to the four-star general then leading U.S. Air Force cyber capabilities, and as a result Iran would soon "be a force to be reckoned with" in cyberspace.[55]

Second, it is possible to argue that these attacks did not violate explicit norms or redlines. Yet in a fast-moving area like cyber conflict, it is reasonable for policymakers to decide *post facto* that a transgression has occurred. The Iranian government did not, to our knowledge, specifically forbid the destruction of their uranium-enrichment infrastructure through cyber attacks. Nor was the U.S. electoral system, at the time of the Russian interference in 2016, specified as critical infrastructure and thus off-limits under stated U.S. norms. Surely, it is not unreasonable to expect a U.S. reaction nonetheless.

Third, it is possible that these trends may not indicate intensification as much as increased digital dependence or technological advancement. As the numbers of connected devices and networks skyrocketed over 40 years, it would be no surprise if attacks and organizations scaled as well. We are not convinced by this argument: The statements of participants in cyber incidents repeatedly and specifically denounce the intransigence and audacity of others, ratcheting up their response. Nor do we find the advancement of technology to be a satisfactory explanation. Adversaries took progressively more risks during the 40-year period under examination: Even technically similar attacks increased in intensity over time. The 2016 election interference was through the hacking of emails — a kind of cyber incident that was neither rare nor advanced in 1998. Only the Russian audacity to release those emails to influence an election was novel. In 2008, both the Obama and McCain presidential campaigns suffered Chinese (and also possibly Russian) intrusions, but only as passive intelligence collection. The campaigns had apparently little concern that the stolen information would be doctored or released.[56] By 2018, conflict had inten-

51 Ludwig Leinhos, "The German Cyber and Information Domain Service as a Key Part of National Security Policy," The Centre of Ethical Education in the Armed Forces, April 1, 2017, http://www.ethikundmilitaer.de/en/full-issues/20191-conflict-zone-cyberspace/leinhos-the-german-cyber-and-information-domain-service-as-a-key-part-of-national-security-policy/.

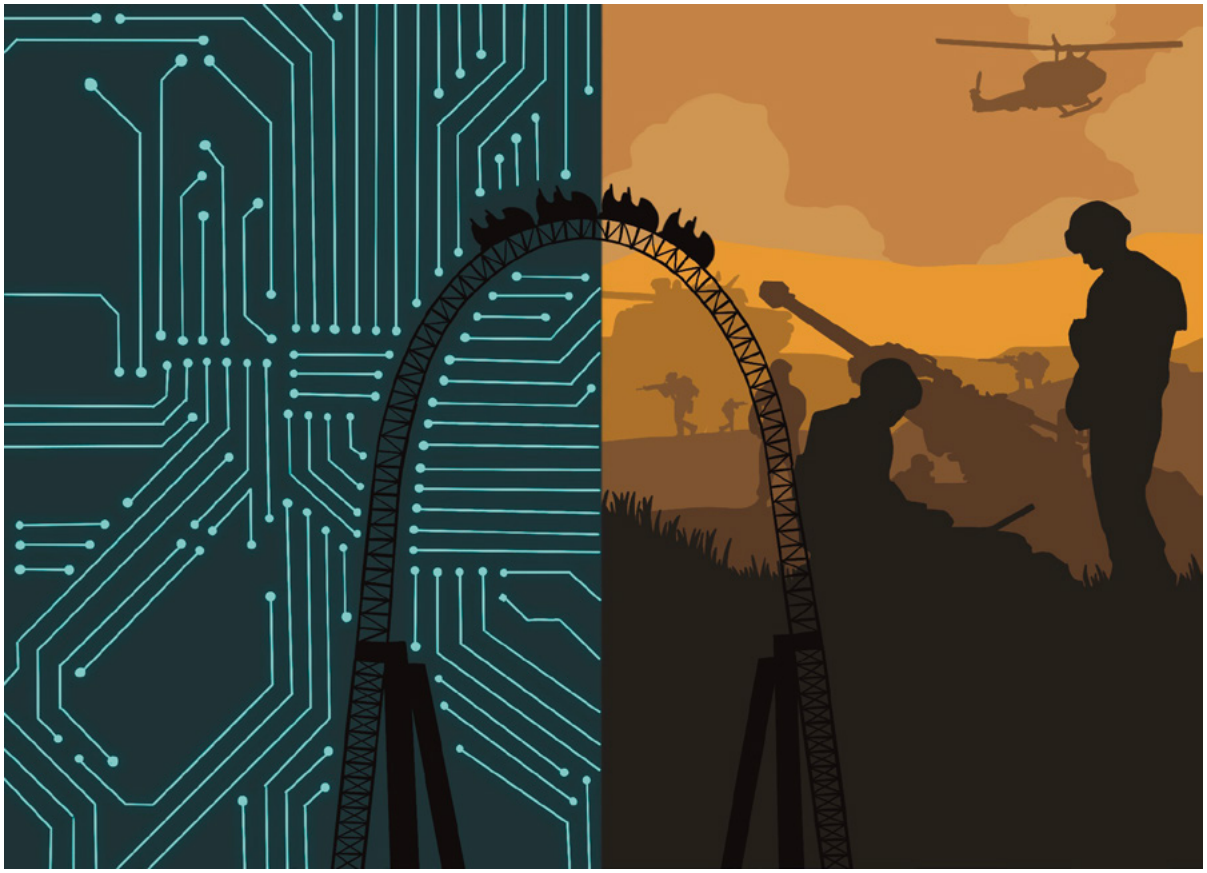
52 Keith B. Alexander, "Building a New Command in Cyberspace," 7.

53 Valeriano et al., "Cyber Strategy: the Evolving Character of Power and Coercion," 70.

54 Paul M. Nakasone, "An Interview with Paul M. Nakasone", *Joint Forces Quarterly* (1st Quarter 2019): 4–9, https://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-92/jfq-92_4-9_Nakasone-Interview.pdf.

55 Andrea Shalal-Esa, "Iran Strengthened Cyber Capabilities After Stuxnet: U.S. General," *Reuters*, Jan. 18, 2013, https://www.reuters.com/article/us-iran-usa-cyber-idUSBRE90G1C420130118.

56 Michael Isikoff, "Chinese Hacked Obama, McCain Campaigns, Took Internal Documents, Officials Say," *NBC News.com*, June 10, 2013, http://www.nbcnews.com/id/52133016/t/chinese-hacked-obama-mccain-campaigns-took-internal-documents-officials-say/; Jeff Stein, "Exclusive: How Russian Hackers Attacked the 2008 Obama Campaign," *Newsweek*, May 12, 2017, https://www.newsweek.com/russia-hacking-trump-clinton-607956.



sified so that none could have such assurances.

We ask unconvinced skeptics to consider not the hard version of the argument — that cyber conflict is *definitely* intensifying and therefore may spark a conflict — but a softer one: that cyber conflict is possibly intensifying or might do so in the future.

A Dangerous Mix

Cyber conflict presents a situation that has no obvious parallels in military history. States covertly experiment with capabilities below the threshold of armed attack and implant them in adversary systems well before hostilities, creating an "environment in which multiple actors continue to test their adversaries' technical capabilities, political resolve, and thresholds," as Director of National Intelligence James R. Clapper testified in 2015.[57] The testing of capabilities and resolve will always increase the chances of miscalculation and mistakes. In few other fields of military endeavor are such aggressive activities so routinely conducted within adversary critical infrastructure during peacetime.

The major cyber powers — and more than a few minor ones — behave greedily in cyberspace. Unhappy with the cyber status quo, they seek to seize as much "territory" (computers and servers in other countries — "gray space" in the U.S. euphemism) and "high ground" (such as core internet routers) as they can.[58] Since no one else seems to be showing much restraint, it may seem a sucker bet to do so, especially with the growing sense that the advantage lies in seizing the initiative.

