Cyber Power and the International System

Shawn William Lonergan

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2017

www.manaraa.com

ProQuest Number: 10620131

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 10620131

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

www.manaraa.com

© 2017
Shawn William Lonergan
All rights reserved

www.manaraa.com

ABSTRACT

Cyber Power and the International System

Shawn William Lonergan

This dissertation is comprised of three separate papers that address how cyber power contributes to national power and the implications for international security posed by cyber operations. The first paper, "Cyber Power and International Stability: Assessing Deterrence and Escalation in Cyberspace," posits that there are unique attributes that define the cyber domain and that have direct implications on deterrence and escalation dynamics between state actors. The second paper, "Arms Control and Confidence Building Measures for the Cyber Domain," explores at various mechanisms that states have traditionally used to foster stability and prevent inadvertent conflict and assesses their applicability to controlling cyber operations. Finally, "The Logic of Coercion in Cyberspace" delves into the role of cyber operations as both inadvertent and deliberate signals and assesses their utility as a coercive instrument of statecraft.

www.manaraa.com

Table of Contents

www.manaraa.com

List of Charts, Graphs, Illustrations

Tables

Diagrams

www.manaraa.com

Acknowledgements

This dissertation greatly benefited from the writings, teachings, and guidance of Richard Betts and Robert Jervis, who served on my dissertation committee. Jason Healey, Jeffrey Lax, and Scott Silverstone also provided insightful comments during the defense, which are reflected in the final submission. I am also grateful for the support and candid thoughts of Lieutenant General Paul Nakasone, Colonel Suzanne Nielsen, and Colonel (Ret.) Gregory Conti provided throughout the duration of this project. Additionally, though they may not have realized it at the time, my long conversations (and occasional arguments) about cyber conflict with Brian Blankenship, Theo Milonopoulos, and Blake Rhoades immensely helped with theory development. Finally, I must acknowledge my co-author, Erica Borghard, who, despite my Pacific Northwest upbringing, is the one who helped me separate the trees from the forest.

www.manaraa.com

To my parents, Cheryl and Frank

www.manaraa.com

Cyber Power and International Stability:

Assessing Deterrence and Escalation in Cyberspace

Shawn W. Lonergan

Introduction

Although the cyber domain in general, and cyber conflict in particular, are not "new," academia is only at a nascent stage in exploring and understanding how many of the core theories of international security apply to state behavior in cyberspace. Perhaps the most fundamental question posed by the emergence of a new domain of strategic and military competition between states is the extent to which it contributes to, or undermines, international stability. During the Cold War, the two interrelated theoretical concepts that were central to the notion of systemic stability or instability were deterrence and escalation. On the one hand, mutual deterrence through the reciprocal threat (and fear) of catastrophic strategic nuclear war was hypothesized to inject stability into relations between the two superpowers. On the other hand, escalation dynamics risked undermining the very stability achieved through mutual deterrence—deliberate escalation as encapsulated in Herman Kahn's escalation ladder envisioned a world in which states could deliberately escalate to the use of nuclear weapons, and theories of inadvertent escalation hypothesized how states could unintentionally stumble into nuclear war.

This paper explores the stability of the cyber domain using the lenses of deterrence and escalation theory. In particular, I assess the extent to which deterrence and escalation theories are applicable to cyberspace and, in doing so, identify the attributes of the cyber domain that differ from other domains of warfare and that have theoretically significant implications for how

www.manaraa.com

scholars should think about deterrence and escalation. The prevailing assumption in the nascent literature on cyber conflict is that deterrence is incredibly difficult in the cyber domain and, therefore, it is likely to be defined by escalatory spirals. However, in my analysis, I find that traditional conceptions of both deterrence and escalation are problematic and that, while deterrence by punishment is hard, deterrence by denial is possible. Moreover, I find that the domain is also not as escalatory as the literature suggests. My divergent conclusions are in large part driven by the fact that, unlike the vast majority of scholars, I do not find the cyber domain to be offense dominant—particularly due to the target-specific nature of offensive cyber operations—and the destructive potential of cyber conflict is relatively minimal. Thus, while deterrence and escalation may be flip sides of the same coin in the context of the traditional security studies literature, I assert that the same is not the case in cyberspace.

This paper proceeds as follows. First, I introduce the three distinct, theoretically important characteristics of cyberspace that carry implications for how we can understand the application of deterrence and escalation to the domain. Then, I assess three different deterrence logics: punishment, denial, and cross domain. The paper then transitions to explore both inadvertent and deliberate escalation dynamics in cyberspace.

Three Unique Attributes of Cyberspace

Three attributes distinguish the cyber domain from others in ways that have implications for the logic and feasibility of deterrence and escalation. In particular, I focus on a comparison between cyber and nuclear weapons due to the context in which deterrence and escalation theories were developed.[1] The greatest similarities between nuclear and cyber weapons are that

[1] While I also consider the applicability of conventional deterrence to the cyber domain, my comparison of weaponry focuses on nuclear versus cyber arms due to the overwhelming influence of the nuclear deterrence literature and the conceptual difficulties associated with

www.manaraa.com

both possess tactical as well as strategic utilities; and the logic of offense-defense theory appears to have relevance to the two fields, as will be explored in multiple parts of this paper.[2] However, there are important differences between these weapons that stem from the nature of their employment, particularly offensive applications, and their respective capacities for destruction.[3]

---

comparing an entire class of armaments (e.g., conventional) with a specific type of weapon. Several scholars have made the nuclear-cyber comparison before. For instance, see Stephen J. Lukasik, "A Framework for Thinking about Cyber Conflict and Cyber Deterrence with Possible Declaratory Policies for These Domains," and Patrick M. Morgan, "Applicability of Traditional Deterrence Concepts and Theory to the Cyber Realm," in *Proceedings of a Workshop on Deterring Cyber Attacks: Informing Strategies and Developing Options for U.S. Policy*, National Research Council (Washington, DC: The National Academies Press, 2010); and Joseph S. Nye Jr., "Nuclear Lessons for Cyber Security?," *Strategic Studies Quarterly* 5, no. 4 (Winter 2011).

[2] Offense-Defense Theory purports that the probability of interstate conflict increases when there is the perception by state actors that offensive capabilities have an advantage over defensive ones. Furthermore, the theory suggests that the probability of conflict decreases and the probability of cooperation between states increases when defense is dominant. For further reference see Robert Jervis, "Cooperation Under the Security Dilemma," *World Politics* 30, no. 2 (1978): 167-214; and Stephen Van Evera, "Offense, Defense, and the Causes of War," *International Security* 22, no. 4 (Spring 1998): 5-43. For a specific reference to the offense-defense balance and cyber see, Ilai Saltzman, "Cyber Posturing and the Offense-Defense Balance," *Contemporary Security Policy* 34, no. 1 (2013): 40-63; and Keir Lieber, "The Offense-Defense Balance and Cyber Warfare," *Cyber Analogies* 96, no. 107 (2014).

[3] There are other important differences between cyber and nuclear weapons. However, those not identified do not carry the same level of significance on the theoretical implications of the analysis presented in this paper as the ones stated. For instance, nuclear weapons have only been offensively employed twice in human history—the atomic bombings against Japan at the close of World War II—compared to a plethora of publicly-known significant cyber attacks, which likely represent only a subset of actual offensive cyber attacks. Offensive cyber operations are mainstream and not subjected to the non-use taboo surrounding nuclear weapons [For further discussion on the nuclear taboo see Nina Tannenwald, *The Nuclear Taboo: The United States and The Non-Use of Nuclear Weapons Since 1945* (Cambridge: Cambridge University Press, 2007).]. Indeed, where Bernard Brodie noted that nuclear weapons changed the nature of war regardless if they are used, cyber weapons have changed the nature of war in the sense that it is an almost certainty they will be employed [Bernard Brodie, "Implications for Military Policy," in *The Absolute Weapon: Atomic Power and World Order*, ed. Bernard Brodie (New York: Harcourt, 1946), 83]. Be it on the battlefield or against the distant homeland of an invader, the employment of cyber weapons will only become more common in future conflict. Fourth, fears of the widespread proliferation of nuclear weapons, which has been a paramount concern since the dawn of the nuclear era, never materialized as opposed to the rapid development and proliferation of cyber arms [see, for instance, Kurt M. Campbell, Robert J. Einhorn, and Mitchell

www.manaraa.com

First, in the cyberspace, secrecy is an inherent characteristic of the domain such that it creates two related complications for deterrence and escalation: the difficulties states encounter when attributing the physical origin and/or political responsibility for a cyber attack, and the decision to reveal information about attribution; and the fact that secretly operating in the domain is essential to mission success. Second, the defining feature of cyber weapons—the fact that they are non-physical weapons that are comprised of zeroes and ones—has implications for vulnerability and survivability, as well as the nature of their lethality. Third, at least at the current technological state, there are fundamental limitations, both material and psychological, to the costs that cyber weapons can generate against targeted states. Below, I outline these interrelated characteristics of the cyber domain, which serve as the bedrock for understanding how the logics of deterrence and escalation apply to cyberspace.

Secrecy is Inherent to the Cyber Domain

Successfully operating in cyberspace demands that states conceal their activities in the domain, both because espionage is a critical element of offensive cyber operations and this, by definition, is a secretive activity; and because revealing information about access and capabilities undermines efficacy. Espionage is the foundation of offensive operations in cyberspace because

---

Reiss, eds. *The Nuclear Tipping Point: Why States Reconsider Their Nuclear Choices* (Washington, DC: Brookings Institution Press, 2004)]. Indeed, unlike nuclear weapons (thus far), offensive cyber capabilities have proliferated across most states and have not been restricted to governmental actors. Indeed, corporations, criminal elements, and even individuals can develop cyber arms. Additionally, unlike the traditional domains of warfare—land, air, sea, and space—the cyber domain is the one domain where there is near offensive parity between many states; there may even be parity between some states and non-state actors. Much of this proliferation is due to low barriers of entry into the cyber field as offensive capabilities can be acquired via multiple online market places or through various cyber security firms [For example, see, "The Digital Arms Trade," *The Economist*, March 30, 2013, http://www.economist.com/news/business/21574478-market-software-helps-hackers-penetrate-computer-systems-digital-arms-trade]. However, this is not meant to imply that offensive cyber weapons are low cost capabilities. As is discussed in this paper, they may come with significant price tags for the initial developer.

www.manaraa.com

successfully delivering an effect against a target requires that a state first (covertly) gather intelligence about it, to include ascertaining how to gain access, particularities about the target's capabilities and vulnerabilities, and mapping the target's virtual network.[4] While a nuclear weapon is a "fire and forget" capability that delivers the same effect regardless of the target with which it comes into contact, employing cyber weapons requires intelligence not only of the target system the state intends to affect, but also of the network vulnerabilities upon which access can be established in order to launch the tool. Specifically, a cyber attack requires three distinct operations. The first is collecting information on the targeted system or process (e.g., the manufacturer of the system, its operating system, its most recent firmware update, etc.). This could be conducted through various intelligence means or scanning the system in question via an assortment of cyber techniques. The second operation necessitates gaining access to the targeted system. This access functions as a foothold that enables further intelligence collection or could serve as a keyhole through which an offensive cyber weapon could be launched against the targeted system. The final step involves executing the malicious code against the targeted system via the access provided.

Therefore, beyond espionage, secrecy is imperative to carrying out the requisite steps of a cyber operation described above. In particular, secrecy is necessary to preserve access, because if the target uncovers that a state has gained access to its networks, it can marshal defenses and take measures to patch its vulnerability and render the attacker's access capability moot. Secrecy is also critical to preserving the capability of the weapon itself, because revealing that information

[4] Generally, cyber operations come in four varieties: cyber defense, cyber penetration (in some military circles "operational preparation of the environment (OPE)" depending on the intended end-state), cyber espionage, and cyber attack. Where the first is conducted entirely on one's own networks, the later three are conducted on foreign nets. Though the intent and the effect delivered for the last three types of operations vary greatly, they can also support cyber defense if employed to collect intelligence on another entities' cyber capabilities.

www.manaraa.com

allows the target to develop defenses against it and it may also reveal the attacking state's targeting strategy and broader set of capabilities.

Because states have an enormous incentive to conceal their capabilities, accesses, and activities in the cyber domain, and due to the technical nature of cyber weapons, states that are the targets of offensive cyber attacks confront significant hurdles to attributing the origins of a cyber attack and ascertaining political responsibility.[5] The technical design of the Internet itself complicates attribution because its virtual nature enables nearly-invisible operations and facilitates attackers obfuscating the point of departure of an attack by using multiple proxies and other anonymizing capabilities.[6] Furthermore, attribution of a cyber attack may take weeks or months and only result in a degree of confidence in the true source of an attack. More importantly, attribution involves a strategic, political calculus to assign responsibility to an entity for a cyber attack, which is confounded by deliberated efforts by governments to obscure command and control for cyber attacks, such as employing cyber proxies.[7] This is fundamentally distinct from nuclear weapons where, due to the limited numbers of states possessing bombs, the

[5] Thomas Rid and Ben Buchanan, "Attributing Cyber Attacks," *Journal of Strategic Studies* 38, no. 1-2 (2015): 4-37.

[6] For further reference see Clement Guitton and Elaine Korzak, "The Sophistication Criterion for Attribution: Identifying the Perpetrators of Cyber-Attacks," *The RUSI Journal* 158, no. 4 (2013): 62-68.

[7] Rid and Buchanan argue that, at the strategic level, "attribution is a function of what is at stake politically," 7. Also see Erica D. Borghard and Shawn W. Longeran, "Can States Calculate the Risks of Using Cyber Proxies?" *Orbis* 60, no. 3 (May 2016), 395-416, for a discussion of why states delegate authority to proxy actors to conduct cyber attacks on their behalf.

www.manaraa.com

ability to detect launches, and the unique signature of nuclear enrichment processes, attribution was not a significant concern for early nuclear scholars.[8]

Cyber Weapons are Distinct

The virtual nature of cyber weapons makes them fundamentally distinct from nearly any other category of weapon in a way that has significant implications for the functions of deterrence and escalation. In particular, cyber weapons are simultaneously uniquely invulnerable and specific in their employment. Because cyber weapons are virtual—comprised of lines of code that could always be regenerated—the tools themselves are intrinsically invulnerable and survivable. This is a crucial distinction from nuclear weapons. Concerns about ensuring the invulnerability of nuclear arsenals were paramount because survivability undergirded the stability of nuclear deterrence by ensuring a state could credibly threaten to retaliate against a nuclear attack.[9] Indeed, Wohlstetter argued that deterrence fails due to temporal constraints if one side can launch a surprise attack that destroys the opponent's retaliatory capability.[10] However, Brodie noted that due to the destructive nature of nuclear weapons, retaliation did not have to be certain, just probable enough to serve a deterrent value.[11] Initially Brodie believed that the ability to respond to a nuclear attack depended on the survivability of a nation's armed forces and their ability to operate autonomously from major urban centers (assuming that cities would

[8] For instance, see Bernard Brodie, "War in the Atomic Age," in *The Absolute Weapon*, 74. That said, depending on how they are conducted, underground nuclear explosions could hinder attribution efforts.

[9] Indeed, secrecy was important to nuclear weapons because it guarded against proliferation and also contributed to survivability.

[10] Albert Wohlstetter, "The Delicate Balance of Terror," *Foreign Affairs* 37 (1958): 211-246, 211.

[11] Bernard Brodie, "Implications for Military Policy," in *The Absolute Weapon*, 74.

www.manaraa.com

fall victim to counter value targeting).[12] However, following the development of thermonuclear weapons in the early 1950s Brodie amended his statement and noted that survivability of nuclear deployment systems was essential to stabilizing deterrence. Concerns over survivability of weapons systems that marked much of the nuclear era has little bearing on cyber weaponry. Cyber weapons do not require airfields, missile silos, or submarine pens to protect them from attack. Cyber weapons could easily be concealed on a hard drive, a disk, or other type of removable media. Furthermore, the majority of cyber tools could be employed from virtually any area that has Internet connectivity regardless of geographic location.[13] Unless an actor is foolhardy enough to store all of her cyber armaments on a centralized server that can be targeted, it is unlikely that an outside actor could destroy another actor's ability to retaliate with offensive cyber capabilities. Even if the latter occurred, the virtual nature of cyber weapons means that an actor could always simply write new code. However, unlike nuclear weapons, which are nearly guaranteed to deliver an immensely destructive effect against any target, successfully deploying cyber weapons is contingent on maintaining access to a target's networks and systems. While cyber weapons cannot be physically destroyed, access is highly precarious and can be vulnerable in unpredictable ways.

Related to the access requirement, cyber weapons are unique in that they lack universal lethality. While nuclear weapons, or nearly any other munition, are target-agnostic, cyber weapons must be tailored to a specific target set, or type of target (which underlies the importance of espionage). Cyber weapons are nothing more than code that interacts with other code that directs a system to do something that the owner did not intend. Yet developing these

[12] Ibid., 87.

[13] The caveat to this point is that there are some isolated targets that operate on closed networks (i.e. not connected to the Internet) that may require proximity to gain access.

www.manaraa.com

capabilities is hard for two reasons. First, states typically secure the critical systems that might be targeted because of their importance to the functioning of a society and national security; and second, these critical targets often employ custom developed supervisory control and data acquisition (SCADA) software that is unique to that system and only intimately understood by the original developers. This means that developers of cyber weapons must not only gain intelligence on a target that most likely is well defended and probably not connected to the Internet, they must also possess intimate knowledge of the specific technology that the system was built on. Developing a capability that can interface with a custom-built system is difficult, but it is by orders of magnitude more difficult to develop the mastery necessary to manipulate the system to do something that it may have been designed to resist. Additionally, obfuscating one's virtual presence from defenders throughout the entire process requires a level of tradecraft that few possess. Indeed, Herbert Lin notes that, "…the cost of a cyber weapon, which is almost entirely in R&D, cannot be amortized over as many targets as would be the case for a kinetic weapon. This fact necessarily increases the cost-per-target destroyed."[14] This is significant for both deterrence and escalation because there is no metric for measuring relative power between cyber actors, while there are both qualitative and quantitative metrics for assessing the relative power of nuclear states—one cannot simply count the number of cyber weapons an actor has in its arsenal in the same way that one can count the number of nuclear-equipped missiles a state has stockpiled. There are important implications stemming from the inability to measure relative power, particularly for escalation because states may be reticent to ratchet up against an adversary if there is endemic uncertainty about their relative power. Moreover, the lack of a

[14] Herbert Lin, "Oft-Neglect Cost Drivers of Cyber Weapons," *Council on Foreign Relations-Net Politics* (online blog), December 14, 2016, http://blogs.cfr.org/cyber/2016/12/14/oft-neglected-cost-drivers-of-cyber-weapons/.

www.manaraa.com

measure of relative cyber strength between actors may inhibit economies of scale. This is further complicated by the fact that cyber weapons have a shelf life. The access upon which they rely, and the effect they deliver, is only an update away from being patched.

Cost Generation in Cyberspace

Finally, cyber weapons cannot generate costs for a target at a magnitude comparable to that of conventional munitions, let alone nuclear weapons.[15] The utility of military instruments of power inheres in their abilities to inflict significant damage and harm on a target state to achieve a political objective. Cyber weapons could be used to cause disruption of an adversary's networks and systems—overwhelming them such that they lose the ability to function or the target loses confidence in their reliability—or causing destruction by destroying data resident on these systems or, in rarer circumstances, producing a physical effect.[16] While conducting multiple cyber attacks against a targeted state's critical national infrastructure, for example, could in theory cause significant destruction, replicating a cyber attack across numerous targets to produce a strategic effect may be beyond the realm of all but the most capable states. To draw an analogy, a war is comprised of tactical engagements that all contribute to the overall cost of the war. Tactical victories in cyberspace, taken individually, may appear to be relatively costless and net a high probability of success, but combining many of these victories together into a campaign that achieves a desired strategic end state is unlikely because the costs of producing such capabilities is extreme. Even if technology improves in a way that enables states to achieve strategic effects on the battlefield, or against civilian assets using cyber weapons, sustaining

[15] See Erik Gartzke, "The Myth of Cyberwar," *International Security* 38, no. 2 (Fall 2013): 41-73; and Thomas Rid, *Cyber War Will Not Take Place* (Oxford: Oxford University Press, 2013).

[16] For a more detailed discussion of this point see Erica D. Borghard and Shawn W. Lonergan, "The Logic of Coercion in Cyberspace," *Security Studies* 26, no. 4 (2017): 452-481, 461-463.

www.manaraa.com

costly campaigns over time is confounded by the fact that it takes time to develop capabilities and accesses (especially due to the absence of universal lethality of cyber weapons); and that the adversary is a strategic actor who can take action to mitigate existing vulnerabilities once the attacking state's tools and accesses are revealed as they are employed. Therefore, as states employ cyber weapons over the course of a campaign, they are likely to use their most decisive capabilities at the outset and will be less effective and/or precise over time, generating effects against most vulnerable rather than more deliberate targets.

Relatedly, while it is possible to employ cyber weapons to produce a destructive effect in the physical realm, no one to date has died as the direct result of a cyber attack. Thus, while it may be possible to generate comparable costs to a target using cyber and conventional military means—in terms of the financial cost of a particular attack—decision-makers (and domestic publics) are likely to perceive the "cost" of physical and virtual destruction differently. This may be similar to the psychological perceptions surrounding certain classes of weapons that societies have deemed heinous, even though their destructive capabilities may be similar to, or even less potent than, weapons deemed acceptable.[17] The perception of the likely cost of an attack is important because it factors into a decision-maker's calculus regarding the relative payoffs of various courses of action in the context of deterrent threats, and it also helps define the thresholds according to which a government assesses decisions to escalate.

Together, these three attributes of the cyber domain—the integral role played by secrecy, the distinct nature of cyber weapons, and the differences in cost generation—affect the logics of deterrence and escalation as applied to the cyber domain. In the context of escalation, the nature of strategic interaction in the cyber domain creates self-dampening effects and injects breathing

[17] See, for instance, the literature on the nuclear weapons taboo or the chemical weapons ban.

www.manaraa.com

room into crisis situations that, together, mitigate the risks of escalatory spirals. However, these same attributes also problematize efforts at achieving successful deterrence because they undermine effective signaling and create significant difficulties in discerning and generating the requisite capabilities for a deterrent threat. In the following section, I turn to the problem of deterrence in cyberspace and explore how these factors complicate state efforts at deterrence.

Deterrence in Cyberspace

Deterrence, at its core, is prevention by fear. Robert Art defines deterrence as, "the deployment of military power so as to be able to prevent an adversary from doing something that one does not want him to do and that he otherwise might be tempted to do by threatening him with unacceptable punishment if he does it."[18] For much of the Cold War the United States and the Soviet Union pursed a strategy of *mutual* deterrence to reciprocally prevent the onset of nuclear war. To many scholars, it was the successful application of deterrence by both sides that kept the Cold War from becoming hot.[19] Theoretical canons were developed during the twentieth century that sought to explain the origins and continuation of stable relations between the Cold War superpowers based on a strategy of deterrence, and sought to identify how perceived momentary military advantages could undermine the preservation of the status quo.

---

[18] Robert J. Art, "To What Ends Military Power?" *International Security* 4, no. 4 (Spring 1980): 3-35, 6. However, it is difficult to ascertain whether a given deterrent threat was successful because deterrence has a negative object.

[19] But see the following for debates about the efficacy of deterrence strategy during the Cold War: Robert Jervis, *The Illogic of American Nuclear Strategy* (Ithaca: Cornell University Press, 1984); Lawrence Freedman, *The Evolution of Nuclear Strategy* (London: Macmillan, 1989); Austin G. Long, *Deterrence: From Cold War to Long War: Lessons from Six Decades of RAND Deterrence Research* (Santa Monica, CA: RAND Corporation, Vol. 636, 2008); and Francis J. Gavin, *Nuclear Statecraft: History and Strategy in America's Atomic Age* (Ithaca: Cornell University Press, 2012).

www.manaraa.com

More recently, scholars and policymakers alike have attempted to apply deterrence models to the cyber domain, given the widely-accepted notion that there is a need to establish norms of acceptable behavior in cyberspace and limit offensive cyber operations and actors. However, while there is some scholarly consensus that achieving deterrence in the cyber domain is far more difficult, if not impossible, than it was for practitioners during the Cold War, the existing cyber literature has not systematically assessed the various logics of deterrence and the mechanisms through which deterrence may or may not work in cyberspace. Therefore, in this section, given the three distinct attributes of cyberspace introduced above, I explore how the traditional logics of deterrence by punishment, denial, and cross-domain deterrence apply to the cyber domain. I will demonstrate that, across all three types of deterrence, the opaqueness of the cyber domain, the unique nature of cyber weapons, and the problem of cost generation in cyberspace make deterrence nearly impossible by undermining the ability of states to credibly signal capability and resolve; and to generate and convey proportionate deterrent threats. However, I also demonstrate that the application of these three types of deterrence—punishment, denial, and cross-domain—to cyberspace is complicated by logics and variables that are unique to each form of deterrence.