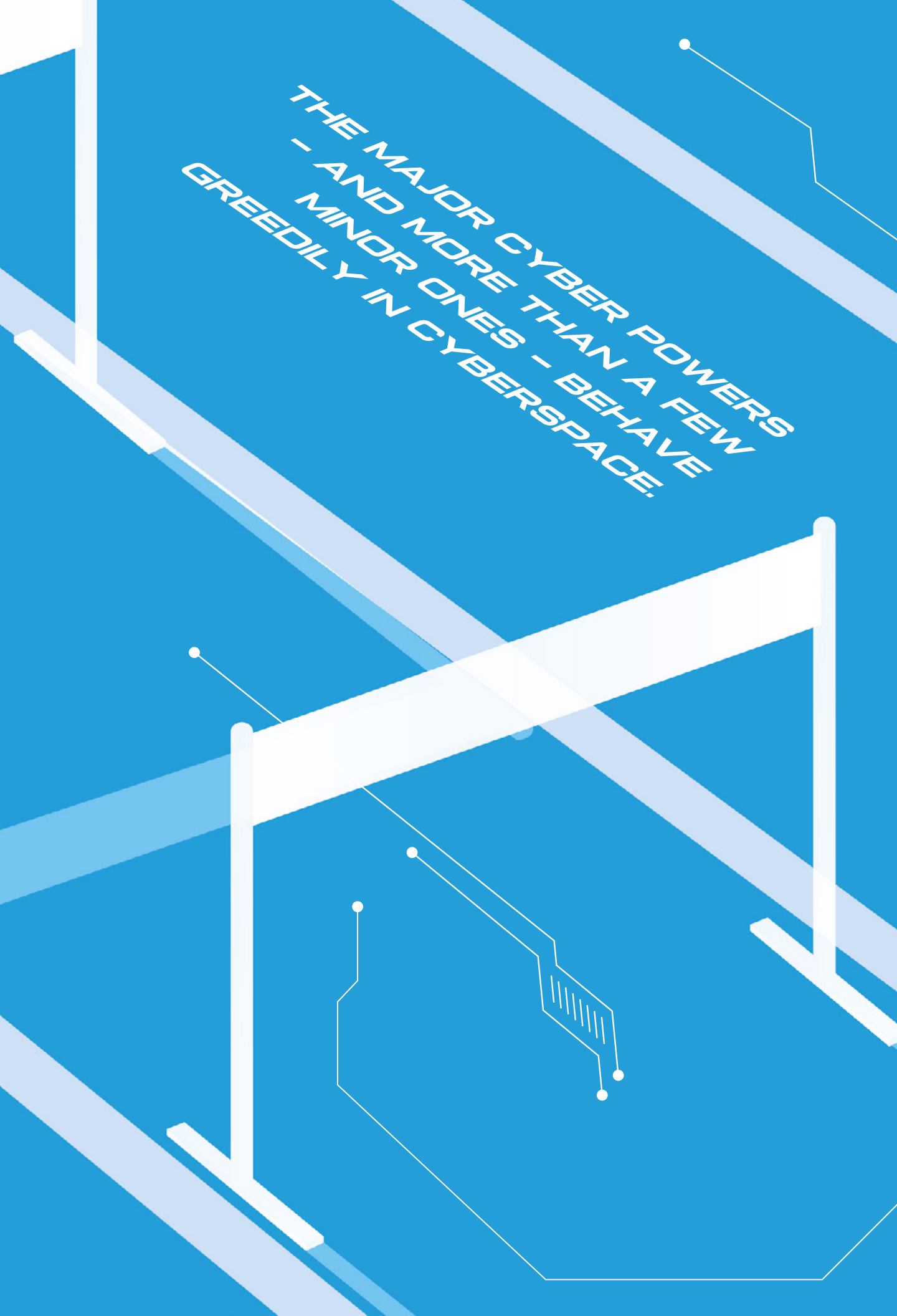
As U.S. cyber operations are said to play "nice" and don't spread wildly or cause collateral damage,[59] many argue that the "status quo is deteriorating into norms that by default are being set by adversaries."[60] Such conclusions, with the United States loudly asserting its victimhood, are based on a selective choice of evidence. It

57 Director of National Intelligence James R. Clapper, "Worldwide Threat Assessment of the US Intelligence Community," Statement for the Record, Senate Armed Services Committee, Feb. 26, 2015, https://www.armed-services.senate.gov/imo/media/doc/Clapper_02-26-15.pdf.

58 Ben Buchanan cites Canadian documents which confirms states "acquire as many new [Operational Relay Boxes] as possible in as many non 5-Eyes countries as possible." Buchanan, *The Cybersecurity Dilemma*, 48.

59 Zaid Shoorbajee, "Playing Nice? FireEye CEO Says U.S. Malware Is More Restrained than Adversaries'," *CyberScoop*, June 1, 2018, https://www.cyberscoop.com/kevin-mandia-fireeye-u-s-malware-nice/.

60 Richard J. Harknett, "United States Cyber Command's New Vision: What It Entails and Why It Matters," *Lawfare*, March 23, 2018, https://www.lawfareblog.com/united-states-cyber-commands-new-vision-what-it-entails-and-why-it-matters.


THE MAJOR CYBER POWERS
- AND MORE THAN A FEW
MINOR ONES - BEHAVE
GREEDILY IN CYBERSPACE.

is easy, when reading U.S. official documents, to forget that the United States was a predator long before it was prey.

American leaders have no problem recognizing that "autocratic governments ... view today's open Internet as a lethal threat to their regimes." Yet they have more difficulty making connections between cause and effect or seeing the situation through the eyes of their rivals.[61] Adversaries perceive that the United States first broke the status quo (by dominating the early internet, pushing for a borderless cyberspace, and building a massive early lead in cyber espionage) and are hitting back, not acting first. To such states, calls to act "responsibly" may appear indistinguishable from demands that they acquiesce to a cyberspace inimical to their survival.

Adversaries can believe that the United States does not play by its own rules. According to the U.S. intelligence community, President Vladimir Putin of Russia was convinced the release of embarrassing financial data from the Panama Papers was a U.S. covert action. This was partly the cause for Putin's decision to interfere in the U.S. elections, which in turn was met with disruptive attacks on the main Russian troll farm by U.S. cyber operators.[62] Chinese leaders may believe that U.S. confidence building and transparency measures, such as discussing a new cyber strategy, are swaggering moves meant to cow Beijing.[63] Iran's cyber operations were almost entirely focused on dissidents until they were hit by the U.S.-Israeli Stuxnet attack, after which Iran raced to build and use its own capabilities. After the revelations of Edward Snowden, European allies were astonished by the scope of U.S. espionage and its lack of restraint.[64]

President Trump, in 2018, reportedly approved the CIA to conduct significantly more operations under less oversight, including "cyber attacks on Iranian infrastructure" and "covert hack-and-dump actions aimed at both Iran and Russia."[65] Any Russian or Iranian attacks since then may have been reprisals, though this would be unknown to researchers, U.S. citizens, and senior government officials and members of Congress who did not have the need to know. Because of compartmentalized knowledge, there are few who know what punches a country is taking, which it is throwing, and the causal relationship between the two.

It is misguided to base any cyberspace policy, theory, or strategy on statements that ignore the role U.S. cyber operations have had in shaping the status quo. We don't argue there is any *ethical* equivalence between the cyber operations of the United States and other nations. Rather, there may be an *escalatory* equivalence when no one thinks anyone else is paying attention to complaints, redlines (tacit or explicit), or perceived norms.

To sum it up: Cyber-induced crises which escalate into larger geopolitical crises are more likely in the coming years, fed by this intensification of operations, insensitivity to the perceptions of others, and a fear of existential digital risks. States will increasingly feel angry, paranoid, trigger-happy, and vengeful, and they will turn to their militaries for salvation — a chaotic recipe, ripe for error, and potentially overwhelming any dampening effects of cyber capabilities. Cyberspace is no longer the preserve of researchers, e-commerce sites, and nerds. It is now existential to a growing number of states. Advanced states rely on connectivity, including the Internet of Things, not just for communication but control of the economy and industry. Cyber conflict may be an intelligence contest, as Rovner and others contend — but if that is true, it is a contest taking place inside a $1.35 trillion digital economy (and that's just the contribution to the United States) and across insecure technologies that hold citizens' most intimate secrets.[66]

61 Admiral Michael S. Rogers, "Statement of Admiral Michael S. Rogers Commander United States Cyber Command Before the Senate Armed Services Committee," Statement Before the Senate Armed Services Committee, March 19, 2015, https://fas.org/irp/congress/2015_hr/031915rogers.pdf.

62 Office of the Director of National Intelligence, "Background to "Assessing Russian Activities and Intentions in Recent US Elections: The Analytic Process and Cyber Incident Attribution," Jan. 6, 2017, https://www.dni.gov/files/documents/ICA_2017_01.pdf; Ellen Nakashima, "U.S. Cyber Command Operation Disrupted Internet Access of Russian Troll Factory on Day of 2018 Midterms," *Washington Post*, Feb. 27, 2019, https://www.washingtonpost.com/world/national-security/us-cyber-command-operation-disrupted-internet-access-of-russian-troll-factory-on-day-of-2018-midterms/2019/02/26/1827fc9e-36d6-11e9-af5b-b51b7ff322e9_story.html.

63 Adam Segal, "What Briefing Chinese Officials On Cyber Really Accomplishes," *Forbes*, April 7, 2014, https://www.forbes.com/sites/adamsegal/2014/04/07/what-briefing-chinese-officials-on-cyber-really-accomplishes/.

64 Adam Segal, *The Hacked World Order: How Nations Fight, Trade, Maneuver, and Manipulate in the Digital Age* (New York: PublicAffairs, 2016), 143-151, and off-the-record conversation at Munich by one of the authors and a European head of state, and members of the European parliament, and other policymakers, February 2014.

65 Zach Dorfman, Kim Zetter, Jenna McLaughlin, and Sean D. Naylor, "Exclusive: Secret Trump Order Gives CIA More Powers to Launch Cyberattacks," *Yahoo News*, July 15, 2020, https://news.yahoo.com/secret-trump-order-gives-cia-more-powers-to-launch-cyberattacks-090015219.html.

66 U.S. Bureau of Economic Analysis, "Digital Economy Accounted for 6.9 Percent of GDP in 2017," April 4, 2019, https://www.bea.gov/news/blog/2019-04-04/digital-economy-accounted-69-percent-gdp-2017.

Pull Out the Big Guns: Acute Crises Invite More Aggressive Cyber Moves

Our second concern is that acute geopolitical crises — having little to do with cyber competition — will be more likely in coming years, leading states to become even more risk-seeking. The intensity of acute crises, including the threat of great-power war, will create conditions well outside the scope of theories on the dampening effects of cyber capabilities. States may be unwilling to adhere to the tacit agreements of quieter times and limit themselves to the relative restraint of an intelligence contest. Rather, cyber capabilities would be used by states in more provocative ways to match the perceived dangers of the crises. If participants are not strongly committed to limiting the conflict, using cyber capabilities will not be a reliable pressure release as it had been in the past.

As one of us has written with Jack Snyder:

> Cyber competition has developed during a period of relative peace and stability between major powers. Perhaps cyber competition has been below the threshold of armed attack simply because after the Cold War, post-1991, adversaries have been (relatively) restrained from armed attack in all its forms, not just cyber. The desire to avoid escalation, and cyber-as-pressure-release, may not be inherent to cyber competition but merely be an inherited characteristic from the global balance of power during the entire period under consideration. A decay of that geopolitical stability could light a match to significantly different and worsening cyber competition.[67]

Harknett and Fischerkeller acknowledge the scope conditions of their own work, clarifying their prescriptions only apply to the "competitive space *short of* armed conflict" and not the "competitive space *of* armed conflict."[68] The barrier between the two may be quite thin. A higher risk of crises also weakens the dampening effects cited by Borghard and Lonergan. States will use their stockpiled capabilities — they will accept the higher risk of using uncertain capabilities and care less about the trade-offs.[69] Adversaries on the receiving end of riskier, more dangerous attacks during a geopolitical crisis will feel less restraint in choosing harsh, even kinetic, responses.

Escalation Inversion: Dynamics Tempt Early Use in Acute Crises

The third and related concern of cyber situational stability is that the use (or fear) of cyber capabilities will escalate acute geopolitical crises. When major national interests are at stake, with the real threat of war, different dynamics of cyber conflict come into play. Indeed, for Erik Gartzke and Jon Lindsay, "The same strategic logic that leads us to view cyberwar as a limited political instrument in most situations also leads us to view it as incredibly destabilizing in rare situations."[70] The "gray zone" below the level of armed conflict may be narrower than policymakers, practitioners, and academics expect.

As crises intensify, the perceived advantage of going first will tempt many adversaries to make cyber attacks they might have withheld otherwise, overstressing the normal pressure-release mechanisms and encouraging rather than dampening escalation. Cyber capabilities may be to World War III as mobilization timelines were to World War I.

It is not terribly relevant whether cyber capabilities can actually have such a strategic, surprise impact. Policymakers and elites seem to believe they can, as is made evident by the intensification discussed above, the reinforcing of critical infrastructure against cyber attacks, and the nearly 30-year lifetime of the concept of a Cyber Pearl Harbor — a sudden and major cyber attack that is carried out with no warning. States may launch a major attack hoping for a surprise, strategic impact. Taking such a shot and missing may lead to similar backlash as succeeding, unless the successful defenders — in the middle of a major geopolitical crisis — decide to shrug off a strategic attack. If the cyber attack becomes publicly known, the policymakers may have no choice but to make a muscular response.

Because cyber capabilities are seen to favor the attacker or the actor taking the initiative, the "incentives to strike first will turn crises into wars."[71] This effect is exacerbated if a nation simultaneously has ineffective defenses yet brags, as the Chairman of the U.S. Joint Chiefs of Staff has done, of

67 Jason Healey and Jack Snyder, "Strategic Equilibrium and Persistent Engagement in Cyberspace," draft paper, June 2020.

68 Fischerkeller and Harknett, "What Is Agreed Competition in Cyberspace?"

69 Borghard and Lonergan, "Cyber Operations as Imperfect Tools of Escalation."

70 Eric Gartzke and Jon R. Lindsay, "Thermonuclear Cyberwar," *Journal of Cybersecurity* 3, no. 1 (March 2017): 1, https://doi.org/10.1093/cybsec/tyw017.

71 Robert Jervis, "Cooperation Under the Security Dilemma," *World Politics* 30, no. 2 (January 1978): 211, https://doi.org/10.2307/2009958.

"incredible offensive capability" to "deter [adversaries] from conducting attacks."[72] As one of the authors has written elsewhere:

> Sixty years ago, during the Cold War, the preferred plan of Strategic Air Command (SAC) was to maximize striking potential by basing nuclear-armed bombers as close as possible to the Soviet Union. Albert Wohlstetter wrote in a RAND Corp. report that this invited a surprise attack: The bombers and tankers parked on those bases would be both existentially threatening to the Soviet Union and themselves vulnerable to a Soviet nuclear attack. ... The combination of a terrifying offense and weak defense would create perverse incentives for the Soviet leadership to launch a disarming strike as early as possible in any crisis. ... [S]ome adversaries will choose the surprise attack rather than waiting to face off with the deadliest gunfighter around. Indeed, the more the gunfighter improves on and boasts about his deadliness, the more he brandishes his pistols, the more incentive there is to get the drop on him, especially if a fight seems inevitable anyhow.[73]

A report on U.S.-Russian crisis stability co-authored by Jim Miller, the former third-ranking Pentagon official, notes these larger dynamics of drawing first, before the other guy draws on you:

> Cyberspace and outer space offer the attacker a very attractive combination: the potential for high impact on the other side's military, with the potential for limited, or even no, direct casualties ... [T]here are likely to be strong incentives on each side to use these capabilities in large doses early in a major conflict to gain coercive and military advantage – and to attempt to prevent the other side from gaining such advantage.... Combatants may worry that an adversary will take measures to reduce its cyber vulnerability, providing reason to strike early while the window to do so effectively appears open.[74]

Miller believes that "the incentives to start[ing] any military conflict with a significant attack in cyberspace and outer space," and to do so before an adversary, "are enormous."[75] This effect is magnified if an adversary believes that strategic weapons systems (especially nuclear weapons or nuclear command and control) and space-based intelligence and detection systems may be vulnerable to a blinding or disarming cyber strike.[76]

As it is nonlethal and reversible, a cyber potshot may seem less escalatory, tempting adversaries to take shots they wouldn't otherwise. Since the U.S. military may seem otherwise unbeatable, an adversary's "weakness may compel him to compensate with audacity in order to redress the balance."[77]

In this situation, the sense that cyber is a pressure-release valve becomes positively dangerous. Optimism can be a self-denying prophecy. If decision-makers believe that the system will be stable regardless of their actions, they will act uncaringly, in a way that ultimately destabilizes that system. If a little cyber conflict is stabilizing, then a lot more cyber conflict should be even better.

The findings of Kreps and Schneider, based on surveys of the American public, suggest a firebreak (a clear delineation, perhaps even associated with a taboo) between cyber and kinetic conflict. In their experiment, a cyber attack with a given impact (such as the destruction of a power plant) was seen as less severe than a kinetic effect with the same impact. Americans were "considerably more restrained when it comes to aggressive retaliatory actions involving the use of force" to respond to cyber attacks.[78] This finding may tell us less about firebreaks than about potshots. If the United States won't take a surprise cyber attack too seriously, even if it caused death and destruction, why not take such a shot? Rather than seeing this survey as soothing evidence, we fear it demonstrates worryingly destabilizing dynamics.

As it was for the Japanese in December 1941, the question may become: If not now, when? And if not

72 Mark Milley, "Gen Milley Chairman Confirmation Testimony," Testimony Before the Senate Armed Services Committee, July 11, 2019, https://www.c-span.org/video/?c4806722/user-clip-gen-milley-chairman-confirmation-testimony.

73 Jason Healey, "Getting the Drop in Cyberspace," *Lawfare*, Aug. 19, 2019, https://www.lawfareblog.com/getting-drop-cyberspace.

74 James N Miller Jr and Richard Fontaine, *A New Era In U.S.-Russian Strategic Stability* (Washington DC: Center for a New American Security, 2017), 48.