The General Argument

Achieving successful deterrence during the Cold War came with incredibly high stakes due to the devastating costs of nuclear war. The primary factor that made nuclear deterrence difficult, therefore, was the credibility of deterrent threats. Under conditions of mutual assured destruction, a credible nuclear deterrent rested on threatening to take actions that would result in the destruction of the threatener's society. Moreover, defining a political objective that would be worth a nuclear war defies basic cost-benefit calculations and is inherently anti-Clausewitzian.

www.manaraa.com

Indeed, this is why Thomas Schelling's seminal work on the topic, *Arms and Influence*, devotes so much time to the problem of credibility and all of the ways in which states could enhance the credibility of their threats (e.g., hands-tying, brinksmanship, the art of commitment, to name only a few).[20] The credibility of a threat is a function of capability—possessing the means to carry out the terms of a threat—and resolve—the willingness to do so. For nuclear deterrence, capability is an unproblematic aspect of credibility because all states are aware of the immensely destructive nature of nuclear weapons, and states can easily demonstrate their capabilities through shows of force. Resolve, on the other hand, is incredibly difficult to convey because there is an inherent *in*credibility in threatening nuclear war for the reasons stated above. Therefore, scholars and policymakers during the Cold War focused on means of conveying a willingness to use an unthinkable weapon.

Deterrence in cyberspace is problematic for precisely the inverse reason. The offensive employment of cyber weaponry is hardly unthinkable. Indeed, states have developed and employed these capabilities since the 1980s, and their use is only becoming more prevalent. Thus, cyber warfare is more immune from the so-called "credibility problem" faced by nuclear strategists. There is little reason to doubt the *general* credibility associated with the use of offensive cyber weapons because states have demonstrated a willingness to do so, and even use cyber capabilities to target an adversary's critical infrastructure. However, cyber weapons possess a distinct credibility issue in the context of cross-domain deterrence, which stems from the difference in the perception of virtual versus physical destruction. More specifically, a state may doubt the credibility of a deterrent threat that promises to respond to a cyber attack

[20] Thomas C. Schelling, *Arms and Influence: With a New Preface and Afterword* (New Haven, CT: Yale University Press, 2008). As Schelling notes, the problem of credibility is even more severe for extended deterrence, where the United States during the Cold War had to convince the Soviets that Berlin or Paris, for instance, was as integral as New York or Washington.

www.manaraa.com

(resulting in *virtual* damage that may produce an effect in the physical world, but does not directly cause death) with a kinetic one (resulting in *physical* damage). Additionally, conveying credibility in a specific situation is problematic in the cyber domain due to the requirements for operating secretly. While recent scholarship has demonstrated that it is possible to credibly signal in secret, specifically through the use of covert actions, this logic rests on the fact that local actors can observe covertly-operating states taking costly and risky actions on the ground in an area of conflict.[21] However, the manner in which secrecy affects cyber operations falls outside of the scope of this kind of reasoning. Setting aside the fact that there are serious problems associated with assessing the relative cost of cyber attacks (and, therefore, their value as a signaling tool), actions that reveal a state's cyber weapons and accesses can provide decisive information to the target of a threat that enables her to take protective measures that render the threat moot. Thus, cyber attacks are poor tools of costly signaling, even when conducted in secret because a signal, once sent, may likely not be repeated.[22] Additionally, even covert signaling through conventional means relies on a reasonable assessment about the identity of the perpetrator of an attack. In cyberspace, targets are not only confronted with an attribution problem, discussed above, but there may also be a significant time lag between when an operation is carried out and when it is perceived by the target, unless the attacker self-identifies (invoking all the risks associated with operating overtly in the cyber domain). Therefore, this suggests that while there is no general credibility problem associated with the use of offensive cyber weapons, conveying credibility in specific instances may be deeply problematic for reasons that are inherent to the cyber domain.

---

[21] See Austin Carson and Keren Yarhi-Milo, "Covert Communication: The Intelligibility and Credibility of Signaling in Secret," *Security Studies* 26, no. 1 (2017): 124-156.

[22] For more, see Borghard and Lonergan, "The Logic of Coercion in Cyberspace."

www.manaraa.com

Beyond these distinct credibility problems, the fundamental conundrum of deterrence in cyberspace is capability. Nuclear capability is easy to demonstrate and the staggering costs of employing nuclear weapons are clear to policymakers.[23] Capability in cyberspace, however, is incredibly difficult to measure; to understand in relative or proportionate terms; and to signal. As will be described below, measuring cyber capabilities is challenging due to the importance states assign to secrecy and the absence of universal lethality of cyber weapons—the fact that specific accesses and tools must be developed for specific targets; that gaining and maintaining accesses is contingent; and that attacks may be unpredictable in terms of cost, scope, and effect. This confounds deterrence for numerous reasons. As described above, a deterring state may be loath to reveal capabilities (which enhances the credibility of deterrence) because the act of revealing them renders them impotent. The deterring state and/or the target of deterrence may not be able to assign a value to the cost associated with the threatened punishment if the latter defects relative to the value the latter ascribes to defection. Relatedly, the threatened punishment may simply not be sufficiently costly to affect the target's calculus, or the target may be willing to gamble that a threatened action may not produce the effect intended by the deterring state due the often-unpredictable nature of cyber operations. Finally, signals regarding capability may go unrealized, or be misinterpreted or misperceived.

Policymakers and scholars, recognizing the inherently problems of credibility of nuclear deterrence, attempted to devise means of overcoming the credibility gap and grappled with whether it could be possible to plan for limited nuclear options. The question for the cyber

[23] Indeed, this is what underlies Kenneth Waltz's arguments regarding the stability of nuclear deterrence. See Kenneth N. Waltz, "The Spread of Nuclear Weapons: More May be Better," *The Adelphi Papers* 21, no. 171 (1981).

www.manaraa.com

domain is whether the capability conundrum and other problems associated with cyber deterrence could ever be overcome.[24]

Deterrence by Punishment

Deterrence is a "coercive strategy" that seeks to prevent an actor from taking an unacceptable action.[25] As Robert Art describes, deterrence involves "…the threat of retaliation… (whose) purpose is to prevent something from happening."[26] Deterrence succeeds, therefore, when "the risks and cost of military action are very high."[27] The literature distinguishes between two different types of deterrent threats: deterrence by denial, and deterrence by punishment. Deterrence by denial, according to Glenn Snyder, involves threatening to physically impede the adversary's ability to successfully carry out a military operation or making it too costly to do so. This form of deterrence, therefore, works by targeting the adversary's military capabilities (counterforce targeting in the context of nuclear weapons) and/or shoring up one's own defenses such that offensive operations are inordinately costly for an attacking state. Deterrence by punishment, on the other hand, rests on the credible threat to wreak devastating pain and suffering on a target's civilian population such that the perceived costs of an action are deemed unacceptably high.[28] The notion that threatening punishment to affect a target state's behavior

---

[24] One remedy may be through the creation of confidence building measures. See Shawn W. Lonergan, "Arms Control and Confidence Building Measures for the Cyber Domain" (working paper).

[25] Lawrence Freedman, *Deterrence* (Cambridge: Polity, 2004), 26.

[26] Art, "To What Ends Military Power?" 6.

[27] John J. Mearsheimer, "Nuclear Weapons and Deterrence in Europe," *International Security* 9, no. 3 (Winter 1984/85): 19-46, 21.

[28] Glenn H. Snyder, *Deterrence and Defense* (Princeton, NJ: Princeton University Press, 1961), 14-15.

www.manaraa.com

gained prominence during the strategic bombing campaigns of the Second World War, where Allied strategists calculated that unleashing terror from the skies against German and Japanese populations would prompt the citizenry to rise up against their governments and force them to concede.[29] In the context of nuclear weapons, deterrence by punishment involved counter value targeting against the adversary's population centers. Lawrence Freedman assert that, "[f]or a sanction of tough punishment to be an effective deterrent it is necessary for a would-be offender to know that there is a high chance of (a) being apprehended and (b) being punished severely."[30] During the nuclear age, the ability of the Soviet Union and the United States to mutually deter the outbreak of strategic nuclear war rested on the reciprocal threat of unimaginable punishment which, paradoxically, created stability in the international system. The stability of mutual deterrence through punishment hinged on the survivability of second strike nuclear forces such that "neither [side], in striking first, can destroy the other's ability to strike back."[31]

In the cyber domain, broadly speaking, offensive operations are typically thought of as being synonymous with "cyber attacks," which are commonly defined as operations that disrupt, deny, destroy, or degrade access to some networked asset. However, offensive cyber operations may also include information operations that are designed to influence an individual or group's decision-making process. Indeed, in recent times, hacked accounts and leaked emails—even those that are altered—have influenced election outcomes and created reputational costs for the

[29] Giulio Douhet, *The Command of the Air*, trans. Dino Ferrari (New York: Coward-McCann, 1942). However, also see the following for a strong critique: Robert A. Pape, *Bombing to Win: Air Power and Coercion in War* (Ithaca: Cornell University Press, 2014). Note that both authors provide a discussion of punishment as it relates compellence, not deterrence, but both illustrate the punishment logic.

[30] Freedman, *The Evolution of Nuclear Strategy*, 83.

[31] Thomas C. Schelling, "Surprise Attack and Disarmament," *Bulletin of the Atomic Scientists* 15, no. 10 (1959): 413-418, 414.

www.manaraa.com

targets. Table 1 identifies the different types of offensive operations that may be employed to create costs against an adversary. However, there is only one type of offensive method that, in theory, could be employed for the purposes of punishment. As will be discussed in greater detail below, the limited utility of most offensive cyber operations to support a punishment strategy contributes to the conclusion that deterrence by punishment is unlikely to succeed in cyberspace.

Deterrence by punishment relies on the logic that a state could credibly respond to a deviation from the status quo by administering a punishment capable of bringing a society to its knees. As Table 1 demonstrates, the type of offensive cyber operation that is even conceptually able to generate these costs would be a cyber attack on critical infrastructure. In theory, undermining critical infrastructure that is essential to the functioning of a society could force panic-stricken citizens to press their leaders to capitulate. However, Borghard and Lonergan argue that the offensive employment of cyber power is not capable of generating the costs necessary for an effective coercive punishment strategy because offensive cyber operations targeting critical infrastructure (e.g. power grids, pipelines, and transportation nodes, etc.) require highly tailored accesses and exploits which are difficult to scale given the material resources necessary for their development and employment.[32]

[32] Borghard and Lonergan, "The Logic of Coercion in Cyberspace."

www.manaraa.com

Table 1: Types of Offensive Cyber Operations

Offensive Method Employed	Description	Cyber Attack	Information Operation
Hacked Account and Data Leak	Stealing private or classified information and releasing to a third party or the public at large.		X
Website Defacement	Changing the appearance of a website often to spread a political message. Typically involves that loss of the original functionality of the website.	X	X
Cyber Extortion	Requiring an actor to do something that they would otherwise not do (e.g. pay a ransom) or else their website will be overwhelmed with a packet flood, their private information released or destroyed, or their network crashed. Often this involves a loss of access to the extorted data.	X	X
Distributed Denial of Service (aka packet floods)	Overwhelming a system through excessive requests for information, thus forcing a system shut down or severe degradation to operations.	X	
Disruption or Destruction of Critical Infrastructure	Installing an exploit on a critical system necessary for the functioning of society that when activated causes some change from the system's normal operation.	X	
Disruption or Destruction of Non-Critical Infrastructure	Installing an exploit on a non-critical system that when activated causes some change from the system's normal operation.	X	
Undermine Data Integrity	Altering data in transit or at rest so that the originator and/or receiver cannot rely upon its authenticity.	X	

For deterrence by punishment to succeed, three requirements must be met. First, the deterring state must possess a capability powerful enough to affect the target's calculus regarding the costs and benefits of a course of action; specifically, she must be able to hold the target accountable and impose a significant level of damage on the target's civilian population and society. Relatedly, the state must possess a mechanism to deliver the threatened punishment.

www.manaraa.com

Second, the state should be able to credibly signal to the target state the former's intent to carry out the threatened action. Third, the target of deterrence must accurately perceive the deterrent threat and make a risk calculation about its own ability to get away with the undesired action. The effective implementation of these requisites of deterrence by punishment are complicated by the three attributes of the cyber domain identified above, as well be demonstrated in the below discussion. Most importantly, deterrence by punishment that relies solely on offensive cyber operations is likely to be ineffective because cyber warfare is an inadequate means to instill an unbearable level of fear in an adversary.

Possessing the Capability

Measuring capabilities in cyberspace and, therefore, issuing effective deterrent threats are complex and difficult. This difficulty is compounded by the limits in cost generation associated with cyber warfare and the difference in how decision-makers are likely to perceive physical versus virtual destruction. In the nuclear arena, there are methods to estimate a state's arms stockpiles and there are treaties, accords, and international oversight institutions that monitor and limit these weapons.[33] However, as already noted, in the cyber domain there is no measure of relative strength because cyber weapons lack universal lethality; one cannot simply count the number of cyber tools the same way one can count the numbers of missiles or pounds of poison gas a state possesses. Certain types of cyber attacks require minimal, target-specific tailoring and are not access dependent; therefore, these capabilities are more "universal." For example, distributed denial of service (DDoS) attacks can overwhelm a target's systems using botnets-for-hire that are easily customizable and that can adapt to surmount a target's defenses (such as

[33] The same is also true with biological weapons. For instance, see, Gregory Koblentz, "Pathogens as Weapons: The International Security Implications of Biological Warfare," *International Security* 28, no. 3 (Winter 2003/04): 84-122.

www.manaraa.com

blacklisting source nodes). However, deterrence by punishment requires possessing the capability to deliver devastating strategic effects against an adversary's population centers. In cyberspace, these kinds of capabilities are almost always dependent on gaining and maintaining unique accesses and are highly tailored because the targets—critical national infrastructure such as SCADA systems—are typically extremely customized.

Assessing the capability to carry out the threat of punishment is further complicated by the secrecy that pervades the cyber domain. In the context of nuclear deterrence, states were incentivized to conduct shows of force to demonstrate capability. However, the nature of cyber warfare produces the opposite incentive: to shroud a tool set in secrecy so that a would-be defender does not erect counter measures that may thwart a desired course of action. Indeed, Austin Long notes that, for this reason, brandishing cyber weapons may attenuate a deterrence strategy.[34] Moreover, a show of cyber force is a poor indicator of latent cyber capability due to the tailored nature of offensive cyber operations against strategically decisive targets. In other words, within the scope of cyber operations carried out for the purposes of punishment, any one cyber operation is an unreliable predictor of future ones because they often employ niche capabilities.[35]

There is also a human capital element of assessing capabilities that is uniquely important for cyber warfare due to the highly-skilled nature of operations developed for the purposes of

[34] Austin Long, "Deterrence: The State of the Field," *New York University Journal of International Law and Politics* 47 (2014): 357-377, 374.

[35] A potentially more useful assessment of capability does not come from the target of deterrence, but from whatever indices exist about the aggressor's cyber capabilities and doctrine. For further reference, see Borghard and Lonergan, "The Logic of Coercion in Cyberspace."

www.manaraa.com

punishment strategies.[36] Some tacticians may be driven to count the number of cyber forces that a government openly reports or note estimates from military journals to assess capability, but these sources do not consider critically important differences in skillsets. For this reason, comparing operators across states is problematic. For example, China is reported to possess legions of cyber operators but these units, such as its Information Warfare Militia, lack widespread standardized training pipelines and, therefore, there is substantial variation in skill within and across these units. The Israel Defense Force's Unit 8200, in contrast, is a highly skilled, creative, and motivated cyber force but, as it is a military unit, it suffers the disadvantages of short enlistment times, high turnover, and requirements to train new personnel with limited technological experience.[37] When carrying out a punishment strategy in the cyber domain, possessing the most "trigger pullers" is not as important as possessing the right trigger puller, armed with the right capability, with access to a vulnerable target.

Beyond simply possessing a capability, a state seeking to deter through the threat of punishment must also possess a means the deliver the punishment; otherwise, the capability itself is irrelevant. The nuclear triad—the B-52 bomber, the nuclear-armed submarine, and the intercontinental ballistic missile—were the paradigmatic images of the Cold War because they were the delivery systems that made the threat of nuclear war real. The comparable analogy in cyberspace is access to an adversary's networks and critical infrastructure, which must be

[36] Again, "cheap, fast, and easy" attacks are not difficult and marginally skilled individuals can conduct these kinds of operations from the safety of their homes. However, the kinds of cyber operations that would comprise a punishment strategy are incredibly complex to develop and carry out and, therefore, require highly skilled operators.

[37] Given the secrecy of cyber operations, the units that conduct them are often shrouded in secrecy. The most comprehensive and publically available state comparison is- United Nations-Institute for Disarmament Research, "The Cyber Index- International Security Trends and Realities," UNIDIR/2013/3, (2013), http://www.unidir.org/files/publications/pdfs/cyber-index-2013-en-463.pdf.

www.manaraa.com

ensured prior to delivering a cyber payload.[38] Access is particularly important for punishment strategies because governments take measures to protect critical national infrastructure from attack by taking measures such as building the system on a closed network that is isolated from the Internet. However, as already noted, access is necessary to collect intelligence on a target so that an attacker can develop a capability to alter the functionality of the intended target as part of cyber attack.

There are three methods for gaining access to an adversary's network: remotely (through the Internet); physical access; and supply chain interdiction. Hacking, or remote access, involves connecting to the target system through the Internet or a network that an operator can already access. However, finding a target can be a daunting task, requiring extensive network mapping to ensure that the operator has found the right network in which to burrow. The operator must also have adequate exploits to surpass any firewalls and advanced defensive efforts that the adversary has in place. Once inside the network, the operator must then locate the intended target within the new cloud of information it is navigating. While the process of gaining access may be complex, once an operator has established a foothold for access it is easy to return to the target through this backdoor for further operations as long as the operator remains undetected. However, maintaining persistent access is often unpredictable because the target may take measures, such as installing patches and updates to a system, that inadvertently result in the operator being expelled.

Since the majority of cyber incursions come via remote access, firewalls are typically placed at the gateway between the network and the Internet. To protect the most sensitive targets

[38] Note that access is not required for all cyber operations, as already discussed above. However, offensive cyber operations meant to support the threat of punishment, in line with the theoretical framework of the conditions where it may be successful, are almost always access-dependent due to the nature of the target.

www.manaraa.com

(i.e., those that would be targeted as part of a punishment strategy), many actors have constructed closed networks that are not connected to the Internet, with the aim of preventing hackers from gaining access. Typically, this is referred to as an "air-gapped" network since there literally is air between the networks' computers, devices, and other machines that may connect to the Internet. However, there are some types of malware that are designed to "jump" the air gap by transmitting data to and from machines that may connect to the Internet. This malware functions by translating code into high frequency sounds that can then be detected by the microphones that are now common on most computers.[39] Therefore, most malware programs that are designed to target closed networks possess a means to overcome the air gap.[40]

Closed networks are not immune from adversary penetration, as supply chain interdiction as well as human operators can be used to physically gain access. In the latter method, a human being is used to physically engage with the target. This approach allegedly enabled the actors behind Stuxnet to gain access to Iran's critical infrastructure through physically introducing the virus into the Natanz nuclear facility via a thumb drive or other personal computing device.[41] However, gaining physical access is difficult and risky because it requires a human operator on

---

[39] Dylan Love, "Hackers Can Infect Your Computer Even if It's Not Connected to the Internet," *Business Insider*, March 5, 2014, http://www.businessinsider.com/what-is-air-gap-malware-2014-3.

[40] The above discussion on the methods to gain access, by design, is devoid of many of the technical complexities of these cyber operations. For a more technically detailed explanation see Chapter 2, "Technical and Operational Considerations in Cyberattack and Cyberexploitation," in eds. William A. Owens, Kenneth W. Dam, and Herbert S. Lin, *Technology, Policy, Law, and Ethics Regarding US Acquisition and Use of Cyberattack Capabilities* (Washington DC: National Academies Press, 2009).

[41] Ralph Langner, "Stuxnet's Secret Twin," *Foreign Policy*, November 19, 2013, http://www.foreignpolicy.com/articles/2013/11/19/stuxnets_secret_twin_iran_nukes_cyber_attack. The actor in this case could either have been a clandestine agent or an unwitting or disgruntled employee.

www.manaraa.com

the ground often behind an adversary's lines. A less risky option, therefore, is through supply chain interdiction. Under this method, a surreptitious actor would interdict some physical piece of the network (i.e. a printer, a modem, a hard drive, etc.) and install an implant prior to it reaching its intended destination. This implant is designed to "call back" through the Internet to its originator to await further directions or provide whatever information it was programmed to garner. This method is perhaps the most effective way to gain access, assuming that the perpetrator is aware of the shipment, has a capability and the time to interdict, and that it is ultimately delivered to where it is intended without trigging the recipient's suspicion.

In short, while there is no system that is theoretically beyond the reach of a state seeking to deter a target through punishment, the ability to gain access to sensitive systems does not guarantee that punishment is possible in the cyber domain. This is because there is a fundamental problem associated with cost generation in cyberspace; there is no guarantee that a state possesses a cyber weapon that could inflict an unbearable level of punishment upon an actor it is seeking to dissuade. In fact, given the extant technological state of the field, it may simply not be feasible to impose the kinds of costs on a target that would be required for deterrence by punishment to succeed.[42] While it is theoretically possible that cyber operations could lead directly to a loss of life, no one has reportedly died to date as a result a cyber attack, despite over 30 years of recorded cyber conflict. The typical scenario associated with a punishment strategy in cyberspace conjectures that loss of power stemming from a cyber attack on a power grid could lead to a loss of life. However, even in this instance, the conceivable deaths from the loss of power over an extended period of time are far from those that are required for an effective punishment strategy. To draw a comparison, when Hurricane Sandy hit the United States'

[42] The concept of punishment is covered in greater detail in Borghard and Lonergan, "The Logic of Coercion in Cyberspace," 477- 478.

www.manaraa.com

eastern seaboard in late October 2012 over 8.5 million people were left without power, with many going weeks and even months before it was brought back on line.[43] Yet, a U.S. National Hurricane Center postmortem of Hurricane Sandy reported that of the 159 people in the United States killed either directly or indirectly from Hurricane Sandy only, "[a]bout 50 of these deaths were the result of extended power outages during cold weather, which led to deaths from hypothermia, falls in the dark by senior citizens, or carbon monoxide poisoning from improperly placed generators or cooking devices."[44] If a cyber attack took out power of a similar magnitude and duration of Hurricane Sandy, it is conceivable that a comparable number of casualties would result. It would be highly suspect to suggest that this would generate the costs necessary for an effective punishment strategy.[45]

Credible Intent

Even under conditions in which a state possesses the capability to inflict cyber punishment on a target and a reliable means of delivering it, for deterrence to succeed the target must believe that the deterring state intends to carry out the terms of threat.[46] Indeed, making a credible threat requires the target to believe that the issuer possesses both the capability to impose the threatened cost and the will to employ it if the target does not comply with the

[43] Eric S. Blake et al., "Tropical Cyclone Report- Hurricane Sandy (AL182012) 22-29 October 2012," National Hurricane Center, February 12, 2013, http://www.nhc.noaa.gov/data/tcr/AL182012_Sandy.pdf, 14-15.

[44] Ibid., 14.

[45] Also see Borghard and Lonergan, "The Logic of Coercion in Cyberspace," for a discussion of how changing in technology, specifically the increased interconnectedness of societies through the Internet of Things, could change the nature of vulnerability.

[46] This article does not extensively plumb the depths of credibility in cyberspace as this topic is covered at some length in Borghard and Lonergan, "The Logic of Coercion in Cyberspace," 464-472.

www.manaraa.com

issuer's demands. Yet as previously noted, secrecy surrounds a nation-state's cyber capabilities and creates a situation of acutely imperfect information from which a policymaker must judge the threat's credibility. This is because the typical route to enhancing credibility—taking public measures to reveal capabilities or tie one's hands—undermines the likelihood of a cyber operation's success because secrecy is necessary to preserve access and capabilities. Therefore, there is an inherent tension between the incentive to conceal and the requirements of deterrence that are irreconcilable. Moreover, past cyber operations are unreliable measures of current credibility because cyber capabilities are not universal and replicable over time across different target sets. The result of this informational imbalance is that a state must look for signals and indices that are exogenous to cyberspace—such as diplomatic communications and increases in cyber resource allocations—to assess the other actor's intentions. The criteria that drive these assessments may be unique to the perceiving policymaker but will typically include an assessment of the strategic culture that has evolved in the state for an environment with a dearth of international laws, norms, or the possibility of iterative reciprocity as well as the use of whatever indices the receiver has developed.[47] While differences in strategic culture may complicate signaling across all of the domains of warfare, the problem is particularly acute in cyberspace due to the absence of norms and indices. Furthermore, the employment of cyber power offers no credible assurances that the confluence of offensive and defensive elements can be brought to bear at the appointed time and place to provide an unbearable punishment sufficient to dissuade the target from undertaking an undesirable action due to the unpredictability of access necessary to deliver an effect.

[47] For a thorough discussion on cyber norms see Martha Finnemore and Duncan B. Hollis, "Constructing Norms for Global Cybersecurity," *American Journal of International Law* 110, no. 3, (2016): 425-479. Also see Borghard and Lonergan, "The Logic of Coercion in Cyberspace," 456-459.

www.manaraa.com

The Target's Perception

Finally, for deterrence to succeed, it is imperative that the actor that is to be deterred understands what she is expected to do and believe that she cannot get away with taking the unwanted action. As already noted, the deterrence literature notes the dangers of misperceptions and the need to make intentions clear to avoid inadvertent conflict.[48] In cyberspace, understanding the intent of any cyber operation is exceptionally difficult given the mission's invisible nature. *Ex ante*, it is difficult to surmise the intent of any given cyber operation alone because discovering that an actor has penetrated a target's networks does not provide the target with any information regarding intent—penetration could support legitimate espionage operations, or it could reflect preparations for an offensive attack. Thus, intent can only be discerned after the execution of the mission, which by definition means that deterrence has failed—the threat of punishment alone should be sufficient to prevent the target from taking undesirable actions. Given the inability to understand intent from a cyber operation alone, a deterring state must couple cyber operations with established means of signaling in the physical domains, such as issuing a formal threat via a diplomatic channel so that the party to be deterred accurately understands the mandate.[49] However, as already described, there are incentives to keep the terms of the threat (the punishment to be meted out) secret because revealing capabilities and accesses undermines mission success.

A critical element to the successful functioning of deterrence in cyberspace is the target's calculations regarding whether she can get away with the action meant to be deterred by taking

[48] Robert Jervis, *Perception and Misperception in International Politics* (Princeton, NJ: Princeton University Press, 1976), 84, 113.

[49] For more on this point see Borghard and Lonergan, "The Logic of Coercion in Cyberspace," 456-489.

www.manaraa.com

advantage of the attribution problem. This presents a conundrum for deterring states because, if they cannot know with some measure of acceptable reliability the identity of the perpetrator, they may be hesitant to respond due to the desire to avoid an escalatory response against the wrong target.[50] Ascertaining the source and, more importantly, the political responsibility of a cyber incursion can be exceedingly difficult due to offensive techniques that use multiple proxy actors, servers, and distributed command and control networks to obfuscate an operation's true source. Though it is no guarantee, there are methods of cyber attribution. The most reliable attribution method requires the party carrying out the operation to acknowledge responsibility and the targeted state to believe that the self-identified aggressor possessed both the capability and motivation to perform the operation. Of course, a state being deterred through the threat of punishment would be unwise to publicly acknowledge its responsibility for an attack. Another means of attribution requires the targeted state to have prior access to the would-be coercer's network from which the attack originated and witness the operation's execution. This is believed to be how the United States was able to attribute the cyber attack against *Sony Pictures Entertainment* to the North Korean Government.[51] This method is not as desirable because, once attributed, the access to the targeted system will most likely be lost as the target wipes or replaces infected devices. Therefore, the long-term intelligence loss must be evaluated against

[50] This logic informs Adam Segal's proposition that, potentially, Edward Snowden's leaks containing information about the National Security Agency's sprawling capabilities could deter American adversaries through "omniscience." In other words, if U.S. adversaries believe that "the NSA sees all," they may be more susceptible to being deterred. See Adam Segal, *The Hacked World Order: How Nations Fight, Trade, Maneuver, and Manipulate in the Digital Age* (New York: PublicAffairs, 2016), 57.

[51] David E. Sanger and Martin Fackler, "N.S.A. Breached North Korean Networks Before Sony Attack, Officials Say," *The New York Times*, January 18, 2015, http://www.nytimes.com/2015/01/19/world/asia/nsa-tapped-into-north-korean-networks-before-sony-attack-officials-say.html.

www.manaraa.com

the need to publically attribute. The final method of assigning attribution is attained when the attack's signature is so unique that it can be traced to a specific actor. This is the least reliable method, yet recently there have been advances in signature recognition software that are designed to scour millions of lines of code in order to obtain unique profiles for developers.[52] Additionally, there have been situations where attribution has been attempted through code analysis, such as the use of a Hebrew reference in Stuxnet.[53] In Stuxnet's case, inserting this code suggested that the operation could be attributed to Israel—although strategic actors may intentionally plant false clues to obfuscate true responsibility. On the other hand, Stuxnet may have contained these references to signal to Iran that the United States and Israel have the means and determination to destroy its nuclear enrichment infrastructure and were therefore using an offensive cyber mission for the purposes of sending a deterrent signal, although it is important to note that this would have been an example of deterrence by the threat of denial, rather than punishment.

Therefore, despite improved techniques to alleviate the attribution problem,[54] the use of cyber power to meet the needs of deterrence by punishment is confounded for two reasons. First, the deterring party may find it a pointless exercise to issue deterrent threats in cyberspace if it cannot reliably identify the perpetrator of the action meant to be deterred. This suggests that, as

---

[52] Jack Goldsmith, "The Significance of Panetta's Cyber Speech and the Persistent Difficulty of Deterring Cyberattacks," *Lawfare* (online blog), October 15, 2012, http://www.lawfareblog.com/2012/10/the-significance-of-panettas-cyber-speech-and-the-persistent-difficulty-of-deterring-cyberattacks/.

[53] John Markoff and David E. Sanger, "In a Computer Worm, a Possible Biblical Clue," *The New York Times*, September 29, 2010, http://www.nytimes.com/2010/09/30/world/middleeast/30worm.html?pagewanted=all&_r=0.

[54] Leon Panetta, "Remarks by Secretary Panetta on Cybersecurity to the Business Executives for National Security," U.S. Department of Defense, New York City, October 11, 2012.

www.manaraa.com

will be discussed in greater detail below, other elements of national power may prove more viable for the purposes of deterrence. Second, though an operation such as Stuxnet may have contained insights into its origin as part of a deterrent threat, there is no guarantee that the target will trace a threat back to the deterring party unless it is coupled with a clear message communicated via an established platform. Thus, the attribution problem fundamentally undermines the efficacy of deterrence because it adds uncertainty both to the target's assessment of the coercing party cyber capabilities and to the target's calculations about its own ability to get away with unsanctioned behavior.

Deterrence by Denial

While the threat of punishment through civilization-ending nuclear war shaped the bulk of academic writing on deterrence during the nuclear era, efforts were made to explore whether tactical or theater nuclear weapons employed on the battlefield could serve the purposes of deterrence by denial, thus extending the literature on conventional deterrence to the nuclear age. Theater nuclear weapons were considered a "deterrent backstop to conventional defense" and could "deter by promising territorial denial."[55] Similarly, Glenn Snyder distinguishes between investing in deterrent forces that can threaten another state versus investing in defensive forces that mitigate the consequences of enemy aggression.[56] Schelling has a slightly different take on the distinction between deterrence and defense. To Schelling, the distinction was centered on force structure, but was related to the intent of force employment. If the intent were to ensure an

[55] Colin S. Gray, "Theater Nuclear Weapons: Doctrines and Postures," *World Politics* 28, no. 2 (January 1976), 305. Also see John J. Mearsheimer, "Nuclear Weapons and Deterrence in Europe," *International Security* 9, no. 3 (Winter 1984-1985), 19-46; and Bernard Brodie's discussion of tactical nuclear weapons in "The Development of Nuclear Strategy," *International Security* 2, no. 4 (Spring 1978), 76-78.

[56] Snyder, *Deterrence and Defense*, 3-5.

www.manaraa.com

attacking enemy could not succeed, it would constitute defense. However, "[i]f the object is to induce… [the enemy] to proceed, by making his encroachment painful or costly, we can…call it…'deterrent' defense."[57] Snyder's view of deterrent forces and Schelling's 'deterrent defense' both were designed to impose costs on an adversary with the goal of rendering an offensive military action too costly for an adversary to pursue, whereas defense was about survival. Over time, the literature evolved to regard the threat of an attacker meeting Schelling's "deterrent defense" as deterrence by denial. Indeed, conveying to an enemy that they will not be able to achieve success on the battlefield at an acceptable cost is a form of deterrence.[58]

John Mearsheimer asserts that deterrence by denial tends to be most associated with the employment of conventional forces to prevent an enemy from achieving battlefield success.[59] In the conventional realm, the efficacy of deterrence is conditional on two variables: first, the costs the attacker would have to pay in terms of material resources, military and civilian casualties, and other related costs associated with the mobilization, deployment, fighting, and maintenance of the force; and second, the probability of success, which is a function of time.[60] To wit, if an attacker believes she has a high probability of achieving victory in a short time span, then she is less likely to be deterred. However, if victory is only going to be achieved after a protracted

---

[57] Schelling, *Arms and Influence*, 78-79.

[58] John J. Mearsheimer, *Conventional Deterrence* (Ithaca: Cornell University Press, 1983), 14-15.

[59] However, Mearsheimer does note that there is a small school of thought that discusses nuclear deterrence by denial and lists amongst other works the following: Bernard Brodie, *Escalation and the Nuclear Option* (Princeton: Princeton University Press, 1966); Samuel T. Cohen, "Mini-Nukes and Strategy," *Orbis* 15 (Spring 1971): 178-193; Henry A. Kissinger, *Nuclear Weapons and Foreign Policy* (New York: Harper and Row, 1957); and John P. Rose, *The Evolution of U.S. Army Nuclear Doctrine, 1945-1980* (Boulder: Westview, 1980).

[60] Mearsheimer, *Conventional Deterrence*, 23.

www.manaraa.com

conflict then it is less likely to succeed because of the perceived costs that would accumulate over time.[61] In sum, when costs of war are low and the probability of success is high, deterrence is unlikely to work.

Unlike deterrence by punishment, which is hampered by fundamental problems that make the likelihood of success minimal, deterrence by denial is attainable in cyberspace. Indeed, there is already emerging literature making claims to this effect. For instance, in recent scholarship, both Patrick Morgan and Martin Libicki have suggested that the logic of deterrence by denial applies to cyber conflict, but is problematic to implement. Indeed, Morgan notes that achieving deterrence by denial would be less costly and also bring no damage to the state compared to deterrence by punishment, but he notes that it may be difficult to demonstrate to an attacker the cost they might incur by attacking.[62] Libicki also notes that psychological factors may confound effective deterrence by denial. Specifically, he argues that "the dynamic nature of cyberspace can convince one that targets that seem impregnatable today may be vulnerable tomorrow simply because things change all the time, so keep trying."[63] Yet, neither has explored the concept in depth or exposed it to much academic scrutiny. Therefore, by building on the nuances in the deterrence literature on the distinctions between deterrence and defense, it is imperative to consider the separate logics of deterring an adversary in cyberspace through building up one's

[61] Ibid., 24.

[62] Morgan, "Applicability of Traditional Deterrence Concepts and Theory to the Cyber Realm," 55.

[63] Martin Libicki, *Cyberspace in Peace and War* (Annapolis, MD: Naval Institute Press, 2016), 271.

www.manaraa.com

own defenses versus doing so through targeting the adversary's offensive capabilities, both of which fall under the rubric of deterrence by denial.[64]

Deterrence by denial is possible in the cyber domain if a state builds up its defensive capabilities such that the would-be attacker calculates that the costs of an offensive cyber operation would be unreasonably high and/or the probability of success small. There is a widespread assumption in the academic and policy literature that offensive cyber operations are relatively easy and that defense is difficult, if not impossible. For instance, while serving as U.S. Deputy Secretary of Defense William J. Lynn III wrote that,

> In cyberspace, the offense has the upper hand. The Internet was designed to be collaborative and rapidly expandable and to have low barriers to technological innovation; security and identity management were lower priorities. For these structural reasons, the U.S. government's ability to defend its networks always lags behind its adversaries' ability to exploit U.S. networks' weaknesses.[65]

However, the manner in which cyber weapons are developed and employed belies this assumption regarding the ease of offense. The target-specific nature of most cyber weapons that can produce strategic effects (e.g., those that target critical national infrastructure or military systems) means that adequate defenses render a highly specific capability in which a state invested research and development time and costs essentially useless. In other words, once a cyber weapon is discovered, it becomes a sunk cost because protections can be built relatively quickly to make future deployments ineffective, which means that an attacking state would have to invest time and resources to develop new capabilities and exploits and find a means of gaining access to deliver them to continue to pursue its objective, while also holding some degree of

[64] For a very cursory discussion of this, see Charles L. Glaser, "Deterrence of Cyber Attacks and U.S. National Security," George Washington University- Cyber Security Policy and Research Institute, Report GW-CSPRI-2011-5, June 1, 2011, 2.

[65] William J. Lynch III, "Defending a New Domain," *Foreign Affairs* 89, no. 5 (2010): 97-108, 99.

www.manaraa.com

uncertainty about whether the new capability will actually produce the intended effect. This challenges the widely held assumption that offense has the advantage in cyberspace because the target-specific nature of cyber weapons means that calculating the relative cost of an offensive operation does not involve the simple additive cost of deploying more of the same assets against a given target. Moreover, the specificity of offensive cyber operations suggests that the relative ease of offense or defense cannot be logically calculated as a systemic variable. Effective defenses, therefore, affect a target's cost calculus because they would require her to invest in building a potentially entirely new set of capabilities that could again be neutered by further defensive measures. Depending on the political objectives of the would-be attacker, she may find it more appealing to simply use cyber capabilities to target a less well-defended state. In other words, in the cyber domain, defense is relative; a target's networks and systems simply have to be slightly less difficult to penetrate than another.[66] Even in response to a persistent and dedicated attacker, there are a plethora of defensive measures that a state can take to make an attack against it unpalatable and, therefore, serve a deterrent function. [67] Defense of networks and systems in cyberspace is layered, as illustrated in Table 2.

Defense-in-depth in cyberspace, as Table 2 illustrates, encompasses physical, human, network, and host security. This may appear to imply, therefore, that an attacking state has many avenues through which it can penetrate a target and, thus, deterrence by defense is a fruitless exercise. However, the dynamics of defense in cyberspace in fact suggest that, even when an

[66] However, this is more likely to be applicable to criminal entities using cyber tools for the purposes of financial gain, rather than nation-states with specific political objectives that are not easily transferable to other more vulnerable targets.

[67] These defensive measures are distinct from active defenses, to be discussed in the subsequent section, which are directed at attriting the adversary's offensive capabilities rather than defending one's own networks and systems.

www.manaraa.com

adversary has penetrated a defending state's network, the latter can still force the attacking state to pay high costs for attempting to surmount its defenses, and the attacker will quickly come to surmise this fact as it employs a trial and error approach. In the logic of deterrence, a well-defended state can force an adversary to waste time and money trying to surmount its defenses such that it may calculate that the costs of persisting are too high and/or that the likelihood of success is too low.

www.manaraa.com

Table 2: Defensive Measures in Cyberspace

Type of Defensive Measure	Description	Means of Access it Counters
Physical Security of the System		
Wireless access points	Ensure there are no "rogue" wireless access points.	Physical and on-net
Safeguarding supply chain	Prevent adversaries from interdicting supply chain that provides hardware and software components.	Physical
Air gapping	Physically isolate networks and systems at facility from Internet.	Physical and on-net
Facility security	Physically security facilities on which networks and systems reside.	Physical and on-net
Training and personnel management	Ensure all personnel are screen and practicing proper computer hygiene.	Physical and on-net
Network Architecture and Management		
Intranet versus Internet	Control points of access of internal networks to Internet.	On-net
VPN technology	End-to-end securing of remote user's access to network.	On-net
Third party access	Screen and limit vendors and contractors' access to network.	Physical or on-net
Firewalls	Filters network traffic to scan for known attack signatures and preclude access to blocked content; typically emplaced at a gateway between a network and the Internet.	On-net
Network-based intrusion prevention systems	Programs that are designed to block network traffic associated with an attack.	On-net
Network-based intrusion detection systems	Programs that scan network traffic looking for signs of an attack and alert network administrators if they detect indications.	On-net
Compartmentalization	Identifying and segregating sensitive data assets.	On-net
Routine network maintenance	Ensuring firmware is updated on networked infrastructure (e.g., routers, switches).	On-net
Quarantining traffic	Create a "DMZ" for suspect traffic to facilitate identification and early warning of malicious activity.	On-net

www.manaraa.com

Honeypots[68]	Appealing traps to distract and isolate potential attackers.	On-net
Network obfuscation	Prevent attackers from mapping entire network through constructing false network typologies	On-net
Host Security		
Passwords and advanced forms of identification	Security protocol that limits unauthorized access to the network and the end user's computer.	Physical or on-net
Host-based encryption	Encrypting work stations and data storage.	Physical or on-net
Host-based intrusion detection systems	Similar to network detection systems, but based on host machines.	On-net
Personal firewalls	Programs that block malware access to/from the individual computer and the network.	On-net
Anti-malware tools	Programs that run on a host that hunt for known malware	On-net
Application Security		
Routine software updating	Ensure regular and time-sensitive patch management.	On-net

Deterrence by denial through degrading a target's offensive cyber capabilities is more complex, but still possible. One of the defining features of cyber weapons, which has implications for both deterrence and escalation, is that they are nearly invulnerable—because they are comprised of zeroes and ones they can, in theory, always be regenerated and cannot be physically destroyed. However, this does not mean that a deterring state is entirely barred from efforts at undermining a target's offensive capabilities. First, a deterring state could target the other state's attack infrastructure, which refers to the infrastructure that enables the attack to be delivered through the Internet. These include but are not limited to servers and routers that are strategically emplaced even outside of the attacking state's sovereign borders. While these are replaceable, destroying them could serve as a delaying function. Second, a deterring state could engage in a range of "active defense" efforts that cut off the attack vector at its source. These

[68] Honeypots could also be used for the purposes of disinformation, to undermine an attacking state's confidence in the integrity of stolen data if information is revealed regarding the use of honeypots.

www.manaraa.com

defensive measures occur outside of one's network and include activities such as engaging with a third party (e.g., telecommunications providers) to cut off the origin of an attack; or emplacing a "logic bomb," which is malicious code that resides on the defending state's network that an attacker takes back to her own network and triggers a negative impact such as wiping files. Third, defending states could engage in so-called "hacking back," which entails penetrating an attacker's networks during the course of or following an attack for the purposes of destroying information with which the attacker absconded, stealing adversary information, wiping their networks, or other retaliatory measures. Finally, a state could employ Counter-Computer Network Exploitation (CNE) that involves going after the adversary's toolkit and developing custom defenses against it; or revealing information about the threat such that others can develop patches against the vulnerability that the cyber weapon exploits and in effect degrading the attacking state's offensive capabilities by rendering them useless.

Cross-Domain Deterrence

The above discussion was limited to deterrence by punishment and denial within the cyber domain. However, there is an emerging discussion among policymakers and scholars about the potential utility of cross-domain deterrence of cyber operations.[69] Indeed, in response to allegations of foreign interference in the 2016 U.S. Presidential Election, President Obama stated "I think there is no doubt that when any foreign government tries to impact the integrity of our elections ... we need to take action. And we will — at a time and place of our own choosing.

[69] For instance, see: Erik Gartzke and Jon Lindsay, "Cross-Domain Deterrence: Strategy in the Era of Complexity," (unpublished manuscript), July 15, 2014; Vincent Manzo, "Deterrence and Escalation in Cross-domain Operations: Where Do Space and Cyberspace Fit," *Joint Forces Quarterly*, no. 66 (Third Quarter 2012); James A. Lewis, "Cross-Domain Deterrence and Credible Threats," *Center for Strategic and International Studies*, July 2010; Libicki, *Cyberspace in Peace and War*, 264-265.

www.manaraa.com

Some of it may be explicit and publicized; some of it may not be."[70] The concept of cross-domain deterrence is far from new. As Robert Jervis points out, during the Cold War deterrence was not framed solely in terms of nuclear politics but, rather, was multidimensional. Deterring different domains and technologies "…were not segregated from one another but rather mixed in the cauldron of international politics."[71] This section defines cross-domain deterrence and discusses some of the issues with this approach to deterring offensive cyber operations.

While deterrence rests on the promise of inflicting a more costly (but within a similar category of) response, cross-domain deterrence relies on using unrelated elements of national power to deter an action. In other words, where deterrence relies on the notion of employing *like* to deter *like*, cross-domain deterrence relies on an *unlike* capability to threaten costs.[72] Given the ineffectiveness of deterrence by punishment and the potentially unpredictable efficacy of deterrence by denial in the cyber domain, the incentive for decision makers to employ cross-domain deterrence is high. This incentive may be even higher for states that possess an asymmetric advantage over others in a different element of national power. However, cross-domain cyber deterrence is plagued by several issues. First, cross-domain responses are typically sent in the open where both domestic and international audiences are privy to the response, whereas a cyber response can be sent clandestinely, and avoid public scrutiny. Second, some non-cyber mechanisms, such the imposition of economic sanctions, may require a multilateral response to be effective. Both issues reflect the fact that cross-domain responses require shining

[70] Barack Obama, "Obama on Russian Hacking: 'We Need to Take Action. And We Will,'" *NPR- Morning Edition*, December 15, 2016, http://www.npr.org/2016/12/15/505775550/obama-on-russian-hacking-we-need-to-take-action-and-we-will.

[71] Robert Jervis, "Some Thoughts on Deterrence in the Cyber Era," *Journal of Information Warfare* 15, no. 2 (2016): 66-73, 67.

[72] Gartzke and Lindsay, "Cross-Domain Deterrence," 2.

www.manaraa.com

a light on activities that would otherwise be kept in the dark and, to justify them to the public, the deterring state may feel compelled to share attribution information or other intelligence it may otherwise prefer to keep secret. Third, developing a proportionate cross-domain response to virtual damage is difficult because no norms exist that equate virtual damage with either economic or material costs (this will be explored in greater detail below in the discussion of escalation).

The enumerated issues above suggest that cross-domain cyber deterrence is beset by serious complications. In particular, any multilateral cross domain response would require common agreement on what actions are to be deterred. This is problematic because great uncertainty exists regarding the "rules of the road" in the cyber domain. For instance, when Estonia came under crippling cyber attacks in 2007 the question arose if Estonia President Jaak Aaviksoo could invoke Article 5's mutual defense clause of NATO's Charter.[73] In the wake of the attack, NATO responded by developing the NATO Cooperative Cyber Defence Centre of Excellence located in Tallinn, Estonia to address issues of this nature. In mid-2016, NATO Secretary General Jens Stoltenberg went on the record and stated that cyber attacks would meet the requirements for an Article 5 response, however the harm that must first be inflicted remains an ambiguous threshold.[74]

[73] Joshua Davis, "Hackers Take Down the Most Wired Country in Europe," *Wired*, Issue 15.09, August 21, 2007, http://archive.wired.com/politics/security/magazine/15-09/ff_estonia.