75 James N. Miller, email to the authors, May 29, 2020.

76 U.S. Department of Defense, Defense Science Board, *Task Force on Cyber Deterrence* (Washington DC, February 2017), 17–24, https://apps.dtic.mil/dtic/tr/fulltext/u2/1028516.pdf.

77 Richard K. Betts, *Surprise Attack: Lessons for Defense Planning* (Washington, DC: Brookings Institution, 1982), 129.

78 Kreps and Schneider, "Escalation Firebreaks in the Cyber, Conventional, and Nuclear Domains," 5.

this way, how? The short shelf life of cyber capabilities may force use-or-lose choices once an adversary expects a conflict. If you have secret torpedoes that can be used in shallow harbors like Pearl Harbor, and conflict seems inevitable, why not use these weapons in a surprise attack before the adversary can counter your exquisite advantage?

Answering the Stability Question: Does Cyberspace Encourage Positive and Negative Feedback?

These concerns of situational cyber stability depend on whether cyberspace and cyber conflict are marked primarily by positive or negative feedback. The accuracy of any analysis of cyber stability relies on the answer to this question, yet it is rarely asked. Strategies and theories are often built on an implicit assumption of relative stability — that since it has been stable in the past it will continue to be so in the future.

If the overall system is marked by negative feedback, then it is like a nice, solid car, engineered for balance and tolerant of mistakes by young and inexperienced drivers. If this holds for cyber conflict, the fluctuations caused by aggressive cyberspace moves by states, even during acute crises, will calm over time. The concerns of Spark, Pull Out the Big Guns, and the Escalation Inversion mechanisms will remain largely theoretical in the face of continued Pressure Release.

If the system is marked by positive feedback, though, then it is more like a clunky jalopy driven on icy roads. Relatively tiny inputs are all it can take to induce wild swings, which amplify unless they are actively and expertly countered by an alert driver. At some point, the driver is no longer in control, as the dynamics take on a life of their own with little room for steering input (or strategic choices). Cyber attacks, in this model, beget worse cyber attacks, and eventually throw the system out of whack, especially through Spark, but also Pull Out the Big Guns or an Escalation Inversion.

Our own preliminary conclusion is that cyber conflict seems to favor positive feedback. In 1978, one of us wrote that security dilemmas of spiraling escalation between rivals would be "doubly dangerous" if it is hard to distinguish offense from defense and if the offense has the overall advantage. Each side would see even defensive moves as escalatory. Because the defense was feckless, the "incentives to strike first will turn crises into wars."[79]

Our analysis of the characteristics of cyberspace leads us to worrying conclusions. Cyber conflict is not merely doubly dangerous, but perhaps quintuply dangerous for these reasons:

1. Offense and taking the initiative are seen to have the advantage — certainly in perception and perhaps in fact.
2. It is hard to distinguish offense from defense, but also from espionage, subversion, sabotage, or contingency preparation for some future attack.
3. There are such low barriers to entry that many states (and nonstate groups) are involved, producing a more complex situation than the dyadic U.S.-Soviet confrontation of the Cold War.
4. Capabilities are not just kept in arsenal but used — covertly and with perceptions of impunity.
5. The complexity of cyberspace means that even expert practitioners cannot understand it well, leading to a significant chance of cascading effects, while its novelty and otherness mean policymakers face greater uncertainties, expanding the role of miscalculation and mistake.[80]

Systems dominated by positive feedback "are characterized by a self-impelled 'switch' or discontinuity between two extreme states."[81] Cyber conflict may be relatively stable now only because the tipping point has not yet been reached. After that, there may be a new, harsher reality — where there are more predators than prey — from which it will be hard to return.

It is understandable for the Department of Defense to pursue offense, which seems to have the advantage, as the best defense. But the cost of the new strategy of persistent engagement to suppress modest operations today may be the creation of even more aggressive and brazen adversaries tomorrow.

The Role of Surprise

Surprise is an important factor in our analysis of situational cyber stability and is worth exploring in more depth. There are no references to surprise in the most recent U.S. Department of Defense *Cyber Strategy*, nor in earlier versions dating back to

79 Robert Jervis, "Cooperation Under the Security Dilemma," 211.

80 As Martin Libicki put it: "Normal human intuition about how things work in the physical world does not always translate." From Martin C. Libicki, *Crisis and Escalation in Cyberspace* (Santa Monica, CA: RAND, Project Air Force, 2012), https://www.rand.org/pubs/monographs/MG1215.html.

81 William Gosling, *Helmsmen and Heroes: Control Theory as a Key to Past and Future* (London: Weidenfeld and Nicolson, 1994).

2006.[82] Military cyber doctrine has been similarly silent, other than unhelpfully saying that surprise is "germane."[83] The term is also lacking from U.K. cyber strategies and key NATO cyber documents.[84]

Scholars, fortunately, have covered surprise in more depth. Emily Goldman, John Surdu, and Michael Weaver were among the first to suggest that "Surprise probably plays a larger role in cyberspace than in any other domain."[85] Gartzke and Lindsay concluded that in cyber conflict, one element of surprise — deception — is more central than in other kinds of warfare: "[A]ttackers who fail to be deceptive will find that the vulnerabilities on which they depend will readily be patched and access vectors will be closed."[86] Buchanan, among others, focuses less on the likelihood of surprise than on its impact. States hide their operations and capability. To reduce surprise, adversaries must use intrusive cyber operations of their own. Such defensive espionage operations might be misread as (or indeed, repurposed for) a future surprise attack.[87]

James J. Wirtz unpacked the concept of a Cyber Pearl Harbor, which conjures "up compelling images of a 'bolt from the blue' surprise attack in American political and strategic culture," and which might induce "catastrophic paralysis rendering [the United States] unable to develop a military or politically effective response in wartime."[88] Goldman, Surdu, and Warner argue that:

> Conditions could entice an adversary to strike a similar, disabling blow against the United States in the hope of a quick victory that presents America with an undesirable strategic *fait accompli* with the possibility of removing the United States as an active opponent while inflicting minimal casualties or damage to U.S. forces ... The burden of escalation would then shift to U.S. policymakers, who would have to choose war over political compromise.[89]

Here, a surprise cyber attack would not be meant to be debilitating, but intended as a sharp jab to see if the adversary is actually serious about the geopolitical issue at stake. An attacker could also use a sudden cyber raid to "keep the victim reeling when his plans dictate he should be reacting,"[90] or it alternatively could be a *coup de main*, where the attack is the main effort to settle the military question. Other states would of course have a reciprocal fear of such attacks from the United States.

Lawrence Freedman suspects that this is overblown: "There is the question of what happens after the first blow. How would this turn into a lasting political gain?" Cyber troops only occupy virtual territory. Therefore "the victims would be expected to respond, even as they struggled to get the lights back on and systems working," even with a "classical military response."[91]

OUR ANALYSIS OF THE CHARACTERISTICS OF CYBERSPACE LEADS US TO WORRYING CONCLUSIONS. CYBER CONFLICT IS NOT MERELY DOUBLY DANGEROUS, BUT PERHAPS QUINTUPLY DANGEROUS...

Across this literature, "surprise" is often quite a broad and ill-defined term. We find five related meanings — different ways that "surprise" applies to situational cyber stability. First, *deception*, *con-*

82 U.S. Department of Defense, *The Department of Defense Cyber Strategy*, April 2015, https://archive.defense.gov/home/features/2015/0415_cyber-strategy/final_2015_dod_cyber_strategy_for_web.pdf; U.S. Department of Defense, "National Military Strategy for Cyberspace Operations (U)" (Office of the Chairman of the Joint Chiefs of Staff, December 2006), https://www.hsdl.org/?view&did=35693.

83 Joint Chiefs of Staff, "Joint Publication 3-12(R) Cyberspace Operations," Feb. 5, 2013, https://fas.org/irp/doddir/dod/jp3_12r.pdf.

84 UK Ministry of Defence, *Cyber Primer, 2nd Edition* (Swindon, UK: Ministry of Defence, Developments, Concepts and Doctrine Centre, July 2016), 100, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/549291/20160720-Cyber_Primer_ed_2_secured.pdf; NATO, "NATO Cyber Defense," December 2017, https://www.nato.int/nato_static_fl2014/assets/pdf/pdf_2017_11/20171128_1711-factsheet-cyber-defence-en.pdf.

85 Emily O. Goldman, John Surdu, and Michael Warner, "The Cyber Pearl Harbor: The Attacker's Perspective," in *Cyber Analogies*, ed. Emily O. Goldman and John Arquilla (Monterey, CA: Department of Defense Information Operations Center for Research, 2014), 29.

86 Erik Gartzke and Jon R. Lindsay, "Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace," *Security Studies* 24, no. 2 (April 3, 2015): 329, 326, https://doi.org/10.1080/09636412.2015.1038188.

87 Buchanan, *The Cybersecurity Dilemma*, 123–130.

88 James J. Wirtz, "The Cyber Pearl Harbor," *Intelligence and National Security* 32, no. 6 (2017): 7, https://doi.org/10.1080/02684527.2017.1294379.

89 Goldman, Surdu, and Warner, "The Cyber Pearl Harbor: The Attacker's Perspective," 26.

90 Richard K. Betts, *Surprise Attack: Lessons for Defense Planning*, 5.

91 Lawrence Freedman, "Beyond Surprise Attack," *The US Army War College Quarterly* 47, no. 2 (Summer 2017): 12, https://publications.armywarcollege.edu/pubs/3368.pdf.

cealment, and *trickery* are central to almost all cyber operations.[92] Second, cyber capabilities lend themselves to surprise because they can be *unexpected* or *unforeseen* as a new technological capability;[93] an unexpected target;[94] an unforeseen intensity, impact, or timing;[95] unforeseen trends;[96] and unexpected means.[97] Third, cyber conflict is frequently marked by being *sudden* or *fast*.[98] Fourth, they are frequently *audacious* or *daring*.[99] Lastly, but most important for stability, cyber capabilities are likely to be used to *attack early in a conflict*, even as an opening strike.[100] This is, after all, central to the Cyber Pearl Harbor concept.

Cyber capabilities can bypass fielded military forces to affect a nearly limitless range of an adversary's society, economy, and psychology.

Any theory or strategy which limits itself to a subset of these meanings of surprise is likely to fall short. Deception is more relevant to tactical cyber operations than escalation and stability. The middle three (*unexpected* or *unforeseen*, *sudden* or *fast*, *audacious* or *daring*) combine their effects to increase the danger of a spark, the first category of instability in which competition and conflict in cyberspace are the root causes of an acute geopolitical crisis. The last (*early use in conflict*) drives the escalation inversion, where cyber capabilities can accelerate the rush to war.

In most of the major cyber incidents to date, cyber defenders knew that such attacks were possible. After each, there have been experts who said something like, "Well, this shouldn't be a surprise. I've been saying for years it was bound to happen sometime." Indeed, Miller believes "a cyber surprise attack would be the least surprising of all the unsurprising 'surprise attacks.'"[101] As in almost all such attacks, "the striking thing ... is that in retrospect one can never quite understand" how the surprise ended up being quite so surprising.[102] Pearl Harbor was presaged by Port Arthur in 1904 and Taranto in 1940. Even defenders who can extrapolate from past trends are caught out by the specifics: the

92 An attacker's infrastructure will be scattered around the world, in other jurisdictions, obscuring the ultimate "return address." It is relatively easy to hide behind proxy groups and there is an entire class of attacks known as Trojan Horses.

93 Technical capabilities routinely deliver expected surprises, such as a new "zero-day vulnerability," a computer flaw known to the attackers but not defenders. Defenders know these exist but cannot guess the exact details. Technical capabilities can also be unexpected and quite massive, like the U.S.-Israeli joint operation to develop the complex Stuxnet malicious software, exquisitely designed to destroy Iranian centrifuges, the first-ever truly destructive attack.

94 Entertainment companies were surprised to find themselves in the crosshairs when the North Koreans deleted files at Sony Pictures and doxed (publicly released) embarrassing emails and the Iranians conducted a "destructive cyberattack" on the Sands Casino. Sony and North Korea. See, Sanger and Perlroth, "U.S. Said to Find North Korea Ordered Cyberattack on Sony"; for the Iranian attack on the Sands, see, Jose Pagliery, "Iran Hacked an American Casino, U.S. Says," *CNNMoney*, Feb. 27, 2015, https://money.cnn.com/2015/02/27/technology/security/iran-hack-casino/index.html.

95 Such as when China stole millions of government personnel records from the Office of Personnel Management, tens of millions of health records from Anthem, and credit records of over 140 million Americans from Equifax. The cyber security teams of those organizations likely knew, or suspected, they might suffer such intrusions, but could not know when, and of course hoped the impact would not be severe.

96 Such as adversaries building (and using) worrying capabilities far more quickly than expected, by using proxies or buying them from others. A 2009 assessment "wrote off" the North Korean hacking threat, underestimating its ability to quickly boost capabilities, while "Iran has boosted its cyber capabilities in a surprisingly short amount of time," in the words of one U.S. member of congress. For North Korea, see, David E. Sanger, *The Perfect Weapon: War, Sabotage, and Fear in the Cyber Age* (New York: Broadway Books, 2019), 129. For Iran, see, Peter Hoekstra, "Iran's Support for Terrorism Worldwide," Testimony Before the House Committee on Foreign Affairs, Joint Subcommittee Hearing, March 4, 2014, https://www.govinfo.gov/content/pkg/CHRG-113hhrg86958/html/CHRG-113hhrg86958.htm.

97 Stealing emails for espionage purposes was expected by the United States and considered fair game but releasing emails to influence the 2016 presidential elections was never seriously considered. Susan B. Glasser, "Ex-Spy Chief: Russia's Election Hacking Was an 'Intelligence Failure,'" *Politico*, Dec. 11, 2017, https://www.politico.eu/article/ex-spy-chief-russias-election-hacking-was-an-intelligence-failure/.

98 Though planning a sophisticated campaign may take months or years, individual cyber operations occur at "network speed" or at or near the "speed of light." The 2003 SQL Slammer worm spread so quickly that only fifteen minutes after infecting its first computer, "huge sections of the Internet began to wink out of existence." Paul Boutin, "Slammed!," *Wired*, July 1, 2003, https://www.wired.com/2003/07/slammer/.

99 The U.S. government simply did not believe "the Russians would dare to leap the Atlantic and apply" its cyber techniques "to an election in the United States." Sanger, *The Perfect Weapon*, XVIII.

100 Israel apparently used cyber means to hide their strike aircraft from Syrian radars during Operation Orchard in 2007. The Department of Defense has reported that Chinese military writings "advocate targeting an adversary's C2 [command and control] and logistics networks to affect its ability to operate during the early stages of conflict." Fred M. Kaplan, *Dark Territory: The Secret History of Cyber War* (New York: Simon & Schuster, 2016), 160-161; Office of the Secretary of Defense, "Annual Report to Congress: Military and Security Developments Involving the People's Republic of China 2016," April 26, 2016, https://dod.defense.gov/Portals/1/Documents/pubs/2016%20China%20Military%20Power%20Report.pdf.

101 James N. Miller, personal communication to the authors, May 29, 2020.

102 Ephraim Kam, *Surprise Attack: The Victim's Perspective* (Cambridge, MA: Harvard Univ. Press, 2004), 2.

who, when, where, how, and how bad. Surprise in cyberspace will be more destabilizing than in other domains, for four reasons.

We have already fully explored the first two reasons. The *dynamics of cyber conflict* lend themselves to surprising uses across all five meanings of surprise. They rely on deception and trickery; enable the unexpected and unforeseen; are sudden and fast, audacious and daring; and especially useful early in a conflict. There are also significant *first-use pressures*, as they may make a security dilemma quintuply dangerous. Because cyber capabilities are not easily observable, it is extremely difficult to assess an adversary's order of battle or relative strengths, or to detect the equivalent of tanks massing on the border. Any particular attack might have an asymmetric impact, keeping defenders on perpetual and exhaustive high alert.

There is also a *nearly limitless realm of the possible*. Cyber capabilities can bypass fielded military forces to affect a nearly limitless range of an adversary's society, economy, and psychology. The pace of innovation and dependence creates countless paths to attain technical surprise and the use of "existing weapons and forces in new and different ways."[103] Even more so than in other kinds of intelligence warning, "[t]here are few limits on what can be imagined," so defenders have less chance of assessing where a blow may strike.[104] Because everything is interconnected and deeply dependent, cyber capabilities offer an attacker more opportunities to shift the correlation of forces in their favor. Some experts assert that "[c]yber attack does not threaten crippling surprise or existential risk," as past attacks only disrupted computer components which can be replaced relatively quickly.[105] Yet this misses the scope of potential future cyber attacks. With the Internet of Things and cyber physical systems, attacks now impact electrical grids, pipelines, and dams, objects made of concrete and steel. The potential impact of and opportunities for surprise attacks will soar in unappreciated ways.