[74] Jens Stoltenberg, "Press Conference by NATO Secretary General Jens Stoltenberg following the North Atlantic Council meeting at the level of NATO Defence Ministers," North Atlantic Treaty Organization, June 14, 2016, http://www.nato.int/cps/en/natohq/opinions_132349.htm.

www.manaraa.com

Yet Secretary General Stoltenberg's rhetoric did not stop Russian interference in the 2016 U.S. presidential election or that of other NATO members.[75] Following reports of Russian interference during the 2016 election, President Obama became concerned that Russia would use cyber attacks to interfere with US election systems and effectively undermine the integrity of the vote on election day. As mentioned above, Obama chose a cross-domain response. First, during the G20 summit in China, President Obama threatened Russia's President Putin that there would be "consequences" if interference did not stop. As the continued release of hacked emails from the Democratic National Committee continued, a month later Obama used the hotline connection between the Nuclear Threat Reduction Centers, which only three years prior had been bilaterally designated as being used for cyber related events, to convey to President Putin that the laws of armed conflict applied to cyberspace.[76] According to reporting, Russian hacking stopped and the electoral system was not directly attacked.[77] However, it is not clear whether this represents an example of successful cross-domain deterrence. Though US electoral systems were not attacked and the integrity of the vote has not been questioned, Putin may have assessed that his influence

[75] Julie Hirschfeld Davis and Maggie Haberman, "Donald Trump Conceded Russia's Interference in Election," *The New York Times*, January 11, 2017, https://www.nytimes.com/2017/01/11/us/politics/trumps-press-conference-highlights-russia.html. The Editorial Board, "Russian Meddling and Europe's Elections," *The New York Times*, December 19, 2016, https://www.nytimes.com/2016/12/19/opinion/russian-meddling-and-europes-elections.html.

[76] William M. Arkin, Ken Dilanian And Cynthia Mcfadden, "What Obama Said to Putin on the Red Phone About the Election Hack," *NBC News*, December 19, 2016, http://www.nbcnews.com/news/us-news/what-obama-said-putin-red-phone-about-election-hack-n697116.

[77] David E. Sanger, "White House Confirms Pre-Election Warning to Russia over Hacking," *The New York Times*, November 16, 2016, https://www.nytimes.com/2016/11/17/us/politics/white-house-confirms-pre-election-warning-to-russia-over-hacking.html.

www.manaraa.com

operation had already achieved its major objectives. It remains unknown how credible Putin perceived Obama's threat to be.

Assessing Deterrence in Cyberspace

The above analysis explored some of the key problems associated with effective deterrence in cyberspace. Unlike nuclear weapons, where demonstrating resolve was the crux of success and confounded scholars and policymakers due to the immensely destructive potential of nuclear war, the fundamental conundrum of deterrence is cyberspace is capability. Few doubt a state's desire to employ cyber weapons, and indeed, offensive cyber operations have become common state practice. However, the inability to credibly signal cyber capabilities and resolve confounds both deterrence by punishment and, to a lesser degree, deterrence by denial. In the context of deterrence by punishment, cyber actors cannot conduct shows of force to demonstrate their capability because doing so runs the risk that the target can customize its defenses to thwart a future action and potentially render the cyber weapon inert. The incentive to conceal capabilities to protect operational viability trumps public measures to enhance credibility. Furthermore, since cyber weapons lack universal lethality, states are limited by the cyber weapons they have on hand and may not have a proportionate response available in their arsenal. Most important for deterrence by punishment, the destructive potential of even a large-scale cyber onslaught against critical national infrastructure is simply insufficiently painful to affect a target's calculus. Indeed, the limit on destruction is the most significant impediment to successful deterrence by punishment. On the other hand, deterrence by denial presents more feasibility for states because there are a range of actions states can take to defend themselves in cyberspace and, therefore, raise the costs to an attacker, as well as to degrade the latter's offensive capabilities. This suggests, therefore, that the prevailing assumption regarding an

www.manaraa.com

offensive advantage in cyberspace is misguided. Due to the limits of deterrence within the cyber domain coupled with the persistent desire to prevent aggressive cyber actions, states have relied on cross-domain solutions, particularly if the deterring state maintains an asymmetric advantage in another element of national power. However, some of the same hurdles associated with secrecy and proportionately that confound deterrence within the cyber domain also apply to deterring cyber attacks through non-cyber means.

The difficulties of deterrence have led some to suggest that interstate relations in the cyber domain are likely to be plagued by dangerous escalatory spirals. The follow section explores why this logic is fallacious.

Escalation Dynamics in Cyberspace

As the above discussion demonstrated, there are substantial barriers to achieving stable deterrence in cyberspace. *Prima facie*, therefore, one might infer that the cyber domain is fundamentally escalatory. Indeed, the emerging consensus among cyber scholars and practitioners is that the domain is fundamentally escalatory and the risks of escalatory spirals in this domain are significant. However, this presents an empirical puzzle: if there are attributes of the cyber domain that create the conditions for escalatory spirals between states, why have we not yet observed any significant escalation in the domain? For example, in response to Russia's hack of the Democratic National Committee's email server in an effort to influence the 2016 presidential election, President Obama imposed sanctions on Russian intelligence agencies, four operatives, and three Russian companies (the *Special Technology Center*, *Zorsecurity*, and the *Autonomous Professional Association of Designers of Data Processing Systems*); expelled 35 Russian diplomats; and shuttered two diplomatic compounds that the administration alleged were

www.manaraa.com

used for espionage.[78] While these sanctions were more costly than the ones the United States imposed on North Korea in the wake of the 2014 Sony hack, they were "not as biting as previous ones in which the United States and its Western allies took aim at broad sectors of the Russian economy and blacklisted dozens of people, some of them close friends of Mr. Putin's."[79] If the cyber domain were characterized by greater risks of escalation, one might have expected that an attack on a democratic state's election system would have prompted a more significant response. The mismatch between the predominant hypotheses in the emerging literature on escalation dynamics in cyberspace and the empirical reality could be attributed to two categories of factors. First, those who assert that the cyber domain is inherently escalatory may be utilizing a conception of escalation that differs from that of this paper. Specifically, following the traditional academic literature on nuclear escalation, this paper defines escalation as an action that increases the intensity of military action, either in quantitative or qualitative terms (see the discussion below for a more thorough definition). Others, however, may include in their concept of escalation actions that simply respond to another state's behavior at a level consistent with the latter. While it is possible that, given the absence of norms and indices in the cyber domain, responsive actions that do not represent an increase in intensity could risk inadvertent escalation due to misperception, this paper does not count these as examples of deliberate escalation. Similarly, proponents of the escalatory nature of cyberspace may identify escalatory behavior where this paper does not due to divergent conceptions of the role of espionage. In particular, this paper excludes from a conception of escalatory behavior espionage conducted through cyber

[78] David E. Sanger, "Obama Strikes Back at Russia for Election Hacking," *The New York Times*, December 29, 2016, https://www.nytimes.com/2016/12/29/us/politics/russia-election-hacking-sanctions.html. There may have also been a covert response, but there is no publicly available information as to its existence or nature.

[79] Ibid.

www.manaraa.com

means because it considers espionage to be a necessary state practice that is mutually recognized by governments as legitimate. This account for the U.S. (non)response to Chinese cyber espionage that resulted in the theft of tens of millions of records from the Office of Personnel Management. This led the US Director of National Intelligence at the time, James Clapper, to note that "You have to kind of salute the Chinese for what they did."[80] Conversely, if espionage is considered deliberately escalatory, than the empirical record is littered with examples of cyber escalation. However, while cyber espionage in particular (as distinguished from other forms of espionage) could potentially trigger inadvertent escalation due to the fact that governments may be unable to distinguish between access gained for the purposes of espionage versus attack, categorizing espionage as escalatory flouts most conceptions of customary international law.

Therefore, while some scholars and practitioners have broached the topic of cyber escalation dynamics, none have approached it with academic rigor and, therefore, the literature has largely ignored critical aspects of escalation that have been central to the traditional canon on escalation. This section will consider the applicability of the academic literature on escalation dynamics that emerged during the early days of the Cold War to the cyber domain. It presents a theoretical framework assessing the nature, likelihood, and causes of both inadvertent and deliberate escalation in cyberspace. For the purposes of this paper, I limit my theoretical analysis to escalation within the cyber domain, rather than cross-domain escalation, unless otherwise specifically noted.[81] Furthermore, I will argue that the prevailing view among cyber conflict experts that the domain is fundamentally escalatory is not necessarily true under most conditions.

[80] Damian Paletta, "U.S. Intelligence Chief James Clapper Suggests China Behind OPM Breach; Mr. Clapper says China is "leading suspect" in the theft of millions of personnel records," *The Wall Street Journal*, June 25, 2015.

[81] Future research could integrate escalation dynamics in cyberspace with other domains of warfare, similar to Kahn's escalation ladder.

www.manaraa.com

This is due to two interrelated factors: first, the self-dampening nature of carrying out operations in the cyber domain; and second, the distinct role played by time in cyberspace, which creates breathing room for decision-makers to assess options and responses to adversary actions.[82]

The section proceeds as follows. First, I develop a theoretical framework to account for why escalation has not been observed in the cyber domain, despite the issues associated with deterrence. Second, following the existing literature on escalation dynamics in cyberspace, I explore causes of inadvertent escalation in cyberspace. The current literature posits that the cyber domain is escalatory because it implicitly assumes that actors, for a variety of reasons that are unique or particularly acute in the cyber domain, will inadvertently cause escalatory spirals. However, this literature is not well-grounded in the traditional theoretical canon on escalation. Therefore, I review the traditional literature on inadvertent escalation and identify causes of inadvertent escalation. Then, I assess the extent to which these causes are applicable to the cyber domain and the implications for causes of inadvertent escalation in cyberspace. My analysis demonstrates that inadvertent escalation in cyberspace is not necessarily predetermined. Indeed, the decision to escalate may be extremely context-dependent on the risk aversion or acceptance of decision-makers and how they respond to conditions of the pervasive uncertainty that is acutely present in cyberspace.

Next, I assesses the extent to which deliberate escalation as part of a bargaining strategy is possible in the cyber domain. This has largely been ignored by the emerging literature on

[82] Joseph S. Nye, Jr., in a recent article on deterrence in cyberspace, acknowledges that states have demonstrated self-restraint in the domain. He attributes this to, "the sheer complexity and uncertainty of cyber systems. In addition to norms, Brandon Valeriano and Ryan Manness list factors such as replication of dangerous malware that 'escapes into the wild'; uncertainty about collateral damage; unanticipated escalation that involves third parties; and 'blowback' or retaliation." Nye, "Deterrence and Dissuasion in Cyberspace," *International Security* 41, no. 3 (Winter 2016/17), p. 70. However, Nye does not present a theoretical framework for this concept or explore it in any further depth.

www.manaraa.com

escalation dynamics in cyberspace, even though the original escalation literature in the nuclear era wrestled with the question of deliberate (rather than inadvertent) escalation. Specifically, I develop a theoretical framework for understanding deliberate escalation dynamics modeled after Herman Kahn's escalation ladder. I consider the extent to which deliberate escalation in cyberspace differs from, and is similar to, deliberate escalation as it is traditionally understood in the literature. My analysis suggests several important distinctions: first, as described below, there are multiple "worlds" in which deliberate escalation in cyberspace could occur, which are a function of the observability and attribution of attacks; second, there are no clear firebreaks as there are with the employment of nuclear weapons and other conventional weaponry; third, there is no universal metric for assessing virtual damage; and fourth, potential thresholds may be inadvertently crossed due to unintended second order effects. Then, based on the cyber escalation ladder, I derive implications for escalation dominance, brinksmanship, strategies for victory, and de-escalation in cyberspace.

Finally, I conclude by synthesizing the insights from the analysis of both inadvertent and deliberate escalation to explore why escalation has not yet occurred in this domain. In particular, I argue that the cyber domain is inherently self-dampening. Indeed, there is a lack of universal lethality of cyber weaponry which means that following a cyber attack states must elect to do nothing, decide to respond with whatever cyber capabilities they already have (which may be far from the ideal response), or escalate through a cross-domain reaction. However, the latter may be difficult to conceive as there are no targeting norms for assessing proportionate physical responses to virtual damage. Finally, any response, cyber or otherwise, requires attribution, which takes time. Both elements work together to create breathing space and stymie escalatory spirals. These themes are borne out in four case studies that explore the nature of the empirical

www.manaraa.com

puzzle by describing several recent cases where one might have expected to observe escalation, but did not. Indeed, by studying these "dogs that did not bark" one can identify the conditions that may have prevented escalation such as the United States' tempered response to the 2014 hack of *Sony Picture's Entertainment* by North Korea.

Defining Escalation

Escalation, put simply, is the increase in the intensity of military action—quantitatively through doing more of the same (e.g., increasing bombardment of a particular area), or qualitatively through introducing a new type of capability (e.g., moving from conventional to tactical nuclear weapons in a theater of operations)—or the scope (e.g., expanding the scope of military action to go beyond the current geographic theater to target the adversary's allies or homeland).[83] Academic debates regarding escalation dynamics flourished in earnest in the wake of the Cuban Missile Crisis, which crystalized the risks of nuclear annihilation due to crisis escalation.[84] Herman Kahn's seminal work, *On Escalation*, set the theoretical foundation for subsequent decades of scholarly debate on escalation dynamics in the context of nuclear deterrence.[85]

As described by Kahn, escalation is akin to Schelling's "competition in risk-taking" in which, in the context of a crisis, one side attempts to increase its effort to match the other's,

[83] Increasing intensity is often termed "vertical" escalation, while scope is termed "horizontal" escalation. See, for example, the discussion in Forrest E. Morgan, Karl P. Mueller, Evan S. Medeiros, Kevin L. Pollpeter, and Roger Cliff, *Dangerous Thresholds: Managing Escalation in the 21st Century* (Santa Monica, CA: RAND Corporation, 2008), 18.

[84] Indeed, Richard Smoke indicates that the concept of escalation "did not appear in dictionaries, in military or scholarly literature, or in the public statements of government officials, before about 1960." Richard Smoke, *War: Controlling Escalation* (Cambridge, MA: Harvard University Press, 1977), 4.

[85] Also see Smoke, *War;* and Bernard Brodie, *Escalation and the Nuclear Option* (Princeton: Princeton University Press, 1966).

www.manaraa.com

potentially producing an escalating spiral that could lead to nuclear war.[86] Parties to a crisis might choose to deliberately escalate to signal to an adversary and force the latter to back down, or escalation could occur inadvertently through security-dilemma dynamics, where an actor does not anticipate the consequences of what the other side interprets as escalatory actions.[87] Victory is achieved through maintaining what Kahn refers to as "escalation dominance," in which one side has a clear superiority at every level of escalation and can, therefore, prevent a crisis from spiraling out of control and achieve a desired political objective.[88]

Whether it is indeed possible to achieve escalation dominance became a pressing question for academics in the context of nuclear deterrence, because it implied that the United States (or the Soviet Union) could escalate in a rational, linear, and controlled manner to get its way in superpower crises without risking spiraling into nuclear Armageddon. Indeed, according to Kahn and other theorists, the structural condition of the nuclear balance of terror—the disproportionate devastation that would occur with the strategic use of nuclear weapons—was likely to "induce some degree of restraint and prudent behavior on each side."[89] In other words, the nuclear balance of terror during the Cold War was deterrence dominant. However, as the Cold War progressed and the idea of using nuclear weapons as part of a deliberate bargaining

---

[86] However, it is important to note that there is distinction between Kahn and Schelling's conceptualizations of competitions in risk taking, with the latter less controlled and deliberate escalation than the former. Schelling emphasizes the bargaining benefits associated with the risks of letting things spiral out of control, while Kahn focuses more on controlled escalation up and down the ladder.

[87] Herman Kahn, *On Escalation: Metaphors and Scenarios* (New Brunswick, NJ: Transaction Publishers, 2010), 3-4.

[88] Ibid., 23-24.

[89] Kahn, *On Escalation*, 13. But also see, for instance, Wohlstetter, "The Delicate Balance of Terror," 211.

www.manaraa.com

strategy came to be perceived by academics as illogical or anti-Clausewitzian (in that the costs of a nuclear confrontation would far outweigh whatever political objective was sought by their use).[90] Therefore, the academic literature shifted to focus on the conditions under which escalation might occur through inadvertent or accidental means.[91]

Inadvertent Escalation in Cyberspace

Taken together, the above examples suggest a non-escalatory dynamic exists in the cyber domain. However, the prevailing sentiment in the emerging literature on escalation dynamics in cyberspace is that inadvertent escalation is overdetermined due to unique properties inherent in the domain. Indeed, while there is some discussion of the employment of cyber arms as part of a deliberate bargaining strategy, most of the contemporary discussion focuses on how parties to a crisis may inadvertently be drawn into strategic cyber war, or how cyber war might inadvertently escalate into kinetic conflict.[92] For example, Lawrence J. Cavaiola et al. claim that avoiding inadvertent escalation in cyberspace is difficult due to unpredictable collateral damage and risks of unintended contagion, and problems of command and control.[93] Patrick Allen and Chris

---

[90] However, it is important to note that practitioners did not give up on the feasibility of using nuclear weapons. See, for example, Austin Long and Brendan Rittenhouse Green's discussion of counterforce targeting operations in the United States in "Stalking the Secure Second Strike: Intelligence, Counterforce, and Nuclear Strategy," *Journal of Strategy Studies* 38, no. 1-2 (2015): 38-73.

[91] The most notable examples are Barry R. Posen, *Inadvertent Escalation: Conventional War and Nuclear Risks* (Ithaca: Cornell University Press, 1991); and Scott D. Sagan, *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons* (Princeton: Princeton University Press, 1993. Also see Graham T. Allison, *Essence of Decision: Explaining the Cuban Missile Crisis* (Boston: Little, Brown, 1971).

[92] Herbert Lin engages in discussion of deliberate escalation in "Escalation Dynamics and Conflict Termination in Cyberspace," *Strategic Studies Quarterly* 6, no. 3 (Fall 2012): 46-70.

[93] Lawrence J. Cavaiola, David C. Gompert, and Martin Libicki, "Cyber House Rules: On War, Retaliation and Escalation," *Survival: Global Politics and Strategy* 57, no. 1 (February-March 2015): 84-94.

www.manaraa.com

Demchak posit that the existence of patriotic hackers—citizens who, motivated for nationalistic reasons, engage in cyber attacks against perceived enemies of the state[94]—creates conditions for the cyber domain to be particularly escalatory due to divergent motivations of patriotic hackers and the government.[95] Libicki argues that crisis management is likely to be more difficult in cyberspace than in conventional domains when the interests at stake are comparable. He also posits that that inadvertent escalation is more likely in cyberspace due to the fact that parties to a crisis may define thresholds differently and/or if thresholds are private; the involvement of third parties and the ensuing attribution difficulties; and heightened problems of command and control.[96] In an earlier piece, Libicki argues that the factors that distinguish the cyber domain from conventional ones makes the former more escalatory, such as the uncertainty surrounding the effects of cyberattacks; the asymmetric nature of the vulnerability in the domain, which could prompt escalation to conventional kinetic attacks; and the greater credibility of retaliatory threats.[97] Jason Healey has claimed that conflict in cyberspace is "the most escalatory kind of conflict we have ever come across."[98] Roger Hurwitz warns of the risk of conflict and escalation in cyberspace due to "ongoing quantitative and qualitative cyber arms races among state actors, the proliferation of cyber weapons among state and non-state actors, a widely shared

---

[94] Borghard and Lonergan, "Can States Calculate the Risks of Using Cyber Proxies?"

[95] Patrick D. Allen and Chris C. Demchak, "The Palestinian-Israeli: CYBERWAR," *Military Review* 83, no. 2 (March/April 2003). Also see Lin, "Escalation Dynamics and Conflict Termination in Cyberspace," 59-61.

[96] Martin C. Libicki, *Crisis and Escalation in Cyberspace* (Santa Monica, CA: RAND Corporation, 2012), 10, 93-97, 106-108, 114-119.

[97] Martin C. Libicki, *Cyberdeterrence and Cyberwar* (Santa Monica, CA: RAND Corporation, 2009), 69-74.

[98] Keynote address, *CyberTalks,* New York City, September 8, 2016.

www.manaraa.com

conventional wisdom that these weapons favor offense over defense, and no broadly accepted definition of the cyber attacks that amount to the use of force."[99] David C. Gompert and Martin Libicki postulate that escalation is particularly likely in the event of a crisis between the United States and China due to mutual expectations regarding each state's likely cyber strategy in the event of armed conflict.[100]

While it may be the case that some aspects of the cyber domain point to greater risks of inadvertent escalation, the above literature does not systematically review the causes of inadvertent escalation as they have been identified in the existing theoretical canon and assess their relevance or application to the cyber domain. The analysis below, therefore, identifies five causes of inadvertent escalation and assesses the extent to which the mechanisms and logics could play out in cyberspace. My analysis suggests that the current conventional wisdom that the cyber domain is fundamentally escalatory—that escalatory spirals can be easily triggered and, therefore, that relations between cyber rivals is inherently unstable—misses the nuances surrounding the nature of escalation in cyberspace. In fact, there are many factors that point in the direction of caution and deliberation and, therefore, mitigate the risks of inadvertent escalation. The literature on escalation dynamics in the context of nuclear weapons, as well as the traditional security studies literature, provides a wellspring of hypothesizing on the factors that may prompt inadvertent escalation. I explore the extent to which the factors that have been hypothesized to cause inadvertent escalation are applicable to the cyber domain.

---

[99] Roger Hurwitz, "Keeping Cool: Steps for Avoiding Conflict and Escalation in Cyberspace," *Georgetown Journal of International Affairs*, International Engagement on Cyber III: State Building on a New Frontier (2013-14), 17.

[100] David C. Gompert and Martin Libicki, "Cyber Warfare and Sino-American Crisis Instability," *Survival* 56, no. 4 (August-September 2014): 7-22. Also see Avery Goldstein, "First Things First: The Pressing Danger of Crisis Instability in U.S.-China Relations," *International Security* 37, no. 4 (2013): 49-89.

www.manaraa.com

The Security Dilemma and the Spiral Model

The security dilemma was originally conceived of as applying to relations between states during peacetime. However, scholars have used the underlying logic of the security dilemma to gain insight into how security dilemma dynamics could provoke unintended escalation during times of conflict.[101] Several factors lie at the crux of the security dilemma and the spiral model: the adverse effects of incomplete information under anarchy;[102] the cognitive biases of decision-makers; and the uncertainty associated with the inability to predict into the future.[103] The fundamental logic of the security dilemma is that the anarchic nature of the international system creates the conditions for arms races and dangerous spirals between status-quo orientated states. A security dilemma occurs when "an increase in one state's security decreases the security of others."[104] Actions that one state takes to make itself more secure could ultimately undermine its security if another state perceives those actions as threatening and takes measures to arm itself, provoking a reciprocal, potentially escalatory spiral.

Robert Jervis identifies two factors that contribute to the security dilemma: offensive versus defensive advantage; and the distinguishability between offense and defense.[105] Security

[101] Posen, *Inadvertent Escalation*, 12-13.

[102] Indeed, incomplete information has played a prominent role in the literature on the cause of war. See, for example, James D. Fearon, "Rationalist Explanations for War," *International Organization* 49, no. 3 (Summer 1995): 379-414.

[103] Jervis, *Perceptions and Misperceptions in International Politics*, ch. 3. The focus of this section, however, is on structural causes of the security dilemma and spirals, rather than individual psychology or cognitive biases.

[104] Jervis, "Cooperation Under the Security Dilemma," 186.

[105] There is a substantial literature on the offense-defense balance, which will not be reviewed in significant depth here. For further reading, see Charles L. Glaser and Chaim Kaufmann, "What is the Offense-Defense Balance and How Can We Measure It?" *International Security* 22, no. 4 (Spring 1998): 44-82; Stephen Van Evera, "Offense, Defense, and the Causes of

www.manaraa.com

dilemmas occur when defense is dominant ("when it is easier to destroy the other's army and take its territory than it is to defend one's own"[106]), but states cannot easily distinguish between offense and defense. Technology and geography are the primary factors that determine whether offense or defense has the advantage. In terms of geography, things that increase distance, cost, and vulnerability as an attacker advances to the defender contribute to a defensive advantage, while factors that decrease these contribute to an offensive advantage. In terms of technology, the vulnerability of weapons contributes to an offensive advantage because states will feel the need to use them prior to being attacked, while the invulnerability of weapons contributes to a defensive advantage.[107] In the context of nuclear weapons, perceived vulnerability of second strike capabilities plays a critical role in igniting escalation from conventional to nuclear conflict, according to Barry Posen, because it creates an incentive for a state to strike first with its nuclear

---

War," *International Security* 22, no. 4 (Spring 1998): 5-43; Stephen Van Evera, *Causes of War: Power and the Roots of Conflict* (Ithaca: Cornell University Press, 2013); Sean M. Lynn-Jones, "Offense-Defense Theory and Its Critics," *Security Studies* 4, no. 4 (Summer 1995): 660-691; Ted Hopf, "Polarity, the Offense-Defense Balance, and War," *American Political Science Review* 85, no. 2 (1991): 475-493; Stephen Biddle, "Rebuilding the Foundations of Offense-Defense Theory," *The Journal of Politics* 63, no. 3 (2001): 741-774; Jack Snyder, *The Ideology of the Offensive: Military Decision Making and the Disasters of 1914*, vol. 2 (Ithaca: Cornell University Press, 1989); and Jack S. Levy, "The Offensive/Defensive Balance of Military Technology: A Theoretical and Historical Analysis," *International Studies Quarterly* 28, no. 2 (1984): 219-238. Also, see the following critiques, Karen Ruth Adams, "Attack and Conquer? International Anarchy and the Offense-Defense-Deterrence Balance," *International Security* 28, no. 3 (Winter 2003/04): 45–83; Richard K. Betts, "Must War Find a Way?: A Review Essay," *International Security* 24, no. 2 (Fall 1999): 166-198; James W. Davis, et al., "Taking Offense at Offense-Defense Theory," *International Security* 23, no. 3 (Winter 1998/99): 179-206; and Sean M. Lynn-Jones, "Offense-Defense Theory and its Critics," *Security Studies* 4, no. 4 (Summer 1995): 660-691.