There is lastly a high potential for *mistake and miscalculation*. The novel nature of cyber attacks means adversaries are likelier to misjudge how their operations will be perceived by the recipient. The attacker might believe their attack is within the norms, justified because it is a tit-for-tat reprisal, or similar to past operations which were met with indifference. Cyber attacks are likely to flop (or worse, messily cascade) if they are not backed by meticulous intelligence, careful planning, and extensive testing — though these only reduce, rather than eliminate the risks. Mistakes can take the adversary (and indeed, the attacker) by surprise, as this happened to the North Koreans and Russians with WannaCry and NotPetya.[106]

During the Cuban Missile Crisis, U.S. Navy commanders "kept down" nuclear-armed Soviet submarines with depth charges, even at the height of the crisis, because that was the established, doctrinally correct procedure.[107] This nonchalant aggression, based on a standard operating procedure approved in more peaceful times, complicated U.S.-Soviet signaling and courted thermonuclear disaster. Before the peak of the next Cuban Missile Crisis-style emergency, each state will be aggressively burrowing into each other's networks for advantage. Those cyberspace teams — often proxies or groups operating loosely under a command hierarchy — will have even more operational leeway to punch and counterpunch than the U.S. Navy commanders of the 1960s. A large number of tactical commanders, often not under strict command and control, can unleash dangerous cyber capabilities and might be itching for a fight more than their seniors. Any mistake, by any side, might prompt an escalation that is unexpected and unwanted by the leadership of either sides. The tempo of the situation can take on a life of its own, leaving less room for strategic choices.

Lessons for Stability

During relative peacetime, cyber conflict should continue to operate as a pressure release. However, at some point in the future, cyber capabilities will be the root cause of a major geopolitical crisis, through mechanisms of Spark, Pull Out the Big Guns, or the Escalation Inversion. States will engage in riskier behavior during crises, either because the stakes of the game remove their earlier inhibitions or because they will act to get their cyber strike in before the real shooting starts. From

103 Michael S. Goodman, "Applying the Historical Lessons of Surprise Attack to the Cyber Domain: The Example of the United Kingdom," in *Cyber Analogies*, ed. Emily O. Goldman and John Arquilla (Monterey, CA: Department of Defense Information Operations Center for Research, 2014), 7.

104 Robert Jervis, *Why Intelligence Fails: Lessons from the Iranian Revolution and the Iraq War* (Ithaca, NY: Cornell University Press, 2010), 129.

105 James Andrew Lewis, *Rethinking Cybersecurity: Strategy, Mass Effect, and States* (Washington, DC: Center for Strategic & International Studies: Rowman & Littlefield, 2018), 1.

106 Ellen Nakashima, "Russian Military Was Behind 'NotPetya' Cyberattack in Ukraine, CIA Concludes," *Washington Post*, Jan. 12, 2018, https://www.washingtonpost.com/world/national-security/russian-military-was-behind-notpetya-cyberattack-in-ukraine-cia-concludes/2018/01/12/048d8506-f7ca-11e7-b34a-b85626af34ef_story.html.

107 Graham Allison and Philip Zelikow, *Essence of Decision: Explaining the Cuban Missile Crisis*, 2nd edition (Reading, MA: Longman, 1999).

this analysis, we draw important lessons for stability across three areas.

1) New Models Are Required for Stability in Cyberspace

Stability and escalation in cyberspace work differently. To adapt to cyber situational stability, the existing language and models used by the national security community and international relations are insufficient and should be avoided, treated cautiously, or reconceptualized altogether.

Don't Rely on "Ladders of Escalation."

Herman Kahn introduced "ladders of escalation": a hierarchical ranking of a set of actions and responses to understand the relationship of conventional and nuclear war.[108] The concept does not translate well to cyber conflict. Indeed, as Rebecca Hersman has written, the entire "new era of strategic competition" will be less predictable due to "intrusive digital information technologies, advanced dual-use military capabilities, and diffused global power structures" which will open "alternative and less predictable escalatory pathways."[109] Cyberspace underpins every aspect of modern society and economy. Cyber escalation ladders will have such narrow bounds that a cunning adversary can find plenty of asymmetric vectors of aggression. There is not just one ladder, but many — if adversaries cannot escalate on one, they can jump horizontally to another.

Reduce One-Sided Knowledge

During the Cold War, Soviet military moves and capabilities were closely guarded secrets in the West, but relative government transparency and a free press ensured that the United States and NATO were open books in comparison. In cyber conflict, attacks from China, Russia, Iran, and North Korea are regularly splashed across the news, while those of the United States remain heavily classified.

Of the roughly 1.2 million people in the U.S. government who hold at least a top-secret security clearance, probably only a few dozen people — in the National Security Council, Department of Defense, and Intelligence Community — know the totality of U.S. operations against a particular adversary and its own operations against the United States.[110] When there is a leak about U.S. capabilities and operations, U.S. government personnel with clearances are forbidden to look, meaning they may actually know *less* about U.S. operations than their adversaries or the informed public.[111]

American adversaries end up in a similar place but by a different path, as their governments typically have less strict controls over their cyberspace forces.[112] Their leadership may only have a dim sense of what malfeasance is being done ostensibly on their nation's behalf, but they likely still have their own national security experts and cyberspace defenders regaling them with tales of horror of what the United States is suspected of doing.

Accordingly, it is especially hard to develop a balanced, objective, or common understanding of the rights and wrongs, moves and countermoves. Cause and effect become nearly impossible to distinguish. There are few recommendations here other than unilateral ones. The national security community must declassify and break down compartments to combat cognitive bias. The current situation — yelping about the adversary's punches but classifying one's own — is not tenable, leading to a biased view of cyber conflict that is poisonous in an open democracy. The U.S. transparency over Operation Glowing Symphony, the cyber campaign against the Islamic State, is an astounding case study in openness.[113] But more should be done with respect to operations directed against state adversaries who can shoot back, like Iran.

2) Missing Mechanisms for Stability That Must Be Developed

The risks of accidental or inadvertent escalation in situational cyber stability require an emphasis on signaling, firebreaks, and off-ramps to

108 Nadiya Kostyuk, Scott Powell, and Matt Skach, "Determinants of the Cyber Escalation Ladder," *The Cyber Defense Review* 3, no. 1 (Spring 2018): 123–34, https://cyberdefensereview.army.mil/Portals/6/Documents/CDR%20Journal%20Articles/Determinants%20of%20the%20Cyber_Kostyuk_Powell_Skach.pdf?ver=2018-07-31-093725-923.

109 Rebecca Hersman, "Wormhole Escalation in the New Nuclear Age," *Texas National Security Review* 3, no. 3 (Summer 2020): 90–109, https://tnsr.org/2020/07/wormhole-escalation-in-the-new-nuclear-age/.

110 Office of the Director of National Intelligence, "Fiscal Year 2017 Annual Report on Security Clearance Determinations," August 2018, https://www.dni.gov/files/NCSC/documents/features/20180827-security-clearance-determinations.pdf.

111 Ryan Gallagher, "U.S. Military Bans The Intercept," *The Intercept*, Aug. 20, 2014, https://theintercept.com/2014/08/20/u-s-military-bans-the-intercept/.

112 Iran and Russia, for example, use proxies extensively. See, Tim Maurer, *Cyber Mercenaries: The State, Hackers, and Power* (New York: Cambridge University Press, 2018).

113 Dina Temple-Raston, "How The U.S. Hacked ISIS," *NPR*, Sept. 26, 2019, https://www.npr.org/2019/09/26/763545811/how-the-u-s-hacked-isis .

deal specifically with cyber conflict. These must feature more prominently within policies, strategies, and projects.

Lack of Effective Signaling

It is particularly difficult in the cyber arena to signal resolve, intent, or displeasure because there are few accepted rules and no clear escalation ladders.[114] There is little direct communication between major rivals. The mechanisms are either low-level and technical or high-level and political. While helpful, neither is routine, timely, or useful for operational signaling.

China's leadership is still incensed over the U.S. indictment of five army cyberspace officers and has banned military-to-military contacts.[115] While the U.S.-Russian "cyber hotline" does connect the White House with the Kremlin, this is useful only for sending political messages, not for managing fast-moving crises. To punish Russia's invasion of Ukraine, the U.S. Congress outlawed more operationally relevant military-to-military contacts.[116] The United States does maintain direct links between the Department of Homeland Security and its Chinese and Russian counterparts, but these are more useful for exchanging technical information between computer emergency response teams.[117]

Even in the best case, the U.S. government may know the signal it is sending but cannot be sure of the signal being received. Feedback to avert and minimize crises will be delayed, unclear, and not relayed directly between the key participants until new hotlines are created, or substituted with back-channel conversations by former policymakers and flag-level officers. These efforts must be lavishly funded — and will still be comparatively cheap — as a powerful negative feedback hedge to a more aggressive persistent engagement.