[106] Jervis, "Cooperation Under the Security Dilemma," 178.

[107] Ibid., 194-196.

www.manaraa.com

arsenal (or, to organize its nuclear forces so that they launch on warning, triggering the potential for escalation) to avoid an attack on its retaliatory capabilities.[108]

The distinguishability of offense and defense enables states to signal their "types" to others, because status-quo states will arm themselves with defensive capabilities, and it creates the conditions for advance warning of attack because states can observe others maneuvering offensive capabilities into position.[109] Distinguishability also enables status-quo states to devise arms control agreements to limit offensive weapons.[110] Jervis posits that, at most times in history, the international system is characterized by the conditions that enable security dilemmas.[111] Indeed, there are very few inherently defensive weapons and, therefore, a state's intent regarding the employment of weapons may be more significant than the characteristics of the weapons themselves.[112] And, of course, anarchy complicates states' efforts to signal intent because it creates incentives for obfuscation and signals can get lost in translation.[113]

The logic of the security dilemma serves as the foundation for the spiral model, which depicts how status quo states may nevertheless end up spiraling into unintended retaliatory cycles of conflict because each proffers up threats and punishment in an attempt to positively affect the other's behavior, when appeasement and conciliation would have been a more

---

[108] Posen, *Inadvertent Escalation,* 4.

[109] Jervis, "Cooperation Under the Security Dilemma," 199-200.

[110] Ibid., 201.

[111] Ibid., 213.

[112] Stephen Biddle, *Military Power: Explaining Victory and Defeat in Modern Battle* (Princeton: Princeton University Press, 2001).

[113] Robert Jervis, "Signaling and Perception: Drawing Inferences and Projecting Images," in *Political Psychology,* ed. Kristen Renwick Monroe (Mahwah, NJ: Lawrence Erlbaum Associates, 2002): 293-312.

www.manaraa.com

appropriate policy.[114] In other words, states miscalculate the nature of the world in which they are operating.

Applying the Security Dilemma to Cyberspace: Offensive vs. Defensive Advantage

One axis of the security dilemma is the extent to which offense or defense has the advantage. As has been extensively noted in discussions of the offense-defense balance in the literature, this variable is meant to operate at the systemic, rather than dyadic level.[115] Because there is no universal lethality of cyber weapons, it is conceptually difficult to conceive of offense or defense being dominant in the cyber domain. Therefore, the only way (from a technical perspective) to begin to quantify an offense-defense balance in cyberspace is dyadically.[116] With this caveat in mind, the analysis below explores the nature of geography and technology in cyberspace and the implications for an offensive or defensive advantage.

The prevailing wisdom in the emerging literature on escalation in cyberspace is that the "geography" of the cyber domain contributes to an offensive advantage, thus making escalatory spirals more likely.[117] From a superficial perspective, it may seem as though there are no geographic boundaries, buffers, and borders between states (or other actors operating in the domain) because anyone can theoretically reach out and touch anyone else in cyberspace. In other words, in cyberspace all states essentially share a common virtual border. The ostensible

[114] Jervis, *Perception and Misperception in International Politics*, Chapter 3. The implication of cognitive biases for escalation will be explored in greater detail in the section below on military organization and culture.

[115] See, for example- Lynn-Jones, "Offense-Defense Theory and Its Critics."

[116] Rebecca Slayton, "What is the Cyber Offense-Defense Balance?" *International Security* 41, no. 3 (Winter 2016/17): 72-109.

[117] William J. Lynn, "Defending a New Domain: The Pentagon's Cyberstrategy," *Foreign Affairs* 89, no. 5 (2010): 97-108.

www.manaraa.com

absence of cyber geography could, therefore, dramatically compress the time necessary to launch an attack and create heightened incentives to strike first, increasing the risk of inadvertent escalation through advantaging offense over defense.[118]

However, in reality, there are both physical and virtual barriers that impede access in cyberspace. For example, critical infrastructure is typically not connected to the Internet. Self-contained systems that are not connected to the Internet are said to be separated from the latter by an "air gap," which makes it physically isolated and, therefore, more difficult to penetrate. As already noted, gaining accesses to these systems requires close access through either physically penetrating a facility using a human operator or getting in sufficiently close proximity to connect via wireless or other means; interdicting supply chains; or securing a witting or unwitting insider to deploy the tool.

Furthermore, the nature of geography in cyberspace is unique in terms of its fluidity in a way that is not comparable to physical geography. For the most part, physical terrain cannot be altered. However, in cyberspace, accesses and barriers can change rapidly, unpredictably, and even unintentionally. This could occur through multiple means. At the tactical level, entities could employ basic defensive measures such as instituting user training and implementing new hardware and software updates that deny the specific vulnerabilities that attackers were exploiting. Operationally, depending on how data flows into and out of a country or region, states may force the active monitoring of content at gateways, thus creating choke points to detect nefarious activity. Finally, at the strategic level, states may wage a preemptive or preventive attack against the cyber capabilities of another state or come to agreements that limit

[118] Conceptually, this is analogous to the time from launch before an ICBM could pulverize an enemy's population centers. However, the critical distinction between nuclear weapons and other technologies is that the condition of mutual assured destruction makes nuclear weapons deterrence dominant.

www.manaraa.com

actions in the domain. Indeed, this occurred in 2015 when China's President Xi Jinping agreed to suspend cyber economic and industrial espionage against the United States.[119]

The full implications of this for offensive versus defensive advantage and the likelihood of escalation remain uncertain. Decision-makers may not be able to reliably calculate whether offense or defense has the advantage at any given moment given the dynamic and evolving nature of offensive and defensive capabilities. Even if they can do so, they are unlikely to have confidence that these calculations are stable over the short- to medium-term, let alone the long-term. In summary, this depicts a domain that may not necessarily be offense dominant. On the one hand, if an actor currently maintains access and fears that it may lose it in the future, this may create an incentive to strike first and, therefore, provoke an escalatory spiral; in other words, there is a "use-it-*or*-lose it" dynamic at play. On the other hand, the fact that states cannot reliably assess whether offense or defense has the advantage may prompt restraint because cyber weapons have a "use-it-and-lose-it" property and, therefore, deploying them under adverse conditions could result in the loss of an expensive access and/or capability.

The second factor that impacts the relative advantage of offense and defense is the nature of technology and, by extension, its vulnerability. The predominant viewpoint in the cyber literature is that the technology of cyber weapons contributes to an offensive advantage. William Lynn, for instance, posits that the environment is "offense dominant" in cyberspace: "Adept programmers will find vulnerabilities and overcome security measures put in place to prevent intrusions….Cyber warfare is like maneuver warfare, in that speed and agility matter most."[120] The argument that offense is dominant due to the technology of cyberwar largely rests on the

[119] This case will be elaborated below.

[120] William J. Lynn III, "Defending a New Domain: The Pentagon's Cyberstrategy."

www.manaraa.com

notion that there are numerous vulnerabilities that could be exploited and, therefore, defenders are always one step behind attackers. Indeed, as computing power and memory storage have improved, operating systems and applications have been able to provide users with increased functionality, but with increased performance come larger and more elaborate computer programs. For instance, Windows 95 had roughly 15 million lines of code when it debuted in 1995, compared to the 2001 release of Windows XP that contained over 35 million lines of code.[121] It is inevitable that, as more code is released, there will always be bugs that enable a malicious actor to manipulate the software. Microsoft, for example, typically finds 10-20 defects per 1000 lines of code during in-house testing, but releases software with 0.5 per 1000 lines.[122] That may seem like a low defect ratio, but it implies that Windows XP, with 35 million lines of code, contained over 17,500 vulnerabilities.[123] Though not every one of these bugs necessarily provides a means for a hacker to take over a machine, some do and, due to the sheer magnitude and complexity of these programs, it is impossible to write completely bug-free code. However, the skill set necessary to discover these vulnerabilities and develop custom software to exploit them is difficult and very costly, particularly if the targeted system is highly customized.[124] Governments have recognized this fact and, therefore, have adapted over time to protect their own supply chains, purchase locally, and develop highly tailored and customized programs. For

[121] Steve Lohr and John Markoff, "Windows Is So Slow, but Why?" *The New York Times,* March 26, 2007, http://www.nytimes.com/2006/03/27/technology/27soft.html.

[122] Steve McConnell, *Code Complete*, 2nd ed., (Redmond, Washington: Microsoft Press, 2004), 521.

[123] Ibid., 521. However, *Windows* is beating the industry average of 1-25 errors per 1000 lines of code for delivered software.

[124] For a further discussion on this point, see Borghard and Lonergan, "The Logic of Coercion in Cyberspace."

www.manaraa.com

example, China is working with Microsoft through its "Government Security Program" to gain access to the source code, conduct security reviews, and have Microsoft customize programs to meet their security requirements.[125]

Furthermore, it is the case that certain types of cyber attacks favor the attacker—for instance, "cheap, fast, and easy" operations such as DDoS attacks, which are designed to disrupt rather than destroy. However, these types of attacks are unlikely to cause significant harm and, therefore, it is doubtful that they could produce strategic effects.[126] Attacks that are truly costly for the target, such as those of a magnitude comparable to Stuxnet, are also costly to develop and implement. This suggests that, when it matters, the technology of cyber weapons gives an advantage to the defense, rather than the offense.

Another aspect of the argument that technology contributes to the likelihood of escalation is the extent to which a given technology is vulnerable and, therefore, whether there is a "use-it-or-lose-it" incentive. As mentioned above, this vulnerability informs Barry Posen's argument regarding the causes of escalation to nuclear conflict during the course of conventional war. This occurs when conventional military operations come in contact with a state's nuclear forces (or early warning and command and control systems) and threaten to undermine the latter's confidence in their ability to use them. In other words, when a conventional operation imperils the integrity of the adversary's nuclear retaliatory forces, the adversary is likely to escalate, particularly if it relies on a counterforce doctrine.[127] Secure nuclear retaliatory forces are the

---

[125] Joris Evers, "China Gets Access to Microsoft Source Code," *InfoWorld*, February 28, 2003, http://www.infoworld.com/article/2681624/operating-systems/china-next-to-get-access-to-microsoft-source-code.html.

[126] Borghard and Lonergan, "The Logic of Coercion in Cyberspace," Figure 1.

[127] Posen, *Inadvertent Escalation,* 2-3.

www.manaraa.com

bedrock of nuclear deterrence and negate a perceived first strike advantage by either side. When those forces are threatened, the perceived incentive to strike first reemerges, creating the conditions for a state to escalate to use nuclear weapons or position its nuclear forces to launch on warning to signal to the adversary its willingness to escalate.[128] During the 1980s, military planners devised plans to target the weak link in the command and control of the adversary's nuclear forces. Attacks on command and control are "one of the few sources of leverage in a nuclear war."[129] Inadvertent nuclear escalation can be more or less acute, according to Posen, depending on the state's nuclear strategy. In particular, Posen posits that a counterforce strategy increases the likelihood of inadvertent escalation because the perceived harm of conventional operations that target counterforce capabilities is greater and leads to perceptions of first strike advantage. Conversely, countervalue strategies are more immune to these risks.[130]

This model of inadvertent escalation is not a perfect fit for the cyber domain due to the nature of the vulnerability of cyber weapons. In the context of nuclear weapons, deterrence is stable when cities are vulnerable and weapons are invulnerable. Operations that threaten the invulnerability of weapons, therefore, undermine stability and increase the risk of escalation. However, cyber weapons are a unique class of weapons in that they are not in themselves inherently vulnerable. Therefore, the pressure to "use it or lose it" does not apply in the same way in this domain. Cyber weapons, as already described in a previous section, cannot be physically destroyed and, in theory, there is an unlimited capacity to regenerate them. Martin Libicki, for instance, asserts that,

---

[128] Ibid., 4.

[129] Ibid., 8.

[130] Ibid., 9-10.

www.manaraa.com

> The prerequisites of a cyberattack are clever hackers, cheap hardware, some network connection, intelligence on the workings and role of the target system, specific knowledge of the target's vulnerabilities, and tools to build exploits against such vulnerabilities. Cheap hardware possibly aside, none of these can be destroyed by a cyber attack (so, there is no basis for counterforce targeting in cyberwar)."[131]

However, there is an important nuance to this point that Libicki misses. While cyber weapons are themselves inherently invulnerable, states can target the attack infrastructure of their adversaries; yet this is a complex endeavor. For instance, while it is possible to conduct counter-force targeting and to destroy the networks that hold the cyber weaponry of another state, it is unlikely that this would be the only copy a state maintains of these weapons. Furthermore, a state could steal another state's toolkit and incorporate the signatures into its defenses to render a potential attack ineffective. However, capabilities do exist that enable attackers to repurpose a tool to overcome the target's defenses.[132] Another type of counterattack would be to disconnect the target state from the Internet or to destroy the networked infrastructure that it uses to employ the cyber weapon. Yet, while it would delay an attacking state, as discussed in the context of deterrence by denial, the impact of this would be temporary. Therefore, in the long run, cyber weapons have near-complete survivability.[133]

---

[131] Martin Libicki, "Pulling Punches in Cyberspace," *Committee on Deterring Cyberattacks, Proceedings of a Workshop on Deterring Cyber Attacks: Informing Strategies and Developing Options for US Policy*, National Research Council (Washington, DC: The National Academies Press, 2010), 124-125.

[132] One such capability is Metasploit's *Meterpreter*. For further reference see: "Metasploit's Meterpreter," *Metasploit*, https://dev.metasploit.com/documents/meterpreter.pdf. Also see the following for a good background reading on methods to get around defenses when a capability would otherwise be stopped by anti-virus software: xMidnightSnowx, "Tutorial: Evading Anti-Virus Software While Hacking," *Cybrary*, June 22, 2015, https://www.cybrary.it/0p3n/tutorialevading-anti-virus-while-hacking/.

[133] Herbert Lin makes a similar point when he asserts, "there is no conceivable way for one nation to eliminate or even significantly degrade the cyber attack capabilities of another." See "Escalation Dynamics and Conflict Termination in Cyberspace," 55.

www.manaraa.com

The implications of the survivability of cyber weapons on escalation are nuanced. On the one hand, the fact that capabilities could, in theory, always be regenerated removes the pressure to strike first in a crisis and therefore, contributes to stability rather than escalation. However, while tools are nearly invulnerable, accesses are not. Without proper access to deliver a cyber weapon at the appropriate time, the invulnerability of a tool is meaningless. What this means for escalation, therefore, is that the anticipation or the fear of losing an access—rather than the anticipation of losing a tool—could create escalatory pressure. The "use it or lose it" incentive applies in the cyber domain to the access, rather than the weapon. However, as noted above, the fluid and unpredictable nature of accesses in the domain means that states may not know with any reliability when they are likely to lose an access. Therefore, this does not uniformly suggest that this facet of the domain contributes to escalation.

Furthermore, the counterforce targeting described above may be particularly important during times of crisis or conflict when the temporary disruption of another state's offensive cyber capabilities may create a momentary military or bargaining advantage. Though the implications of this type of escalation are entirely theoretical, a state may find an advantage in this kind of escalatory move because it changes (albeit for a limited period of time) the relative vulnerability posture. Another temporary advantage that can change the vulnerability calculus is if one targets the command and control nodes of cyber weapons. For instance, in the United States only the President, and under some conditions, the Secretary of Defense, may order an offensive cyber operation.[134] If this command and control node can be decapitated at a strategic time, then an

[134] Headquarters, Department of the Army, "Field Manual 3-38: Cyber Electromagnetic Activities," February 12, 2014, 3-9, 3-10.

www.manaraa.com

adversary's employment of cyber weapons, at least theoretically, can be prevented until communication is reestablished.[135]

Applying the Security Dilemma to Cyberspace: Distinguishability of Offense and Defense

The indistinguishability of offense and defense also contributes to the likelihood of escalation. An inability to distinguish between offense and defense could stem from the structural conditions of an anarchic international system, where actors have imperfect information, as well as from the nature of the weapons themselves, which could have offensive or defensive purposes depending on how they are incorporated into a state's strategy and doctrine.[136]

There are several factors suggesting that the security dilemma and, therefore, the risk of escalation, may be heightened in cyberspace due to the difficulty of distinguishing between offensive and defensive cyberattacks. Most notably, states face considerable difficulty distinguishing between cyber intrusions that are conducted for the purposes of espionage versus those that are laying the groundwork for an offensive cyberattack. This is because gaining access and absconding with data could just as easily support intelligence objectives as it could offensive information requirements to needed to target the network or specific systems connected to it. Compounding the issue of distinguishability is the predominant role secrecy plays in the cyber domain. Secrecy complicates efforts to distinguish between offensive and defensive capabilities because it is difficult to observe how states are arming themselves, and states rarely get advance warning of an impending attack. Moreover, problems of attribution mean that a targeted state

[135] The conditions under with offensive cyber authorities can be and should be delegated is an area that is need to further research. A good starting point is C. Robert Kehler, Herbert Lin, and Michael Sulmeyer, "Rules of Engagement for Cyberspace Operations: A View From the USA," *Journal of Cybersecurity* 3, no. 1 (March 2017): 69-80.

[136] Biddle, "Rebuilding the Foundations," 741-744.

www.manaraa.com

may face significant barriers to identifying the adversary who committed an attack; and, even if attribution is possible, the intentions of the attacking states may be difficult to discern due to difficulties of signaling in the cyber domain.[137]

Taken together, these suggest that there is an increased chance of escalation due to the ambiguities associated with discerning between offense and defense and, more broadly, inferring intent. However, the factors that contribute to the risk of escalation are mitigated by the role of time. There are unique properties to cyber conflict that inject breathing space into crisis situations. The time lag that it takes between perceiving an attack and having confidence in attribution provides breathing room and political cover to decision-makers to deliberate on the appropriate response (and assess whether the government has the necessary capabilities and accesses at the time to deliver a proportionate response).[138] Problems associated with distinguishing between espionage and offensive cyberattacks, attribution, and intent create the risk that a state may respond disproportionately to a perceived cyberattack, and/or that they may respond against the wrong actor. These risks are as likely to induce caution as they are to incentivize escalation. Risk-adverse states may be more likely to use uncertainty as an excuse to pause before retaliating against a perceived cyberattack; risk-acceptant states may be more likely

[137] Erik Gartzke and Jon R. Lindsay, "Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace," *Security Studies* 24, no. 2 (2015): 316-348; and Borghard and Lonergan, "The Logic of Coercion in Cyberspace."

[138] Though there are no open source studies on the time lag between detection of an intrusion and attribution, Thomas Rid and Ben Buchanan note that, "Analyzing a well-executed operation in a narrow timeframe will be a significant challenge even for the most professional and best resourced teams, firms, and agencies. In serious cases, when high-level decisions will have to be made under pressure, the speed of political developments may outpace the speed of the attribution process." Thomas Rid and Ben Buchanan, "Attributing Cyber Attacks," *Journal of Strategic Studies* 38, no. 1-2 (2015): 4-37, 32.

www.manaraa.com

to use uncertainty as an excuse to retaliate against a perceived cyberattack to fulfill their own political ambitions.[139]

The Stability-Instability Paradox and Military Organization and Culture

There are additional concepts in the canon on inadvertent nuclear escalation that could be applicable to inadvertent escalation in the cyber domain, but remain beyond the scope of this paper and are, therefore, fruitful avenues for future research. First, an extension of nuclear deterrence is the concept of the stability-instability paradox, in which stable deterrence at the strategic level nevertheless creates the conditions for conventional conflict because nuclear war is so unthinkable that states are willing to risk conventional conflict; paradoxically, this could unintentionally spiral into a nuclear exchange, despite the stability of mutual assured destruction.[140] As the deterrence analysis above demonstrates, there is no stable deterrence in the cyber domain and cyber war is not "unthinkable." Therefore, the structural conditions of the stability-instability paradox may appear ill-suited to the cyber realm. However, there is emerging literature that attempts to apply this concept to cyberspace and, therefore, its utility warrants further exploration in light of the findings of this paper.[141] For instance, in future research one might explore the conditions under which low-level provocations and crises in the cyber domain risk spiraling into strategic conflict, given the unique attributes of the cyber domain that produce

[139] Randall L. Schweller, "Neorealism's Status Quo Bias: What Security Dilemma?" *Security Studies* 5, no. 3 (Spring 1996): 90-121; and "Bandwagoning for Profit: Bringing the Revisionist State Back In," *International Security* 19, no. 1 (Summer 1994): 72-107.

[140] Glenn Snyder, "The Balance of Power and the Balance of Terror," in Paul Seabury, ed., *The Balance of Power* (San Francisco: Chandler, 1965), 184-201. Also see, Robert Jervis, *The Illogic of American Nuclear Strategy* (Ithaca: Cornell University Press, 1984), pp. 29-34; and Robert Jervis, *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon* (Ithaca: Cornell University Press, 1989), 19-22.

[141] Jon R. Lindsay and Erik Gartzke, "Coercion through Cyberspace: The Stability-Instability Paradox Revisited," (working paper).

www.manaraa.com

self-dampening dynamics. The logic of the stability-instability paradox could also be explored in a cross-domain context (which, indeed, was how it was originally conceived in the realm of nuclear deterrence) to identify the likelihood of minimally costly but potentially disruptive cyber provocation spilling over into kinetic conflict between adversaries.

There is also an extensive body of literature on the topic of military organization and culture and its effects on escalation through various mechanisms, such as the cult of the offensive, strategic culture, risk-aversion or acceptance, and loose vs. tight coupling, to name a few.[142] For instance, Posen argues that certain types of civil-military arrangements could heighten the risk of inadvertent nuclear escalation. Specifically, inadvertent escalation could occur if there is a divide between civilian leaders and military planners, with the latter having significantly greater knowledge than and closely holding war plans from the former. Furthermore, if civilian leaders perceive nuclear weapons to be "exotic," they may have difficulty foreseeing the consequences of their use and, therefore, take actions that risk escalation.[143] In the cyber domain, lack of expertise among civilian leaders about the nature of cyber warfare and an aversion to understanding what is perceived of as a highly technical field may make it difficult for civilian policymakers to understand and appreciate the escalatory risks of war plans, potentially empowering the military and playing into cult of the offensive logic. There may also be significant variation within military leadership concerning cyber competence, which could result in a small, elite cadre of cyber warriors to whom both civilian and military

[142] Jack Snyder, "Civil-Military Relations and the Cult of the Offensive, 1914 and 1984," *International Security* 9, no. 1 (Summer 1984): 108-146; Barry Posen, *The Sources of Military Doctrine: France, Britain, and Germany between the World Wars* (Ithaca: Cornell University Press, 1986); Elizabeth Kier, *Imagining War: French and British Military Doctrine Between the Wars* (Princeton: Princeton University Press, 1997); Scott D. Sagan, *The Limits of Safety* (Princeton: Princeton University, 1993).

[143] Posen, *Inadvertent Escalation,* 13-14.

www.manaraa.com

leadership grant significant autonomy and decision-making authority, further enhancing the risk of inadvertent escalation. Finally, the critical role played by the private sector and non-state actors more broadly in the cyber domain, and their relationships with the government, introduce another potential source of inadvertent escalation that is distinct from other domains of warfare. For example, the extent to which the private sector is granted discretion to take active measures to defend its networks and systems and the level of oversight exerted by the government could risk that actions taken by non-government entities against foreign targets could draw a government into an unwanted crisis or cyber war. Variation in civil-military and public-private relationships and the implications for escalation is a topic for further research.

Finally, the literature on strategic culture offers rich and extensive opportunities for further research and many axes of variation for scholars to explore. Variation in strategic cyber culture across states could explain differences in states' cyber strategies, and decisions about cyber force structure and employment. Using the emergence of a militarized cyber domain as an exogenous event, researchers could assess the conditions under which states' approaches to the cyber domain are consistent with versus diverge from its existing strategic culture. Within states, variation across cyber actors with different strategic or organizational cultures (public versus private; civilian versus military; military versus intelligence; different military services) could explain divergent approaches to the domain.

A Unique Facet of Inadvertent Escalation in Cyberspace: Cyber Proxies

Perhaps the most likely cause of inadvertent escalation in cyberspace is a factor that the existing literature on inadvertent escalation does not consider, namely, the role played by cyber proxies. As articulated in past research, the desire for plausible deniability in the cyber domain creates an incentive for states to work with proxy actors to conduct cyber warfare on their behalf.

www.manaraa.com

Indeed, cyber proxies provide political cover for states even as the technical ability to assign attribution is improving. When proxy actors are inserted into a crisis, regardless of domain, there is an increased risk that these actors may intentionally or inadvertently exceed their mandate and escalate a crisis beyond what is intended.[144] Nevertheless, in addition to the technical factors that dampen the risk of inadvertent escalation in cyberspace regardless of the existence of proxies, uncertainty regarding command and control of cyber operations may also prompt the target of a cyber attack to pause before issuing an escalatory response rather than respond erroneously against the wrong actor. In this way, employing cyber proxies may actually enable states to get away with cyber attacks and avoid retribution.

Deliberate Escalation in Cyberspace

The above analysis has demonstrated that there are several properties inherent in the nature of the cyber domain that mitigate the risk of inadvertent escalation. A separate consideration is whether deliberate escalation—where states intentionally ratchet up the scale or scope of operations against an adversary in cyberspace—is possible and, if so, what the nature of that escalation could theoretically look like. The elements of cyber operations that alleviate inadvertent escalation risks through generating self-dampening dynamics—for instance, the absence of universal lethality of cyber weapons; the difficulty of generating proportionate responses; attribution difficulties and the role played by time—also complicate efforts to develop and implemented a deliberate escalation strategy. Nevertheless, this section uses the deliberate escalation literature, particularly Herman Kahn's work on escalation, as a framework for theoretically assessing the dynamics of deliberate escalation in cyberspace. I argue below that

[144] Indeed, employing proxy actors could indeed be a dangerous because governments risk a dilemma of inadvertent crisis escalation by empowering proxies with more expansive, or less restrained, political agendas that may exceed their mandates. See Borghard and Lonergan, "Can States Calculate the Risks of Using Cyber Proxies?"

www.manaraa.com

there is a critical conceptual difference between deliberate escalation in the nuclear and cyber realms due to the role played by secrecy in the cyber domain. In the latter environment, deliberate escalation could occur in three possible worlds—obfuscation, covert coupling, and overt coupling—each of which contains its own logic and is distinguished by the manner in which secrecy permeates cyber operations. Furthermore, deliberate escalation in cyberspace is also distinct in that there is no clear firebreak comparable to that between the deliberate non-use versus use of nuclear weapons.

This section proceeds as follows. First, I define deliberate escalation and describe the logic of the three worlds of deliberate escalation in cyberspace. Next, I construct an escalation ladder for the cyber domain. I subsequently explore the implications for how we can conceptualize escalation dominance in cyberspace; brinksmanship; and strategies of victory. I conclude by assessing the prospects of de-escalation in cyberspace.

Deliberate Escalation

In the context of nuclear warfare and nuclear options, the idea of escalation as part of a deliberate bargaining strategy became appealing for American policymakers in the context of the doctrine of Flexible Response, where the U.S. sought "seemingly numerous *possibilities* for *controlled* escalation, in deliberate gradations through an extended range of violence."[145] It was assumed, in this approach to controlled escalation, that policymakers would actually be able to deliberately manipulate levels of violence, ratcheting it up or down in a controlled manner.[146] The seminal framework for conceptualizing deliberate escalation in the nuclear realm was

[145] Smoke, *War*, 13.

[146] Ibid., 4. Smoke also acknowledges that this represented something of a hubris of American practitioners and strategists, who believed that "escalation processes in war can be restrained and halted and war kept limited, and…that escalation can be deliberately managed to manipulate the ongoing level of violence," p. 13.

www.manaraa.com

conceived by Herman Kahn, who used the metaphor of a ladder to epitomize escalation dynamics. The basic framework of the escalation ladder involved "a linear arrangement of roughly increasing levels of intensity of crisis."[147] In Kahn's conception of deliberate escalation, states could intentionally move up or down the escalation ladder to win in a crisis. For Schelling, state could deliberately escalate for the purposes of signaling in the context of coercion.

Three Worlds of Deliberate Escalation in Cyberspace

There are several factors that make deliberate escalation in cyberspace unique. First, a fundamental aspect of nuclear escalation dynamics is that they occur in the context of a two-sided crisis.[148] However, with cyber operations it may be more likely that there are third parties that are participants. Indeed, states that lack appropriate capabilities to carry out offensive cyber operations or seek additional (political) plausible deniability beyond what the cyber domain already provides at a technical level may employ cyber proxies to conduct the operation.[149] That said, for the purposes of this analysis I employ a two-party escalation model for cyberspace, although exploring the implications of multi-party escalation models is a fruitful avenue for future research. Furthermore, a two-party escalation model is theoretically possible even if one or both sides employ proxy actors as long as either side chooses to hold the other state responsible for the actions of a proxy actor.

The most theoretically significant difference between deliberate escalation in the nuclear and cyber realms is the implication of the roles of secrecy and attribution in cyberspace. Deliberate escalation in cyberspace could occurs in three different worlds (see Figure 1

[147] Kahn, *On Escalation*, 38.

[148] Kahn, *On Escalation*, 37.

[149] Borghard and Lonergan, "Can States Calculate the Risks of Using Cyber Proxies?"

www.manaraa.com

below).[150] Each world is based on the extent to which the escalating states prefer to reveal its role in a cyber attack.[151] The more information that is revealed regarding attribution, the less applicable are the self-dampening aspects of escalation in cyberspace (which in the case of this situation are primarily driven by the time required to ascertain attribution). In the first world, which I term "Obfuscation," the escalating state does not make any deliberate efforts to self-attribute; rather, attribution is the responsibility of the target. This world is by far the most common world in which interaction in cyberspace currently occurs. A notable exception to this is when former Secretary of Defense Ash Carter took public responsibility for a U.S. cyber operation against ISIS in Mosul, Iraq.[152] The first world is characterized by the anonymizing features the cyber domain provides to actors and thus, offensive actors have plausible deniability and the targets are forced to pursue attribution. As already discussed above, attribution takes time, resources, and rarely can be leveraged with complete confidence.[153] This process creates

[150] One central puzzle that is beyond the scope of this paper is the dynamics of mutual escalation across multiple domains. One interesting point, is that the world has not observed a military cross domain response to a cyber attack. For instance, if North Korea had invaded the mainland US, broke into the Sony Pictures Entertainment California offices, physically destroyed their email servers and threatened destruction on par with 9/11, then we would most likely have seen a proportionate or escalatory military response. However, presumably because hacking lacks a violent aspect, despite gross damages that may exceed physical destruction, cyber attacks have not netted a military counter response. Further scholarship is necessary to assess what is a comparable measure of virtual versus physical destruction.

[151] Attribution of cyber attacks is comprised of two components: technical attribution, which enables a target to identify to locus of an attack, and political attribution, which refers to the responsibility for the political decision to launch an attack. Achieving technical attribution does not guarantee that political attribution is possible; conversely, targets can use other intelligence sources to ascertain political attribution even if technical attribution is not feasible.

[152] Ash Carter, "In Fight Against ISIS, U.S. Adds Cyber Tools," *National Public Radio*, February 28, 2016, http://www.npr.org/2016/02/28/468446138/in-fight-against-isis-u-s-adds-cyber-tools.

[153] As previous noted, there is a need for research on the mean time to attribution. Rid and Buchannan argue that quality attribution requires tactical, operational, and strategic level data-

www.manaraa.com

breathing space between potential escalating volleys of cyber attacks and responses, even when states are pursuing a deliberate escalation strategy.

Figure 1: Three Worlds of Deliberate Escalation in Cyberspace

In the second world, which I term "Covert Coupling," an escalating state privately takes responsibility for a cyber attack through covert coupling. This private attribution could occur through a secretive diplomatic missive or perhaps even technical messages embedded in attack signatures or lines of code.[154] Indeed, secrecy has been used by actors outside of the cyber domain to convey resolve without signaling to domestic or international audiences.[155] By privately revealing information about the identity of an attacker, the target is provided with a

---

"The tactical goal is understanding the incident primarily in its technical aspects, the *how*. The operational goal is understanding the attack's high-level architecture and the attacker's profile — the *what*. The strategic goal is understanding who is responsible for the attack, assessing the attack's rationale, significance, appropriate response — the *who* and *why*." "Attributing Cyber Attacks," 10. This constellation of data may require inputs from numerous organizations and experts to include, intelligence professionals, technical forensic specialists, and national security experts, and require a concerted effort to sync the data into a coherent and quality assessment. Complete attribution may not be achievable because of lack of data at any level and, as previously noted, this effort may demand significant time and resources which are beyond the abilities of all but the most capable actors.

[154] For further discussion on what this type of messaging might look like so that the target has confidence in where it came from see, Borghard and Lonergan, "The Logic of Coercion in Cyberspace," 459.

[155] For an excellent work on the historic mechanisms and motivations of covert signaling see, Carson and Yarhi-Milo, "Covert Communication: The Intelligibility and Credibility of Signaling in Secret."

www.manaraa.com

decision-making environment in which the reputational costs of not responding are not nearly as acute, because the only actor who would perceive the non-response is the attacking state. Thus, while the latter may make calculation about the target's capabilities or resolve based on a non-response (or a non-escalatory response), the target is shielded from domestic or international reputational costs.

The third world, "Overt Coupling," most closely resembles escalation dynamics in the nuclear realm because it lacks the self-dampening features that typically exist in the cyber domain. In this world, the attacker uses overt signaling to claim responsibility for an attack. This world envisions attribution being coupled with a formal diplomatic and/or a technical message, such as a defacement, that is widely and publicly observed and the authenticity of which is accepted. Therefore, in this world, one may be more likely to observe cross-domain responses to cyber attacks because offensive cyber counter responses may be difficult to rapidly leverage and the public nature of the attack is likely to put pressure on leaders to carry out an expedient response once an attack reaches a certain threshold. Importantly, this threshold is likely to be lower than it would be for a World 1 or 2 type of attack because of the increased reputational costs that are generated.[156]

In this framework, when governments use cyber operations as part of a deliberate bargaining strategy, the initiator has the first choice to decide the world in which it prefers to operate. The choice of a particular world sends a signal to the target (and, if the initiator chooses World 3, international and domestic audiences) about its resolve and perceived ability to weather

[156] Responses to cyber attacks may be diplomatic, economic, or military (be it with cyber power or the use of conventional force). As will be demonstrated in a later section of this paper, diplomatic and economic responses are the most common for World 1 attacks. In the future, there may be publically available data for attacks that occur in World 2 and 3 that will enable further analysis of thresholds and responses in this space.

www.manaraa.com

a cyber or cross domain response. The initiator must also make a risk calculus regarding the likelihood that the target chooses to respond in a different world, which potentially changes the dynamics of the game. In other world, a target may elect to respond to an attack in World 1 by maneuvering into the second or third worlds, which forces the initiating state to decide if it wants to deny its culpability. In doing so, the initiating state may deescalate the situation by fostering domestic and international dissention and, therefore, delegitimizing a potential response. However, it also runs the risk that the target may be willing to publicly provide evidence to support the latter's attribution. The initiating state may also choose to deny culpability through secret channels, but in doing so it could be signaling that its resolve and capabilities are perhaps not as strong as may have seemed at first and may suggest a desire to avoid continued escalation. In another scenario, a target may elect to respond to a cyber attack in Worlds 2 or 3 by a maneuver in World 1, which is a de-escalatory move because its response is not observable, but it also runs the risk that the initiating state (and, potentially, international and domestic audiences who cannot observe the response) may interpret the target's behavior as a signal of its diminished resolve and capabilities and invite further attacks in the future. In sum, this implies that deliberate escalation in the cyber domain is more complex than other arenas because escalatory and de-escalatory maneuvers could occur vertically up and down an escalation ladder, and also through jumping between different worlds.

The Cyber Escalation Ladder

In this section, I use Kahn's 44-rung escalation ladder as a starting point for developing an escalation ladder in cyberspace and explore the extent to which the latter and former overlap and diverge. Kahn's escalation ladder is divided into seven distinct units in which the movement from one unit to another represents a "firebreak" that signals a qualitative and significant

www.manaraa.com

escalatory step.[157] The most significant firebreak, of course, is the "deliberate nonuse of nuclear weapons."[158] While the ladder is organized in a linear, hierarchical fashion, one does not have to continue to climb up the escalation ladder rung-by-rung—one can also go down and de-escalate, or one can skip rungs.[159] A state has the escalatory advantage and can achieve desired political objectives if it enjoys escalation dominance at a given rung of the ladder (ideally, at every rung). Escalation dominance goes beyond simply having military superiority at any given rung of the escalation ladder. It also includes "the assurance, morale, commitment, resolve, internal discipline, and so on, of both the principals and their allies."[160]

Organizationally, the cyber escalation ladder is similar to Kahn's in that, within each world described above, cyber escalation may also work along a fixed structure and is hierarchically organized according to different rungs and thresholds. However, while Kahn's ladder integrated the use of nuclear weapons into the spectrum of instruments of power available to a state (including diplomacy and conventional military activity), this ladder is solely for the cyber domain—hence, there are fewer rungs. This is because, to gain theoretical and conceptual clarity regarding the nature of escalation dynamics in cyberspace, it is important to first understand how cyber operations could be arrayed in relations to one another. An avenue for

---

[157] Kahn, *On Escalation*, 39-40. The seven units are: sub-crisis maneuvering, traditional crises, intense crises, bizarre crises, exemplary central attacks, military central wars, and civilian central wars. A critical firebreak or threshold is the "no nuclear use threshold" between rungs 20 and 21 of the escalation ladder.

[158] Smoke, *War*, 14. Also see Brodie, *Escalation and the Nuclear Option*, 103. Brodie emphasizes the normative aspects of the firebreak between non-nuclear and nuclear uses of forces—the idea that nuclear weapons are psychologically in a different category—noting that this distinction "represents…an idea or conviction to be actively promoted, partly through preaching its merits to the unconverted both at home and abroad."

[159] Kahn, *On Escalation*, 40.

[160] Ibid., 23.

www.manaraa.com

further research is to integrate the cyber escalation ladder with the range of instruments available to states.

Therefore, in the cyber domain, there are theoretically three parallel escalation ladders, which are mirror images of one another but differ in the extent to which any given action is attributable. It is important to note that the relative ordering of rungs on the ladder may vary slightly depending on the world in which the interaction occurs. Like Kahn's ladder, one can skip rungs and the ladder is organized hierarchically according to the degree of harm inflicted. Therefore, attacks that are destructive are placed higher than those that are disruptive. However, there are some exceptions to the hierarchical ordering of the below ladder, which reflects the wide range of targets that cyber operations could theoretically attack. For instance, while in most cases publicly defacing websites (rung 7) creates more harm for the target because it creates reputational costs and could theoretically tie leaders' hands, in some instances beaconing inside a network (rung 6) could be perceived by the target as more harmful if the network is essential for military operations (for instance, nuclear command and control networks). Thus, unlike Kahn's escalation ladder, the specific ordering of these rungs is potentially more fluid in the cyber domain. This is due, in part, to the fact that there is no universal metric for assessing virtual damage. Thus, the degree to which a given action is perceived as escalatory may depend on what a state values, which is a more subjective measure than simply material cost. This is a point that will be further elaborated in the context of the rungs below. However, the key implication is that there is likely to be variation across states and also over time in terms of how harm is defined on variables such as regime type and the systemic norms that govern interstate interactions in cyberspace at a given point in history. Indeed, Kahn makes the point that, for instance, U.S. and

www.manaraa.com

Soviet Union escalation ladders and thresholds may differ, but does not explore the implications of this for crisis stability.[161]

Figure 2: Cyber Escalation Ladder

Rungs 1 through 3 on the cyber escalation ladder mirror Kahn's "subcrisis maneuvering" unit, which he describes as maneuverings that "manipulate, either deliberately or otherwise, the fear of escalation or eruption."[162] The first three rungs essentially begin with verbal, diplomatic instruments: states indicate that a crisis is looming; they make diplomatic/political/economic

[161] Ibid., 23, 217.

[162] Ibid., 41.

www.manaraa.com

gestures; or they issue formal declarations.[163] Though they are the beginning rungs of the cyber escalation, at least as demonstrated by the recent history of cyber conflict, crises do not necessarily have to begin at this level due to the secrecy that surrounds state behavior in the domain. Perhaps a normative implication is that crisis in cyberspace *should* start like this, because it would be more stabilizing, so that another state can fully understand that a crisis is looming, but currently this is not common practice. However, since it could occur in the future it remains an integral part of the presented cyber escalation ladder. Absent these rungs, it is difficult for a state to see that a crisis is looming. Furthermore, it is difficult to interpret behavior at higher rungs if the operation is not coupled with some other type of political signaling. Absent this coupling, the line between ill-conceived coercion and brute force may be blurred or, more generally, the intent behind a particular cyber operation may be misinterpreted. Cyber escalation in World 1 would not begin at these rungs because they enable attribution; cyber escalation in World 2 may involve the private issuing of formal diplomatic statements.

Cyber operations begin at Rung 4. Rungs 4 through 6 start to enter the dangerous space of threatening behavior that can be destabilizing, and depending on the level of attribution can function either as shows of force for the attacker or shows of vulnerability for the targeted entity. Though a traditional show of force would reveal capability, because these attacks do not cause destruction the attacker is not showing any capability other than that it has mapped key networks and, in some cases, gained access. The ability of the attacker to escalate beyond is something which a target would have to make an assessment.[164] However, unique to the cyber domain, cyber operations of this nature also reveal how vulnerable the target is and may encourage the

[163] Ibid., 41-42.

[164] These rungs are stylized. There may be technical variations that enable half steps between them. However, the general logic behind each rung holds.

www.manaraa.com

target to deescalate simply because its security was shaken by the inadequacies of its own defenses.

Rung 4 envisions the aggressive port scanning of key infrastructure of another actor. In other words, it is the technical equivalent of a pointing a finger and saying, "I see you and I know what you value." Additionally, by scanning all ports on a target, an attacker can identify ingress points that are vulnerable to the toolset they possess.[165] However, port scanning sends a signal to the recipient because it can be easily detected in real-time. It is important to note that port scanning does not indicate that the sender has the ability to gain access to or the capability to disrupt or destroy the scanned target. The scanning simply signals that the critical infrastructure can be identified across cyberspace. Rung 5 turns up the invasiveness and envisions attempts to gain access that are "noisy." In other words, the attacker is conducting penetration attempts against the target and is making little to no effort to obfuscate that fact it is doing so. Often this would look like an uptick in attacks against a target using commonly tools and techniques that can easily be detected and defended against. Rung 6 is where the attacker demonstrates that it has access to the target. Causing easily detectable beacons to be activated in the target's key networks and systems invokes a threat of persistent access and also makes the target feel vulnerable because its defenses were undermined. The target may also wonder what else the attacker might have gained and retained access to.

Between Rungs 6 and 7 rests what I coin the "Public in the know" threshold. Beyond this threshold the effect of the operation is generally detectable to the public or the plausible deniability of the *effect* of the operation is difficult to hide (even though the identity of the

[165] This paper, by design, avoids delving into extremely technical aspects of any cyber operation. However, a good reference for port scanning is Patrick Engebretson, *The Basics of Hacking and Penetration Testing: Ethical Hacking and Penetration Testing Made Easy* (Amsterdam: Elsevier, 2013), 59-60.

www.manaraa.com

attacking state could still be hidden).[166] This creates political costs for the target state, even if the initiating state can avoid attribution. Rung 7 features activity such as website defacement and hacked social media accounts. The damage delivered is temporary and recoverable, but paints the website or account owner in an unfavorable light and thus inflicts a reputational cost. Rung 8 envisions DDoS attacks against the Internet facing website of a public or private entity. These attacks can vary greatly in scale and scope. They have been highly coordinated, as was the case in Operation Ababil (discussed below), but many have been largely unsuccessful and resulted in little degradation of service. Rung 9 depicts the release of confidential information from a firm or government entity. Data breaches of this nature can vary greatly in scale and scope. They can be extreme, such as the hack against the US Office of Personnel Management that led to the loss of over 20 million records and has to potential to confound U.S. intelligence operations for a generation; or the DNC hack, which exposed, among other things, the personal emails of Hillary Clinton's campaign manager and were used as fodder by her opposition during the 2016 U.S. presidential race. They also can cause reputational damage for a firm, for instance the 2014 Sony hack led to the release of sensitive corporate communiques as well as a decision to not release the film (although Sony Pictures ended up making more than expected revenue in the wake of the North Korean hack due to greater publicity about an otherwise B-rated movie).[167]

[166] It is possible that highly tailored operations against specific non-critical private or public infrastructure may be hidden from the public, but this varies by state depending on accountability requirements. For instance, the United States requires publically traded firms to report costs associated with data breaches through the Security and Exchange Commission. The important take away is that beyond the second threshold, the attack needs to assume that the public may eventually become aware of the effects of the attack.

[167] Incidentally, though the 2014 Sony breach, which will be discussed later, was initially projected to cost Sony over $100 million it is estimated to have cost around $35 million, mostly in remediation costs, representing a loss of less than 2% of Sony's projected sales for 2014. For further information see Benjamin Dean, "Why Companies Have Little Incentive to Invest in

www.manaraa.com

Between rungs 9 and 10 is a threshold that describes the transition from disruptive to destructive attacks. Attacks of this nature move beyond nuisance and depict physical damage to the network or operating infrastructure of their target. The 2014 attack by Iran against the Sand's Casino corporation which destroyed part of the firms networked infrastructure is an example of an attack at the tenth rung. However, the significance of Rung 9 may be greater in a world in which there is public attribution for the attack. Rung 11 envisions a destructive attack against non-critical military or government networks and systems.

Beyond rung 11 lies another threshold which describes the transition into targeting critical infrastructure.[168] Critical infrastructure is defined as systems so vital to a society that their incapacitation will put pose direct risks to public safety and national security. However, how individual states define what systems fall under the rubric of critical infrastructure may vary significantly or may only be realized after the fact. For instance, in democracies, electoral

---

Cybersecurity," *The Conversation*, March 4, 2015, http://theconversation.com/why-companies-have-little-incentive-to-invest-in-cybersecurity-37570.

[168] Critical infrastructure has become a widely-used term across state actors. In the United States, the *Critical Infrastructures Protection Act of 2001* defines critical infrastructure as "…systems and assets, whether physical or virtual, so vital to the United States that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters," (Title 42 U.S. Code § 5195c, paragraph e). While common place in many Western counties, the "critical infrastructure" distinction is new to China. Article 25 of the draft 2016 Cybersecurity Law specifically addresses the new designation, "The State gives priority to the protection of basic information networks providing public communications and radio & TV transmission services, important information systems of key sectors such as energy, transportation, water resources and finance and public service fields such as power supply, water supply, gas supply, healthcare and social security, military networks, government networks used by state organs at or above the level of a city with districts, and networks and systems having large numbers of users and owned or managed by network service providers." For further reference to defining critical infrastructure see John Moteff and Paul Parfomak, "Critical Infrastructure and Key Assets: Definition and Identification," Library of Congress, (Washington DC: Congressional Research Service, 2004); Alan T. Murray and Tony H. Grubesic, "Overview of Reliability and Vulnerability in Critical Infrastructure," in *Critical Infrastructure* (Berlin: Springer, 2007), 1-8.

www.manaraa.com

systems may count as critical infrastructure, while in autocracies systems that enable the regime to monitor, surveil, and control access to information domestically (such as China's "Great Firewall") may be perceived as critical infrastructure. In the United States, President Obama declared Sony Pictures to be part of under US protection in that it was part of "first amendment" values after the revelation of the North Korean hack.[169] Furthermore, the degree to which an attack is escalatory may vary from one sector of critical infrastructure to another for reasons that are beyond the scope of this paper. For instance, attacking a key military system in the middle of a military crisis that leaves the targeted state vulnerable is arguably more escalatory than crashing an election system, especially if the outage does not occur during an election cycle. Likewise, attacking a gas line that provides heating to much of a state during the winter may be more escalatory than if the same attack occurred in the spring. In other words, the degree of harm inflicted on these key systems is what determines how escalatory the attack is, not simply that critical infrastructure was targeted.[170]

When attacks on critical infrastructure cause a loss of life, then a different threshold has been breached as the damage done is no longer limited to a virtual world—it now has generated costs in physical space beyond the costs of reputation and the replacing affected systems. Furthermore, the first time this threshold is breeched would be significant because, as already noted, to date, no one has ever died as a direct result of a cyber attack.[171] Another unique aspect

[169] Barack Obama, "Obama Says Sony Should Not Have Pulled Film Over Threats," *National Public Radio*, December 19, 2014, http://www.npr.org/sections/thetwo-way/2014/12/19/371894427/fbi-formally-accuses-north-korea-in-sony-hacking.