Difficulty Reaching Global Norms

International norms of behavior for cyber conflict will always be problematic: General principles have huge loopholes and can be ignored by states seeking advantages, while specific norms can usually be circumvented. Many destabilizing, brazen, and reckless attacks have not violated the letter of U.S. norms.[118] Neither the North Korean attack on Sony Motion Pictures in 2014 nor the Russian interference in U.S. elections in 2016 technically violated the stated U.S. norm proscribing attacks on "critical infrastructure." In other cases, it seems that the United States wants norms for thee but not for me. Chinese espionage into the Office of Personnel Management should have been unobjectionable per U.S. statements. It was, in the words of a former head of CIA and NSA, "honorable espionage work" as the office was a "legitimate foreign intelligence target."[119] But the Obama administration still decided to "retaliate."[120]

We believe there is little prospect for norms that are specific, binding, *and* global. Policymakers should instead push for a set of norms that attains at least two of these criteria, while collectively building towards a solution with all three. For example, the 2019 "Joint Statement on Advancing Responsible State Behavior in Cyberspace" brought together 27 like-minded Western democracies to call out specific norms and "work together on a voluntary basis to hold states accountable when they act contrary to this framework [because] there must be consequences for bad behavior in cyberspace."[121]

Defense Is Likely the Best Defense

Attackers have the advantage. It takes a varsity defense to defeat a team of junior varsity attackers. If the attackers bring their own varsity team,

114 See, Ben Buchanan, *The Hacker and the State: Cyber Attacks and the New Normal of Geopolitics* (Cambridge, MA: Harvard University Press, 2020), especially Chapter 9, for more on the difficulties, using the Russian disruption of Ukrainian electrical grid as a case study.

115 From author's experience in discussions with Chinese officials in track 1.5 discussions led by the Center for Strategy and International Security.

116 AFP, "Ukraine Crisis: US Suspends Military Cooperation with Russia," *The Telegraph*, March 4, 2014, https://www.telegraph.co.uk/news/worldnews/europe/ukraine/10674777/Ukraine-crisis-US-suspends-military-cooperation-with-Russia.html .

117 Sean Gallagher, "US, Russia to Install 'Cyber-Hotline' to Prevent Accidental Cyberwar," *arsTechnica*, June 18, 2013, https://arstechnica.com/information-technology/2013/06/us-russia-to-install-cyber-hotline-to-prevent-accidental-cyberwar/.

118 Jason Healey and Tim Maurer, "What It'll Take to Forge Peace in Cyberspace," *CSM Passcode*, March 20, 2017, https://www.csmonitor.com/World/Passcode/Passcode-Voices/2017/0320/What-it-ll-take-to-forge-peace-in-cyberspace.

119 General Michael Hayden quoted by Julian Hattem, "Ex-CIA Head: 'Shame on Us' for Allowing Government Hack," *The Hill*, June 16, 2015, https://thehill.com/policy/national-security/245101-ex-cia-head-shame-on-us-for-allowing-government-hack.

120 David E. Sanger, "U.S. Decides to Retaliate Against China's Hacking," *New York Times*, July 31, 2015, https://www.nytimes.com/2015/08/01/world/asia/us-decides-to-retaliate-against-chinas-hacking.html.

121 U.S. Department of State, "Joint Statement on Advancing Responsible State Behavior in Cyberspace," September 23, 2019, https://www.state.gov/joint-statement-on-advancing-responsible-state-behavior-in-cyberspace/.

the defenders need to have an all-star defense to have a chance. And if the attackers are themselves an all-star team, then few organizations in the world have much chance. There is certainly a role for the new U.S. push for persistent engagement and defending forward. When Russian cyber operatives are disrupting the opening ceremony of the Olympic Games and North Koreans conduct cyber bank heists around the world, it seems disingenuous to badmouth U.S. countermeasures as being escalatory.[122] It is destabilizing, however, to elevate the operational concept of persistent engagement to a strategy, given the likelihood of destabilizing positive feedback.

A better option is for policymakers to prioritize defense and reverse attacker advantage though "leverage." The New York Cyber Task Force analyzed five decades of "technology, operational, and policy innovations which most advantage the defender" and concluded a more defense-advantage cyberspace is possible with technical solutions that can scale across the entire internet (rather than just one enterprise at a time) and fresh investment in operational and process innovations.[123]

If cyberspace were more advantageous to the defender, many of the most destabilizing dynamics would lose force with higher barriers to entry, leading to fewer capable adversaries and fewer serious attacks. Since fewer attacks might be catastrophic, the pressure for counter-offensive operations would be diminished with more room for agreement and norm building.

Measurement

This article has summarized research on cyber stability and instability, escalation and de-escalation. Almost none of this research is based on significant measurements of what actions lead to what responses over time. Previous work co-authored by one of us with Neil Jenkins has proposed several frameworks to measure if persistent engagement is correlated with changes in adversary behavior:

> The advocates of persistent engagement and deterrence suggest it should have a substantial, perhaps unprecedented impact on adversary behavior. Anything other than a correspondingly strong reduction [in such behavior] suggests that the policy may not be working as intended. If the trend significantly worsens, it may be that a hypothesis that the new policy is inciting adversaries is a better fit to the curve.[124]

Such measurements need only be concerned with the direction and magnitude of the vector: Is adversary behavior changing — or cyberspace becoming more stable or instable — and how fast? Categorizing and tracking these over time would be inexpensive and a worthy investment.

3) Hedge Against Cyber Surprise

Military surprise in the initial phase of war usually succeeds, especially against the United States.[125] Our colleague, Dick Betts, wrote 35 years ago: "Some other problems may be more important [than preparing for surprise attack] but most of them are better understood."[126] Research by academics and attention from military professionals, intelligence officials, and policymakers to understand and counter the role of surprise cyber attack will have a low cost but high payoff.

The Detection and Attribution Gaps

During the Cold War, both sides were wary of the danger of a surprise nuclear attack. It was then and is still now stabilizing for each nation to possess a secure second-strike capability, as neither nation needs to worry quite as much about a debilitating first strike, and for each side to have capabilities to rapidly and reliably detect missile launches. The nuclear warfighters of the Strategic Air Command (and presumably their brethren in the Soviet Union) may not have liked the reduction of operational surprise, but the need for stability meant policymakers had an easy time overruling their concerns. Such "national technical means" were critical to stability and arms control and both nations agreed to have "open skies" to one another, allowed observation of major ex-

122 For Russia's attacks on the Olympic games, see, Greenberg, "Inside Olympic Destroyer, the Most Deceptive Hack in History"; for the North Korean heists, see, Ben Buchanan, "How North Korean Hackers Rob Banks Around the World," *Wired,* Feb. 28, 2020, https://www.wired.com/story/how-north-korea-robs-banks-around-world/.

123 New York Cyber Task Force, "Building a Defensible Cyberspace," Columbia University, School of international and Public Affairs, Sept. 28, 2017, https://sipa.columbia.edu/ideas-lab/techpolicy/building-defensible-cyberspace.

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

125 Betts, *Surprise Attack*, 4.

126 Betts, *Surprise Attack*, 3.

ercises, and reported major troop presence and movements in Europe.

None of these stabilizing factors apply to cyber conflict. The value of cyber operations, and the critical need for them to stay unobserved and covert, means steps to improve mutual visibility are impractical. Because the primary use of cyber capabilities today is espionage, mutually-beneficial surveillance is impossible, leaving weaker powers feeling distinctly insecure. For example, one reason China may have difficulty agreeing to cyber norms is China's weak attribution vis-a-vis the perceived strength of the U.S. government and commercial intelligence expertise.[127]

Here it is far from clear what practical recommendations to make. It is unthinkable that the United States might, in the name of stability, assist China to boost its attribution capabilities to better detect U.S. cyber operations. Nor is it feasible for Russia and the United States a develop virtual "open skies" to freely transit each other's networks.

Reduce the Probability of Surprise

The United States must act to reduce the *probability* of surprise. Increased intelligence and warning are useful but not game-changers unless the intelligence is particularly exquisite, such as persistent access to adversaries' networks. Such dominance is expensive, fleeting, and adds its own destabilizing pressure. More useful gains can be had by expanding defenders' imaginations and experience through exercises, experimentation, and curiosity about future forms of cyber conflict.[128]

U.S. and allied militaries must recognize that an initial surprise attack is both likely to occur and likely to succeed. And since nonstate actors "possess a greater range of capabilities than at any time in history," and cyber security and technology companies routinely and agilely respond to critical threats, those strategies and doctrines must include cooperative response to deal with surprise.[129] If the United States wants stability, and not merely superiority, then Russia and China (and, to a much lesser degree Iran and North Korea) should also have less fear of a surprise cyber attack.

Reduce the Impact of Surprise

The United States and its cyber adversaries work hard to avoid surprise attacks while simultaneously maximizing their own ability to carry out surprise attacks on foes. This is a solid policy in a stable environment but it is exceptionally risky in an unstable one. Perhaps the only way to meaningfully slice though this dilemma is through the "defense is the best defense" approach discussed above. The United States, the European Union, and China could cooperate to change the physics of the internet through new standards and engineering. This would stabilize the entire system, reducing the ability to surprise and the gains to be had. It would reduce their own offensive capabilities some but potentially drastically reduce those of criminal actors, Iran, North Korea, and third-tier adversary powers.