[170] It is important to note that states do not equally value key systems. For instance, some states do not value loss of life to the same extent. For a longer discussion of this argument see, Borghard and Lonergan, "The Logic of Coercion in Cyberspace," 461-463.

[171] In Thomas C. Reed's book *At the Abyss* he discusses that during the 1980s the United States created flawed software that intelligence had suggested the Soviet Union was trying to steal.

www.manaraa.com

of this threshold is that the attacker may not have intended to cross it. Indeed, critical infrastructure is often critical because its loss can directly impact public safety, therefore an attack against it may have unanticipated consequences, including the loss of life. Therefore, these types of attacks at the thirteenth rung are extraordinarily risky and any rational attacker should be conscientious that controlling escalation once critical infrastructure is targeted may be exceedingly difficult.

Rung 14, which lies beyond the final threshold, envisions widespread destructive attacks. These attacks would cross multiple critical sectors and not be geographically localized. Attacks of this nature cause catastrophic devastation and fear among a populace as a coordinated offensive cyber campaign is carried out. Once this rung is hit, it could very easily look like the cyber equivalent of Herman Kahn's famous "war-gasm," *initially*. However, once this "cyber-gasm" is hit, the complexity of the cyber attacks are likely to gradually subside as the tailored accesses and tools that were necessary to conduct the initial offensive campaign become inert due to responding defenses and simply because the absence of universal lethality means that cyber weapons they often lose their utility once they have achieved their desired objectives. Therefore, for the attack to continue over time, attackers are likely to pursue fast, cheap, and easy means to hit remaining vulnerable targets (although, the remaining targets are likely to be the less vulnerable than those targeted in the initial wave). This type of targeting is unlikely to be effective against critical infrastructure, which often requires custom tools and accesses.

Allegedly the software made its way back to the Soviet Union and triggered a massive pipeline explosion. Thomas C. Reed, *At the Abyss: An Insider's History of the Cold War* (New York: Presidio Press, 2005), 268-270. Though, Jason Healey, a noted expert on the history of cyber conflict, suggests that this event needs further verification. Jason Healey, ed. *A Fierce Domain: Conflict in Cyberspace, 1986 to 2012* (Vienna, VA: Cyber Conflict Studies Association, 2013), 29.

www.manaraa.com

Therefore, causing widespread attacks against vulnerable targets across all sectors could quickly become a war of attrition.

It is important to note that rungs 9 through 13 require access in order to carry out the desired effect. At least conceptually, each rung requires a greater level of capability and access than afforded at the rung below. That said, there exists an incentive for a state that does not have preexisting accesses and capabilities to skip directly to targeting the weak underbelly of the state. This "cherry picking approach" to targeting, if effective, amounts to a war of attrition. Indeed, if the scope and scale of the onslaught is wide enough, the effect could lead to the faltering of most of the state's IT infrastructure as users lose confidence in the reliability and integrity of their systems and data.

Implications: Strategies of Victory; Escalation Dominance; and Brinksmanship

If deliberate cyber escalation can ever be useful for states, they must be able to achieve political objectives and extract concessions from adversaries. This section discusses the implications of the escalation ladder for how states could construct strategies of victory; the possibility of achieving escalation dominance in cyberspace; and the dynamics and potential risks of brinksmanship. According to Kahn, there are two strategies one can employ to achieve victory in the context of deliberate escalation. The first type of strategy involves "mak[ing] use of features of the particular 'agreed battle' that is being waged in order to gain an advantage."[172] In other words, one side in a crisis may have a threshold or limit beyond which it will not escalate, and will instead focus on achieving victory within a predetermined sphere. These thresholds could be clearly and credibly communicated to the adversary, or they could be kept secret. The second type of strategy eschews limits to escalation, instead engaging in

[172] Kahn, *On Escalation*, 7.

www.manaraa.com

brinksmanship/games of chicken.[173] Essentially, this is the notion that one party to a crisis can escalate and risk "eruption" to force the other side to concede (Schelling also discusses this).[174]

Kahn's "agreed battle" presents issues for the cyber domain when thresholds are not clearly communicated and it may be possible to inadvertently escalate further than intended due to opaque communication and the possibility of cascading effects. Furthermore, the lack of overtly and clearly defined thresholds fundamentally undermines the creation of an agreed battle space. If there are explicit thresholds, states have certainly kept them secret, which Kahn suggests "may run the risk of a full-scale pre-emptive eruption."[175] Furthermore, in a crisis where trust may also be an issue between adversaries, the only way to currently limit the conflict space would be through some sort of diplomatic agreement. However, the likelihood that one side would use the ambiguity of the domain to gain an advantage is high, thus undermining incentives for agreement.

Intrinsically linked to escalation is the concept of brinkmanship. Schelling conceptualizes brinkmanship as a game of nuclear chicken; it involves deliberately manipulating the mutual risk of nuclear-armed states in a crisis such that things may get beyond an actor's control and both parties could fall over the brink.[176] Employing a strategy of brinkmanship is possible in the context of cyber operations. Parties to a cyber crisis could manipulate the risk of escalation through tying their hands using two mechanisms. First, states may choose to employ cyber proxies and essentially arm them with capabilities and thus turn over command and control. This

[173] Ibid., 7.

[174] Ibid. 7-9.

[175] Ibid., 7.

[176] Thomas C. Schelling, *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1960, Chapter. 8.

www.manaraa.com

ties a state's hands because, if the attack against another state lacks centralized command, it could generate escalatory effects that may be hard to control. For example, when Estonia came under a far-sweeping cyber attack following the removal of a Soviet-era war monument, over 128 botnet control nodes were registered.[177] This meant that the actor (most likely Russia) behind the attacks had delegated authority to proxy actors that they may not have had influence over. Secondly, states could tie their hands in the cyber domain through automaticity, which could take several forms. First, states could create the capabilities for automatic counter cyber attacks. In other words, an autonomous offensive response to an incursion that is made without human decision-making. This would involve significant technical complexities, such as the ability to ensure the hacking back capability functions the way in which it is intended against an unknown target, and the need to develop artificial intelligence to ensure that it does not get manipulated by the attacker. However, DARPA's 2015 Cyber Grand Challenge demonstrated that super computers can indeed be used to find vulnerabilities and attack another computer while themselves under attack.[178] Another way to create automaticity in cyberspace would be to employ honey pots, which are a type of cyber security mechanism designed to attract attackers to certain types of data resident on a network. If these are deployed in such a way that the data that the attacker exfiltrates contains a logic bomb that is activated when it is back on the host's own network, then an automatic counter response is possible. In this sense, the widespread employment of honey pots is analogous to a minefield that, if properly arrayed, only hurts the attacker.

---

[177] Andreas Schmidt, "The Estonian Cyberattacks," in *A Fierce Domain*, ed. Jason Healey, 182.

[178] Kelsey Atherton, "DARPA's Cyber Grand Challenge Ends in Triumph," *Popular Science*, August 5, 2015, http://www.popsci.com/machines-win-darpas-cyber-grand-challenge; New York Cyber Task Force, "Leverage: Getting to a Defensible Cyberspace," co-chairs: Merit Janow, Gregory Rattray, and Phil Venables, (forthcoming), 27.

www.manaraa.com

Beyond successfully employing a brinksmanship strategy, states could theoretically win an escalation contest if they enjoy escalation dominance, ideally at every rung of the escalation ladder. If one state can defeat another at every rung of the escalation ladder, it can control the dynamics of a crisis either through deterring the other side from escalating or (paradoxically) escalating to a rung in which it has escalation dominance in order to de-escalate. While the concept of escalation dominance is difficult to operationalize in traditional circumstances, it is even more complex in the cyber domain. There have been some claims in the cyber literature that escalation dominance is possible, but these arguments ignore the implications of the lack of universal lethality of cyber weapons and fluid nature of access in the domain, upon which delivering an effect is predicated.[179] Together, these complicate states' efforts to hit the right, rather than the most available or most vulnerable, target. Thus making achieving escalation dominance nearly impossible. Moreover, this effort would have to be replicated for every potential adversarial state as accesses and required capabilities are unique for each target.

Furthermore, even if escalation dominance were physically possible in cyberspace, there are other aspects of the domain that complicate it in practice. For instance, one shortcoming of Kahn's work is that he does not discuss the implications of different conceptions of the escalation ladder, as well as poor understandings of the other side's ladder. A theoretical implication of this for achieving victory through escalation dominance is that, if thresholds are not codified in doctrine, a state may inadvertently escalate more than intended—perhaps to a rung in which it lacks an advantage over its adversary. Furthermore, the fact that these thresholds are relatively new and have mostly remain untested suggests that, even if they are made known,

[179] See, for example, Stephen J. Cimbala, "Cyber War and Deterrence Stability: Post-START Nuclear Arms Control," *Comparative Strategy* 33, no. 3 (2014): 279-286. Also see Cavaiola et al., "Cyber House Rules," 84, 99.

www.manaraa.com

it is likely that they will gradually tested. Indeed, Schelling's famous analogy of wading into the water to discover the true threshold applies in this space.[180] Furthermore, beyond ambiguity regarding thresholds, the secrecy surrounding cyber operations and the potential involvement of third party proxies impedes states' calculations regarding its adversary's capabilities at any given rung of the escalation ladder, making it difficult to assess which side may have escalation dominance.

A final key facet of deliberate escalation is the ability of the escalation state to control the magnitude of the escalation. Kahn, for example, describes three ways an actor could escalate in the context of a crisis: increase the intensity of its effort (by doing more of the same, or by violating targeting and weapons norms and, in extreme cases, using nuclear weapons); widen the area of operations (by increasing the geographic scope of the conflict to include local sanctuaries or neighboring countries); or engage in compound escalation (by targeting beyond the local area of operations to hit the principal adversary or her allies).[181] While it is possible to contain a cyber attack to a select target or group of targets that are all located within a certain region or state, a state needs to anticipate that unforeseen outcomes that extend beyond the area of hostilities. Cyber operations can have cascading effects due to unanticipated behavior as well as command and control issues that may result in targeting a principal or an ally unintentionally. Additionally, limiting cyber operations to certain types of targets may be complicated due to the nature of targeting over time. As already discussed above, over the course of a cyber campaign, a state's access, exploits, and their utility expire. Therefore, states may find that increasing intensity

[180] Schelling, *Arms and Influence*, 66-69.

[181] Kahn, *On Escalation*, 4-6.

www.manaraa.com

means that they are targeting the most vulnerable targets because they present the easiest means of access, which could mean targeting civilian assets.

The previous discussion is not to insinuate that escalation can never be controlled and deliberate in cyberspace. Cyber escalation may be controlled through the implementation of several restrictive measures. First, care would need to be made during development and testing of cyber weapons to prevent undesired cascading effects.[182] Second, the operators conducting the attack would need to be closely monitored by the state to ensure they do not exceed their orders. Third, states can develop confidence building measures to mitigate the risks of misinterpretation of signals and intent in the cyber domain.[183]

De-escalation and Ending Wars

The flip side of escalation dynamics is, of course, de-escalation. There are several aspects of cyber operations that facilitate de-escalation more so than in traditional domains of conflict, but also attributes of cyber conflict that make de-escalation difficult. One way that Kahn posits states can de-escalate from a crisis is through developing capabilities for "de-escalation dominance," in which it can unilaterally de-escalate from a situation regardless of what the adversary is doing.[184] Kahn lists some examples of de-escalation: reversing the previous escalation move; resolving a peripheral dispute; free prisoners; make conciliatory statements;

[182] However, despite testing, cyber weapons may act in highly unanticipated ways once they enter the target's networked environment. For instance, Stuxnet (which was presumably a highly vetted cyber capability that was only intended to in infect Iranian nuclear enrichment processes) is reported to have infected over 100,000 computers worldwide.

[183] For further discussion of how confidence building measures can be used to signal intent of cyber operations see Shawn W. Lonergan, "Arms Control and Confidence Building Measures for the Cyber Domain," (working paper).

[184] Kahn, *On Escalation*, 231.

www.manaraa.com

replacing individuals; let time pass for the situation to dampen down.[185] The primary mechanism through which cyber operations facilitate de-escalation is the aforementioned dampening effects of operating in the domain. Because there is no universal lethality of cyber weapons, states will eventually exhaust their cyber capabilities or find them rendered inert over time as effective defenses are discovered. States may try to fight this tendency through mobilizing new resources but, even then, there is likely to be a significant time lag that injects breathing space into any crisis between cyber adversaries. This implies that tit-for-tat volleys in cyberspace are likely to be separated by longer stretches of time than in other domains, providing breathing space for governments to decide to ramp down.[186] Furthermore, employing disruptive, rather than destructive, cyber attacks may also facilitate de-escalation because they can be walked back in a pre-coordinated timeline without having irrevocably destroyed an adversary's capabilities or assets.

However, there are also unique aspects of interstate relations in the cyber domain that make some aspects of de-escalation more difficult. Kahn states that parties can de-escalate through "concessions and conciliation," but that de-escalation may be more difficult to achieve because it "is even more sensitive to accurate communication and shared understandings than escalation is."[187] This is because, to de-escalate, adversaries must be able to coordinate on de-escalation moves, and this requires that they have a shared understanding of the situation. In cyberspace, this shared understanding of a situation may be hard to credibly communicate because states have incentives to keep information secret, as revealing it may expose accesses

[185] Ibid., 232.

[186] Libicki, *Cyberspace in Peace and War,* 276-277, 285.

[187] Kahn, *On Escalation*, 230-231.

www.manaraa.com

they use for intelligence collection. Furthermore, even if states could agree on a shared understanding of the situation, the ability to credibly commit to a course of action and cease hostilities is problematic. Indeed, depending on the means of attack, it may be nearly impossible to reassure one's adversary that the attacks will subside. For instance, if a state has employed proxies to conduct the attack, the state would have to find a way to credible exercise control over them to crease the onslaught. This can be difficult if the proxies are decentralized and the state does not have any means to influence them once they have equipped them with tools or targeting information. Furthermore, certain methods of attack can have effects that are difficult to stop either because of the way they are designed to propagate through a network, or once exposed to the target environment they act in unforeseen ways. For example, Stuxnet, which targeted Iranian nuclear centrifuges, it is believed was never supposed to propagate beyond a few targeted systems. However, due to its unique design, the worm quickly spread to other SCADA systems throughout the world.[188] One potential method to de-escalate in such a situation is to share the original source code of the weapon with the party with whom you want to de-escalate. This would promote a shared understanding of the situation and for the targeted state to more rapidly engineer a means to thwart the attack and patch vulnerabilities as opposed to having to wait for an in-depth forensic analysis.

Finally, in the context of ending wars, Kahn briefly explores the implications of the mobilizations of society for war. The First and Second World Wars required the total mobilization of the belligerents' populations to wage modern industrial warfare. Kahn suggests that thermonuclear war, despite being far more destructive, would not be total in terms of

[188] David E. Sanger, "Obama Ordered Wave of Cyberattacks Against Iran," *The New York Times*, June 1, 2012, http://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html. For a discussion of the difficulties of reassurance in cyberspace, see Borghard and Lonergan, "The Logic of Coercion in Cyberspace," 471-472.

www.manaraa.com

mobilizing society; rather, it could be waged by government authorities and technicians.[189] The implication is, therefore, that thermonuclear war would be easier to end because governments do not have to walk back their entire population from supporting the war effort. Would strategic cyber warfare require the mobilization of civilian society and, if so, what would it look like?

Due to the multi-stakeholder nature of the Internet, cyber power does not solely exist in government. Government networks ride on the backbone of civilian ones; non-governmental actors are powerful players in the domain and may have capabilities that rival even those of states; cyber talent is diffuse and exists across many sectors of society, not just in the military industrial complex; and in many ways the private sector is ahead of government agencies in understanding and operating in the domain. Therefore, in this sense, strategic cyber war may necessitate the mobilization of certain sectors of society in a way that thermonuclear war would not. Mobilization could come through two means: either the mobilization of the technology base to engineer cyber weaponry and supporting infrastructure, or the equivalent of an "electronic levee en masse." Though mobilizing the tech sector in the same way the industrial sector did during WWI and WWII is possible, it would fundamentally be different given the no universality lethality of cyber weapons principle. To be effective, states would need to entrust corporations with targeting information, intelligence, and battle plans well in advance of their implementation for the base to develop requisite capabilities. Though there are means to clear these civilians for access to information that states typically classify, the potential to lose the element of surprise behind future military operations would be high. Furthermore, it would be difficult to scale such a program across all of industry. A more likely scenario would be that states would engage with elements of their defense industrial base that have been pre-vetted to handle sensitive

[189] Kahn, *On Escalation*, 201.

www.manaraa.com

information. States may also consider conscripting these personnel to conduct operations, however their legal status and the oversight of their operations are considerations that would a prudent state would want to address. However, if a state decides to engage in indiscriminate targeting such as that described at rung 14, the cyber-gasm, then this type of mobilization might be exceedingly effective.

A second option envisions combining the computing power of computers resident in a state. This "electronic levee in masse" would require the placement of software on some percentage, if not all, of the computers in a state that enables it to direct unused computing cycles to assist in conducting attacks or breaking advanced encryption standards.[190] The computational power that distributed computing on this level would vary from state to state depending on the scale and scope of the program it implements and the preexisting computing base within the state. In a future where this method is employed, electronic coalition warfare that brings together the combined computing power of multiple states may be a reality. Taken together, this suggests that, in the event of strategic cyber warfare, the relationship between the civilian sector and government may more closely resemble that of the total industrial wars of the twentieth century than a theoretical thermonuclear war, making it more difficult for government to end these kinds of wars.

[190] Though it would arguably be a violation of the constitutions of numerous western states to require an "Electronic Levee en Masse," as a proof of concept it has already occurred. In 1999 the Search of Extraterrestrial Intelligence (SETI) organization created the "SETI@Home" effort which asked computer users to download free software that would run when the user is not using their computer. This enabled the creation of a virtual super computer comprised of Internet-connect computers spanning the globe that can assist in SETI data analysis efforts. Currently, the project represents the world's largest distributed computing platform with over 3 million users. Further information about the project can be found at- http://setiathome.berkeley.edu/

www.manaraa.com

The Dogs That Didn't Bark

Below, I explore four examples of U.S. interactions with competitors in cyberspace: China, North Korea, Iran, and Russia. These examples demonstrate this section's empirical puzzle and provided the thrust for the presented theoretical analysis, with the caveat discussed earlier that scholars and practitioners who take the position that the cyber domain is inherently escalatory may interpret these cases differently due to divergent conceptions of what counts as escalatory (i.e., defining an equivalent response espionage as escalatory behaviors). These cases only address the response action from the United States to a specific cyber attack or type of activity.[191] One shortcoming of this approach is that the individual responses may be part of a larger tit-for-tat strategic response. However, given the secretive nature of cyber operations and the closed nature of the authoritarian governments evaluated, data on their perspective is speculative unless otherwise noted.

The first case considered, China, stands apart from the other three. With China, the United States was responding to a systematic cyber espionage campaign, whereas with the other three examples the United States responded to distinct events that generated effects beyond the theft of trade secrets. In all three of these cases the public was aware of the specific occurrences, as opposed to the case with China where the general US public was aware of systemic Chinese espionage, but perhaps not individual thefts against public and private entities.

China

Speaking before Congress in 2016, former director of the National Security Agency and Commander of the United States Cyber Command, General-retired Keith Alexander described

[191] However, it is possible that, in some of these cases, the non-U.S. state was responding to a U.S. action that is not yet publicly known.

www.manaraa.com

how cyber space has enabled the greatest transfer of wealth in history.[192] Though Alexander was speaking about the cost of the loss of trade secrets from cyber espionage, it is well documented that that the Chinese government has engaged in a systemic campaign of cyber economic espionage against the United States. Many of the attacks against corporate and governmental targets remain unreported to the public, but several incidents have recently come to light. The United States, however, has historically been reticent to publically address Chinese espionage. Indeed, despite China being linked to cyber espionage from the early 1990s the first time the US publically named the Chinese government as being behind the cyber-enabled theft of intellectual property was in a 2011 National Counter Intelligence Executive report that noted that "Chinese actors are the world's most active and persistent perpetrators of economic espionage."[193] This was a significant step because, prior to this report, the United States had been reticent to publically chastise state actors for economic cyber espionage.[194] Other than this report however, there was no other significant public discussion of Chinese cyber espionage by the US government for over 2 years, despite China's continuation of the practice. One senior US military officer with knowledge of US-Chinese exchanges explained that the Chinese would simply deny taking part in cyber espionage and would ask the US to prove their claims. In his

[192] Keith B. Alexander, "Prepared Statement of GEN (Ret) Keith B. Alexander on Digital Acts of War: Evolving the Cybersecurity Conversation before the Subcommittees on Information Technology and National Security of the Committee on Oversight and Government Reform," U.S. House Oversight Committee, July 13, 2016, https://oversight.house.gov/wp-content/uploads/2016/07/Gen-Alexander-Statement-Digital-Acts-of-War-7-13.pdf.

[193] "Foreign Spies Stealing US Economic Secrets in Cyberspace," Office of the National Counter Intelligence Executive, October 2011, i.

[194] Thom Shanker, "U.S. Report Accuses China and Russia of Internet Spying," *The New York Times*, November 3, 2011, http://www.nytimes.com/2011/11/04/world/us-report-accuses-china-and-russia-of-internet-spying.html.

www.manaraa.com

estimation, the Chinese knew full well that the United States did not want to do this because it would expose the intelligence accesses the United States relied upon to spy on China.[195]

This public ambivalence ceased in February 2013 when *The New York Times* published a widely read article that directly named the Chinese People's Liberation Army (PLA) as one of the sources of a series of attacks targeting private corporations in the United States.[196] The article was based off of a seminal investigatory report issued by the US cyber security firm Manidant.[197] Though the Chinese government denied their involvement the day following the release of the report, given the technical nature of the Chinese Internet infrastructure, which is known for widespread government monitoring and filtering, and the detail of the evidence that Mandiant had collected and made publically, it is implausible that the Chinese Government was not aware of, if not sponsoring, these activities.[198] Following the release of this report in March 2013, Tom Donilon, then U.S. National Security Advisor, addressed the Asia Society in New York City. During his remarks, Mr. Donilon issued the first public admonishment by a senior U.S. official of Chinese cyber activities against U.S. corporations and national interests.[199] It was

---

[195] Interview with a senior US military officer with extensive operational cyber experience, September 2014.

[196] David E. Sanger, David Barboza and Nicole Perlroth, "China's Army Is Seen as Tied to Hacking Against U.S.," *The New York Times*, February 18, 2013, http://www.nytimes.com/2013/02/19/technology/chinas-army-is-seen-as-tied-to-hacking-against-us.html.

[197] "APT1- Exposing One of China's Cyber Espionage Units," *Mandiant*, February 2013, https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf.

[198] David Barboza, "China Says Army Is Not Behind Attacks in Report," *The New York Times*, February 20, 2013, http://www.nytimes.com/2013/02/21/business/global/china-says-army-not-behind-attacks-in-report.html.