Secure Cyber, Space, and Strategic Systems

The most dangerous temptation is for a state to believe it can blind or disarm its rival's cyber capabilities, space systems, or nuclear weapons/command and control. States must spend resources to secure those systems most essential to great-power deterrence and strategic stability. The U.S. Defense Science Board proposed a cyber-resilient "thin-line" of strategic forces to reduce the impact of surprise attack.[130] As Jim Miller shared with us, cyber resilience "may be as important as dispersing bombers and deploying Polaris were in the early days of the Cold War."[131] Securing even a slice of space-based intelligence and warning systems reduces the temptation for a surprise attack. Space, strategic, and cyberspace forces do not need to be 100 percent resilient, just secure enough that an attacker could not have a realistic hope of a disarming attack.

127 Scott Warren Harold, Martin C. Libicki, and Astrid Stuth Cevallos, *Getting to Yes with China in Cyberspace* (Santa Monica, California: RAND, 2016), https://www.rand.org/content/dam/rand/pubs/research_reports/RR1300/RR1335/RAND_RR1335.pdf.

128 Jason Healey, "What Might Be Predominant Form of Cyber Conflict?," in *2017 International Conference on Cyber Conflict (CyCon U.S.)* (Washington, DC: IEEE, 2017), 36–44, https://doi.org/10.1109/CYCONUS.2017.8167511.

129 On general nonstate capabilities in national security, see, Barry Pavel, Peter Engelke, and Alex Ward, *Dynamic Stability: US Strategy for a World in Transition* (Washington, DC: Atlantic Council Strategy Papers, March 2016), 17. https://www.atlanticcouncil.org/wp-content/uploads/2015/04/2016-DynamicStabilityStrategyPaper_E.pdf. For a cyber strategy built around nonstate entities and capabilities, see Jason Healey, *A Nonstate Strategy for Saving Cyberspace* (Washington, DC: Atlantic Council Strategy Papers, 2017), https://www.atlanticcouncil.org/wp-content/uploads/2015/08/AC_StrategyPapers_No8_Saving_Cyberspace_WEB.pdf.

130 U.S. Department of Defense, *Task Force on Cyber Deterrence*, 44; U.S. Department of Defense, Defense Science Board, *Task Force Report: Resilient Military Systems and the Advanced Cyber Threat* (Washington DC, January 2013), https://defenseinnovationmarketplace.dtic.mil/wp-content/uploads/2018/04/Resilient_Military_Systems_Cyber_Threat.pdf.

131 James N. Miller, personal communication to the authors, May 29, 2020.

Next Steps for Situational Cyber Stability

In the film comedy *Zoolander*, a group of not-too-bright male models have a gasoline fight at a filling station. Everyone watching is in on the joke: It is only a matter of time before one of these imbeciles, oblivious to the danger, lights a match. The punchline, a massive fireball, is a surprise to no one.

We hope this analogy to cyber conflict remains a silly one — there is no comparison to states playing a dangerous game, soaked in vulnerabilities and being complacent that no one will light up. But the dynamics of cyber conflict drive nearly all states to be greedy, expansionist powers. Every adversary is deeply vulnerable and obeying broadly the same imperatives — to collect intelligence, lay the groundwork for future attacks, and seize terrain in cyberspace to contest an adversary's operations — and assuming all others are maximally doing the same.[132] This competition is not carried out over physical territory, but over network infrastructure and information, owned by the private sector and the lifeblood of modern economy and society. This drives positive feedback, possibly spiraling out of the willful control of the participants.[133]

If states are frustrated in the competition to achieve meaningful strategic gains through cyberspace, this may just fuel additional escalation. Each side will go back to their legislatures or paymasters, asking for a larger budget and looser rules and pointing to the other side's newly aggressive forward defense as proof of their intransigence. Since each side views the other as aggressive, there is "no reason to examine one's own policies," nor is there a "need to make special efforts to demonstrate willingness to reach reasonable settlements."[134] If concessions will not alter the other's actions, then restraint can seem a fool's choice — unless everyone is soaking in gasoline.

Stability and restraint may not be likely unless adversaries seek stability and act with restraint. This will be particularly hard now that the participants are engaged in relentless, persistent engagements. Conflict (especially conflict that can never really end, like that in cyberspace) can lead to heightened emotions, unwillingness to compromise, and self-righteousness.[135] The United States believes, probably rightly, that it has showed restraint by eschewing large-scale disruptive operations or espionage for commercial gain, and sets great store in how this restraint highlights U.S. interests for a peaceful cyberspace. But these self-imposed limits have been overshadowed by near-limitless political-military espionage. American claims that its "pervasive, persistent access on the global network" is "just espionage" fall flat.[136] Adversaries (and allies) could be forgiven for doubting U.S. restraint, given their existential dependence on technology largely invented and created in a country seeking to bask in lasting cyberspace pre-eminence.

The technology community has been concerned about Balkanization of the internet — what was once unified is now split by national borders like China's Great Firewall.[137] But cyberspace is also being Balkanized in another sense, in that those involved are incapable of forgetting or forgiving insults they have suffered from others and blind to those they themselves have inflicted. Such long and selective memories are likely to be as destabilizing in the virtual world as in the real. For this and related reasons, "state's strategic responses should not be cyber operations," but rather sanctions, indictments, trade and immigration restrictions, or other levers of power.[138]

In the face of situational cyber awareness, the long-term goal might go beyond stability to order — an order which players accept out of their own interest rather than through the pressure of a hegemon. For Russia and China to buy into such an order, it would need to include limits on cross-border flows of information and internet content. Such controls are hard to reconcile with traditional liberal democratic practice, though the trans-Atlantic political pressure on companies like Facebook and Twitter to better police hate speech, terrorists, trolls, and foreign political meddling may make such a grand bargain more palatable in the future.

If the United States wants a universal order, accepted by friends and rivals alike, it will have to make very serious compromises. U.S. decision-makers may decide, either positively or through inaction, that they are unwilling to make such compromises, so that for the duration of the digital age, the United States will have to enforce its preferences through power. Many people, and not just hawks, will accept this bargain

132 See especially, Buchanan, *The Cybersecurity Dilemma*, 113-116.

133 Robert Jervis, *Perception and Misperception in International Politics* (Princeton NJ: Princeton University Press, 1976), 64–65.

134 Jervis, *Perception and Misperception in International Politics*, 353.

135 For more on the role of emotions in cyber conflict, see Rose McDermott, "Some Emotional Considerations in Cyber Conflict," *Journal of Cyber Policy* 4, no. 3 (2019): 309–25, https://doi.org/10.1080/23738871.2019.1701692.

136 Joanna Walters, "NSA 'Hacking Unit' Infiltrates Computers Around the World – Report," *The Guardian*, Dec. 29, 2013, https://www.theguardian.com/world/2013/dec/29/der-spiegel-nsa-hacking-unit-tao.

137 "A Virtual Counter-Revolution," *The Economist*, Sept. 2, 2010, https://www.economist.com/briefing/2010/09/02/a-virtual-counter-revolution.

138 Buchanan, *The Cybersecurity Dilemma*, 183.

gladly. But if cyberspace encourages positive feedback it is unlikely to survive the conflict in anything like its form today. At the very least, the United States should acknowledge that adversaries see U.S. actions and preferences as destabilizing (at least to their own domestic orders). U.S. policy should aim not to combat adversaries but rather destabilization itself. Stability should be the goal and not a side benefit expected from unending confrontation.

In many ways, cyber capabilities possess dynamics opposite to those of nuclear weapons.[139] By radically decreasing the cost of war, even to a state with significant relative disadvantages, cyber capabilities can drastically change world politics.

Jason Healey *is a senior research scholar at Columbia University's School for International and Public Affairs, specializing in cyber conflict and risk. He started his career as a U.S. Air Force intelligence officer, before moving to cyber response and policy jobs at the White House and Goldman Sachs. He was founding director for cyber issues at the Atlantic Council where he remains a senior fellow and is the editor of the first history of conflict in cyberspace,* A Fierce Domain: Cyber Conflict, 1986 to 2012. *He is on the DEF CON review board and served on the Defense Science Board task force on cyber deterrence.*

Robert Jervis *is the Adlai E. Stevenson professor of international politics at Columbia University. In addition to his recent* How Statesmen Think, *he is the author of eight other books and over 200 articles. He was president of the American Political Science Association from 2000 to 2001 and is a member of the American Philosophical Society and corresponding member of the British Academy.*

This work was supported by the Office of Naval Research under the OSD Minerva program, grant number N00014-17-1-2423

Photo: NSA Image

139 Robert Jervis, *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon* (Ithaca, NY: Cornell University Press, 1989). See also, Robert Jervis, "Author Response: Reflections on The Meaning of the Nuclear Revolution, 30 Years Later," *Texas National Security Review*, April 30, 2020, http://tnsr.org/roundtable/book-review-roundtable-the-meaning-of-the-nuclear-revolution-30-years-later/.