[199] "Complete Transcript: Thomas Donilon at Asia Society New York," Asia Society, March 11, 2013, http://asiasociety.org/new-york/complete-transcript-thomas-donilon-asia-society-new-york.

www.manaraa.com

subsequently revealed to the press that, three months prior to Donilon's speech, the U.S. had issued a secret demarche order to the Chinese government in protest of cyber espionage on the heels of over 6 months of unproductive closed door dialogues between the two governments.[200] Though it is unclear what pressed the Obama administration to finally admonish China, it probably occurred, in part, as a result of outcry from elements of the US private sector that had fallen victim to Chinese hacking. This, coupled with the public release of the Mandiant report, enabled the US to highlight evidence of Chinese hacking without endangering any intelligence accesses. Indeed, the Mandiant report appears to be the tipping point for the administration to act. From the U.S. perspective, economic and industrial espionage are considered illegitimate (unlike intelligence collection to fulfill national security requirements). Through this public admonishment, the United States was attempting to employ a "naming and shaming" strategy to create a norm against economic espionage between the United States and China.

Xi Jinping becoming president of the People's Republic of China in March 2013 created an opportunity for the US to advance this foreign policy initiative through diplomacy. Indeed, the issue of collaborating on the protection of intellectual property and cyber security threats was discussed during President Obama's first call to President Xi.[201] This point was reiterated during the president's first meeting in June where cyber security and the protection of intellectual property was a significant point of discussion. During a press conference at the summit, President Xi was asked directly about Chinese cyber attacks against the United States, Xi noted that "The application of new technology is a double-edged sword. On the one hand, it will

[200] Siobhan Gorman, "U.S. Eyes Pushback On China Hacking," *The Wall Street Journal*, April 22, 2013, http://www.wsj.com/articles/SB10001424127887324345804578424741315433114.

[201] "Readout of the President's Phone Call with Chinese President Xi Jinping," The White House, March 14, 2013, https://obamawhitehouse.archives.gov/the-press-office/2013/03/14/readout-president-s-phone-call-chinese-president-xi-jinping.

www.manaraa.com

drive progress in ensuring better material and cultural life for the people. On the other hand, it might create some problems for regulators and it might infringe upon the rights of states, enterprises, societies and individuals."[202] Xi's remarks reflect the fundamental way that China conceived cyber espionage—as a means to further the power of the state by enabling innovation and growth across all economic sectors.[203] However, both leaders realized that there was a need to cooperate on cyber security related threats and agreed to open a dialogue moving forward. The next month, in July 2013, a bilateral working group for cyber security held its first meeting where the issue of cyber espionage was discussed and the need to collaborate via future meetings to address these concerns.[204]

Despite the public admonishments and top-level meetings, public reporting from the Office of the National Counter Intelligence Executive indicated that China continued to aggressively target U.S. interests and demonstrated an increasing level of sophistication.[205] The next public rebuke from the United States of Chinese cyber espionage came in May 2014 with the indictment of 5 PLA officers for many of the activities identified in the 2013 Mandiant

---

[202] "Remarks by President Obama and President Xi Jinping of the People's Republic of China After Bilateral Meeting," The White House, June 8, 2013, https://obamawhitehouse.archives.gov/the-press-office/2013/06/08/remarks-president-obama-and-president-xi-jinping-peoples-republic-china-.

[203] China's use of cyber espionage as a deliberate means to secure economic growth and industrialization is well documented. For instance, see Segal, *The Hacked World Order*, 125-129.

[204] "Senior Administration Officials On the First Day of the Strategic and Economic Dialogue and U.S.-China Relations," U.S. State Department, July 11, 2013, https://2009-2017.state.gov/r/pa/prs/ps/2013/07/211801.htm.

[205] J.J. Green, "Exclusive: NCIX says Chinese hackers 'getting faster and smarter,'" *Wtop*, July 10, 2014, http://wtop.com/j-j-green-national/2014/07/exclusive-ncix-says-chinese-hackers-getting-faster-and-smarter/.

www.manaraa.com

report.[206] The Chinese responded to the indictments by again denying their involvement and by pulling out of the bilateral cybersecurity working group in protest.[207]

In mid-2015, Chinese hacking took a new twist when they breached the personnel records of over 21 million U.S. government employees including the finger prints of nearly 6 million.[208] While this breach was significant due to its scale and the impact it could have on ongoing and future intelligence operations, reporting indicates that the Obama administration was perplexed with how to respond for they had been pressing a norm that intelligence for national security was acceptable. Indeed, James Clapper, the Director of National Intelligence at the time, noted that "…you have to kind of salute the Chinese for what they did."[209] Several months following the U.S. realization of the OPM breach, Presidents Obama and Xi struck a deal to stop economic

---

[206] United States District Court Western District of Pennsylvania, "United States of America v. Wang Dong, Sun Kailiand, Wen Xinyu, Huang Zhenyu, Gu Chunhui," Criminal Number: 14-118, Filed: May 2014, https://www.justice.gov/iso/opa/resources/512201451913235846194 9.pdf. But also, see the official press release- Department of Justice, "U.S. Charges Five Chinese Military Hackers for Cyber Espionage Against U.S. Corporations and a Labor Organization for Commercial Advantage," Office of Public Affairs, May 19, 2014, https://www.justice.gov/opa/pr/us-charges-five-chinese-military-hackers-cyber-espionage-against-us-corporations-and-labor.

[207] Shannon Tiezzi, "China's Response to the US Cyber Espionage Charges," *The Diplomat*, May 21, 2014, http://thediplomat.com/2014/05/chinas-response-to-the-us-cyber-espionage-charges/.

[208] Julie Hirschfeld Davis, "Hacking of Government Computers Exposed 21.5 Million People," *The New York Times*, July 9, 2015, https://www.nytimes.com/2015/07/10/us/office-of-personnel-management-hackers-got-data-of-millions.html. David E. Sanger, "Hackers Took Fingerprints of 5.6 Million U.S. Workers, Government Says," *The New York Times*, September 23, 2015, https://www.nytimes.com/2015/09/24/world/asia/hackers-took-fingerprints-of-5-6-million-us-workers-government-says.html.

[209] David E. Sanger, "U.S. Decides to Retaliate Against China's Hacking," *The New York Times*, July 21, 2015, https://www.nytimes.com/2015/08/01/world/asia/us-decides-to-retaliate-against-chinas-hacking.html.

www.manaraa.com

cyber espionage in September of 2015 and discussed means to create a cyber code of conduct.[210] The further creation of cyber norms between the two states has continued through the reinstated bilateral cyber security working group which meets bi-annually. Both the private sector as well as public statements from Obama's US Assistant Attorney for National Security indicated that Chinese cyber espionage decreased in the wake of the accord.[211] Through concerted diplomatic engagement and strategically timed public admonishments, the United States was able to create a norm against economic espionage with China. However, only time will tell the long-term durability of the norm.

North Korea

Sony Pictures Entertainment's development of movie *The Interview,* depicting two American journalists that are conscripted by the Central Intelligence Agency to assassinate North Korea's Kim Jong Un triggered one of the biggest cyber security breaches of 2014. The infamous "Sony hack" was part of a failed coercive ploy to prevent the release of the film that played out over November and December of that year. The first indication that something was

[210] Julie Hirschfeld Davis and David E. Sanger, "Obama and Xi Jinping of China Agree to Steps on Cybertheft," *The New York Times*, September 25, 2015, https://www.nytimes.com/2015/09/26/world/asia/xi-jinping-white-house.html. David E. Sanger, "U.S. and China Seeks Arms Deal for Cyberspace," *The New York Times*, September 19, 2015, https://www.nytimes.com/2015/09/20/world/asia/us-and-china-seek-arms-deal-for-cyberspace.html.

[211] John P. Carlin, "Detect, Disrupt, Deter: A Whole-of-Government Approach to National Security Cyber Threats," (moderated discussion, Center for Strategic & International Studies, June 28, 2016), https://www.csis.org/events/detect-disrupt-deter-whole-government-approach-national-security-cyber-threats. Also, see the report by the cyber security firm FireEye- FireEye, "Redline Drawn: China Recalculated its Use of Cyber Espionage," *FireEye Isight Intelligence*, June 2016, https://www.fireeye.com/content/dam/fireeye-www/current-threats/pdfs/rpt-china-espionage.pdf. Of note, one finding of the report was, "… a notable decline in China-based groups' overall intrusion activity against entities in the U.S. and 25 other countries. We suspect that this shift in operations reflects the influence of ongoing military reforms, widespread exposure of Chinese cyber operations, and actions taken by the U.S. government."

www.manaraa.com

awry came on Friday, November 21st, when several top executives received a cryptic email from a group calling themselves "God'sApstls" demanding monetary compensation or else "Sony would be bombarded as a whole."[212] However, these emails were either filtered by spam filters or ignored entirely by the recipients.[213] Additionally, one of the company's Twitter feeds was taken over and a message stating that the two co-chairs of Sony were "going to hell" was sent out on the same day. The following Monday many of the 3,500 Sony employees came into their Culver City, California corporate headquarters to find that they were locked out of their systems and some were left with a virtual banner depicting a skull with a message stating they had been hacked by the "Guardians of Peace" and specifically that,

> We've obtained all your Internal data including your secrets and top secrets. If you don't obey us, we'll release data shown below to the world. Determine what will you do till November the 24th, 11:00PM (GMT).[214]

The message also contained links to sites where the hackers had already posted some of Sony's files.

On Tuesday, November 25th the first significant data dump came with the posting of four yet-to-be released films online. Shortly after their posting, a senior editor working at *fusion.net* received an email from the "boss of the G.O.P" instructing him to search online for the leaked films and that there was nearly 100 terabytes of Sony's data that was yet to be leaked.[215] Over the next several weeks, the world saw at least eight separate dumps of Sony's corporate data.

---

[212] Mark Seal, "An Exclusive Look at Sony's Hacking Saga," *Vanity Fair*, March 2015, http://www.vanityfair.com/hollywood/2015/02/sony-hacking-seth-rogen-evan-goldberg.

[213] Ibid.

[214] Kim Zetter, "Sony Got Hacked Hard: What We Know and Don't Know So Far," *Wired*, December 3, 2014, https://www.wired.com/2014/12/sony-hack-what-we-know/.

[215] Seal, "An Exclusive Look at Sony's Hacking Saga."

www.manaraa.com

Among these data dumps were salary schedules, human resource files, network information, and credentials that could be used by others to gain further access to Sony's IT assets. Indeed, hackers even included directions on how these data could be used to further exploit Sony.[216] Additionally, thousands of internal emails were also leaked. In the wake of the breach, these leaked emails proved later to contribute to the dismissal of at least one executive.[217] Forensics of the malware employed found that the malware was coded to search for specific servers and either send the information to a server designated by its originator or to simply erase all data on the infected system.

After the third data dump the attack got personal when several Sony employees received emails to the either their private or corporate emails accounts asking them to sign a petition condemning Sony or else they and their families would be in danger.[218] However, it was only after the 7th data dump, on December 16th that the hacker's demands were made clear. In an email to Sony's chief financial officer, the hackers insinuated that "9/11" violence was coming and that employees should seek shelter when the world sees "…what an awful movie Sony Pictures Entertainment has made."[219] To Sony, the message was clear—they were asked to cancel the release of *The Interview*. The coercive tactics worked, causing the studio to cancel their planned Christmas Day 2014 release following the backing out of several major theater

---

[216] Zetter, "Sony Got Hacked Hard: What We Know and Don't Know So Far."

[217] "Ex-Sony Chief Amy Pascal Acknowledges She Was Fired," *NBC News*, February 12, 2015, http://www.nbcnews.com/storyline/sony-hack/ex-sony-chief-amy-pascal-acknowledges-she-was-fired-n305281.

[218] Seal, "An Exclusive Look at Sony's Hacking Saga."

[219] Ibid.

www.manaraa.com

chains.[220] This triggered a public outcry in the United States as it appeared the hackers had been successful in stymieing the release of the movie. Indeed, President Obama stated that he respected Sony's decision, but felt that they "…made a mistake" and "we cannot have a society in which some dictator someplace can start imposing censorship here in the United States."[221] This response from President Obama encouraged Sony executives to allow independent movie theaters the option of airing the movie as well as to work with several online distributers—the film had its Christmas day release after all.[222]

The same day that President Obama expressed his concerns about the Sony hack, the FBI released its report attributing the attack to the North Korean government.[223] The report reflected an unusually high degree of confidence in the attribution, which led some experts to suggest that US must have had preexisting access to the networks from which the attack emanated.[224] This attribution made it possible for the United States to publically respond. On January 2, the Obama administration responded with an executive order imposing economic sanctions against ten senior North Korea officials and the intelligence organization linked to North Korean cyber

---

[220] Michael Cieply and Brooks Barnes, "Sony Cyberattack, First a Nuisance, Swiftly Grew into a Firestorm," *The New York Times*, December 30, 2014, https://www.nytimes.com/2014/12/31/business/media/sony-attack-first-a-nuisance-swiftly-grew-into-a-firestorm-.html.

[221] Barack Obama, "Remarks by the President in Year-End Press Conference, The White House, December 19, 2014, https://obamawhitehouse.archives.gov/the-press-office/2014/12/19/remarks-president-year-end-press-conference.

[222] Seal, "An Exclusive Look at Sony's Hacking Saga."

[223] "Update on Sony Investigation," Federal Bureau of Investigation, December 19, 2014, https://www.fbi.gov/news/pressrel/press-releases/update-on-sony-investigation.

[224] Sanger and Fackler, "N.S.A Breached North Korean Networks Before Sony Attack, Officials Say." Also see, Segal, *The Hacked World Order*, 51-60, for an accounting of the events surrounding the Sony hack and the resulting attribution.

www.manaraa.com

operations.[225] Additionally, several sources reported that there were unspecified covert actions taken as well.[226]

Iran

From September 2012 to March 2013 a group calling themselves the Izz ad-Din Al-Qassam Cyber Fighters, conducted a multi-phase cyber attack.[227] Operation Ababil, as the hackers called it in an online blog post, directed sophisticated distributed denial of service attacks against dozens of the public-facing websites of U.S. financial institutions.[228] According

---

[225] "Executive Order -- Imposing Additional Sanctions with Respect to North Korea," The White House, January 2, 2015,

https://obamawhitehouse.archives.gov/the-press-office/2015/01/02/executive-order-imposing-additional-sanctions-respect-north-korea.

[226] David E. Sanger and Michael S. Schmidt, "More Sanctions on North Korea After Sony Case," *The New York Times*, January 2, 2016, https://www.nytimes.com/2015/01/03/us/in-response-to-sony-attack-us-levies-sanctions-on-10-north-koreans.html. Though the United States has never reported what their "covert" operations were, shortly after Christmas and lasting into January 2016, North Korea's already limited access to the Internet was entirely cut off. North Korea quickly blamed the United States as the culprit and leveraged racist insults against President Obama. For more information see, Martin Fackler, "North Korea Accuses U.S. of Staging Internet Failure," *The New York Times*, December 27, 2014, https://www.nytimes.com/2014/12/28/world/asia/north-korea-sony-hacking-the-interview.html

[227] The Department of Justice indictment of the Iranian Revolutionary Guard Members associated with this attack states that the DDoS attacks started in approximately December 2011, however, it was in September 2012 that they experienced a significant escalation which was in line with the blog posts by the Izz as-Din Al-Qassam Cyber Fighters and the commencement of Operation Ababil. For further information see United States District Court Southern District Of New York, "United States of America v. Ahmad Fathi, Hamid Firoozi, Amin Shokohi, Sadegh Ahmadzadegan, Omid Ghaffarinia, Sina Keissar, and Nader Saedi," 16 Crim 48, https://www.justice.gov/usao-sdny/file/835061/download.

[228] Nicole Perlroth, "U.S. Banks Again Hit by Wave of Cyberattacks," *The New York Times*, January 4, 2013, https://bits.blogs.nytimes.com/2013/01/04/u-s-banks-again-hit-by-wave-of-cyberattacks/. Nicole Perlroth and David E. Sanger, "Cyberattacks Seem Meant to Destroy, Not Just Disrupt," *The New York Times*, March 28, 2013, http://www.nytimes.com/2013/03/29/technology/corporate-cyberattackers-possibly-state-backed-now-seek-to-destroy-data.html.

www.manaraa.com

to the initial Department of Justice (DoJ) indictment, the attacks began in December 2011 and occurred on an occasional basis until September 2012. It was during that time that their frequency came almost weekly and increased in their intensity. DoJ estimates that at least 46 major US financial institutions and other financial-sector corporations were targets of DDoS attacks over a period of at least 176 days. They note that, "[o]n certain days during these attacks, hundreds of thousands of customers were unable to access their banks accounts online" and that "[a]s a result of these attacks, those victim institutions incurred tens of millions of dollars in remediation costs as they worked to mitigate and neutralize the attacks on their computer servers."[229] Given the technical complexity of the attack and other intelligence, the U.S. assessed Iran was the culprit and that it was the probable response to U.S. led sanctions and cyber attacks against Iran's nuclear enrichment processes.[230]

There is no open-source reporting discussing if the U.S. conducted a retaliatory military response against Iran for Operation Ababil, however the United States did respond three years later by indicting seven individuals linked to the Iranian Revolutionary Guard Corps for their involvement in the attacks.[231] It is possible that the United States also responded with a covert

---

[229] United States District Court Southern District Of New York, "United States of America v. Ahmad Fathi, Hamid Firoozi, Amin Shokohi, Sadegh Ahmadzadegan, Omid Ghaffarinia, Sina Keissar, and Nader Saedi," 4.

[230] Nicole Perlroth and Quentin Hardy, "Bank Hacking Was the Work of Iranians, Officials Say," *The New York Times*, January 8, 2013, http://www.nytimes.com/2013/01/09/technology/online-banking-attacks-were-work-of-iran-us-officials-say.html.

[231] Department of Justice, "Manhattan U.S. Attorney Announces Charges Against Seven Iranians For Conducting Coordinated Campaign Of Cyber Attacks Against U.S. Financial Sector On Behalf Of Islamic Revolutionary Guard Corps-Sponsored Entities," U.S. Attorney's Office, Southern District of New York, March 24, 2016, https://www.justice.gov/usao-sdny/pr/manhattan-us-attorney-announces-charges-against-seven-iranians-conducting-coordinated.

www.manaraa.com

attack as reporting suggests that the United States maintains extensive network accesses within Iran and is prepositioned to conduct an offensive cyber campaign against their key infrastructure.[232] However, it seems unlikely because of the effect such an attack would have caused on the already minimal relations between the United States and Iran. It should be noted that during the same time period as Operation Ababil, the United States was pursuing negotiations for the Iranian nuclear deal (also known as the "Joint Comprehensive Plan of Action"). The United States may have found that delaying the indictments, and not responding with a punishing cyber attack, was more prudent until after the deal was signed. Furthermore, by solely issuing indictments, the United States was able to reinforce the rule of law over the domain—a tactic that the Obama administration had already seen work in their relations with China.

Russia[233]

In the weeks leading up to and following the 2016 U.S. presidential election, reports spread via the media suggesting Russian interference in the election.[234] In a joint assessment on the election, the U.S. Intelligence Community described a widespread and multifaceted campaign conducted by the Russian government that attempted to undermine the U.S. election

[232] David E. Sanger and Mark Mazzetti, "U.S. Had Cyberattack Plan if Iran Nuclear Dispute Led to Conflict," *The New York Times*, February 16, 2016, https://www.nytimes.com/2016/02/17/world/middleeast/us-had-cyberattack-planned-if-iran-nuclear-negotiations-failed.html.

[233] Note that at the time of writing there are at least three separate federal investigations into this matter. The results of which unfortunately may significantly rewrite the details of this case and expose Russian cooperation with Americans and the undermining of President Obama's bargaining position.

[234] Eric Lipton, David E. Sanger, and Scott Shane, "The Perfect Weapon: How Russian Cyberpower Invaded the U.S.," *The New York Times*, December 13, 2016, https://www.nytimes.com/2016/12/13/us/politics/russia-hack-election-dnc.html.

www.manaraa.com

through hacking of sensitive information and the spread of propaganda. Indeed, the Democratic National Committee and the networks of several state and local level elections were hacked. Their confidential information was posted to *WikiLeaks* and false reports were disseminated via several media outlets that are widely seen as propaganda organs of the Russian state.[235]

In the wake of the election, President Obama was pressed by the media why he had not responded to the outcome sooner, he noted during a December 16th press briefing the following:

> What we've simply said is the facts, which are that, based on uniform intelligence assessments, the Russians were responsible for hacking the DNC, and that, as a consequence, it is important for us to review all elements of that and make sure that we are preventing that kind of interference through cyberattacks in the future…. Now, with respect to how this thing unfolded last year, let's just go through the facts pretty quickly. At the beginning of the summer, we're alerted to the possibility that the DNC has been hacked, and I immediately order law enforcement as well as our intelligence teams to find out everything about it, investigate it thoroughly, to brief the potential victims of this hacking, to brief on a bipartisan basis the leaders of both the House and the Senate and the relevant intelligence committees. And once we had clarity and certainty around what, in fact, had happened, we publicly announced that, in fact, Russia had hacked into the DNC. And at that time, we did not attribute motives or any interpretations of why they had done so. We didn't discuss what the effects of it might be. We simply let people know -- the public know, just as we had let members of Congress know -- that this had happened. ….And, finally, I think it's worth pointing out that the information was already out. It was in the hands of WikiLeaks, so that was going to come out no matter what. What I was concerned about, in particular, was making sure that that wasn't compounded by potential hacking that could hamper vote counting, affect the actual election process itself. And so in early September, when I saw President Putin in China, I felt that the most effective way to ensure that that didn't happen was to talk to him directly and tell him to cut it out, and there were going to be some serious consequences if he didn't. And, in fact, we did not see further tampering of the election process. But the leaks through WikiLeaks had already occurred. So when I look back in terms of how we handled it, I think we handled it the way it should have been handled. We allowed law enforcement and the intelligence community to do its job without political influence. We briefed all relevant parties involved in terms of what was taking place. When we had a consensus around what had happened, we announced it -- not through the White House, not through me, but rather through the intelligence communities that had actually carried out these investigations. And then we allowed you

[235] Intelligence Community Assessment, "Assessing Russian Activities and Intentions in Recent US Elections," Office of the Director of National Intelligence- National Intelligence Council ICA 2017-01D, January 6, 2017, https://assets.documentcloud.org/documents/3254237/Russia-Hack-Report.pdf.

www.manaraa.com

and the American public to make an assessment as to how to weigh that going into the election.[236]

This extraordinarily candid statement by President Obama describes several aspects of the decision-making behind the U.S. government's response. First, Obama was worried about the escalatory risk of further hacking into the election system that a military response may have provoked. Second, he notes that investigations take time and the importance of attribution. Third, President Obama believed that his direct discussions with Putin played a role in deterring the Russia from conducting further attacks.[237] What is striking about this statement is that Obama's initial response clearly reflects a cautious hand that felt vulnerable to further attacks and sought to avoid a cyber war. However, he acknowledges there is a temporal aspect to consider when responding to cyber attacks. He notes the difficulty of attribution, but stresses that it is possible with time. In other words, attributing injects breathing space into what otherwise might have been a time sensitive crisis.

In the weeks following President Obama's remarks, the United States formally responded by expelling thirty-five Russian diplomats that the United States said were linked to espionage and closing two waterfront estates located in Maryland and New York that had historic ties to being used for spying. Furthermore, through the Department of Treasury, the U.S. placed sanctions on four top Russian officials, a Russian intelligence unit, Glavnoye Razvedyvatel'noye Upravleniye (GRU), and three Russian cyber security companies that had been involved in the

[236] Barack Obama, "Press Conference by the President," The White House, December 16, 2016, https://obamawhitehouse.archives.gov/the-press-office/2016/12/16/press-conference-president.

[237] A more comprehensive discussion of the exchanges between Presidents Obama and Putin on this matter are captured in the Cross-Domain Deterrence section of this paper.

www.manaraa.com

operation according to U.S. intelligence.[238] Despite initial threats from the Russian foreign minister, Putin responded that he would avoid retaliation and indicated that he looked forward to more favorable relations with the incoming Trump administration.[239]

Conclusion

Despite the fact that the cyber domain is not deterrence dominant, it remains puzzling that we have not observed significant escalation—either inadvertent or deliberate. The balance of terror logic that governed nuclear deterrence has little bearing when applied to the cyber realm because states simply cannot leverage enough devastation to threaten the annihilation of another civilization using cyber power alone. Moreover, this is compounded by the attribution problem, because states may calculate that they can get away with a cyber attack as long as they can avoid attribution.

One may infer that in light of the attribution problem that there is little incentive for prudent behavior in the cyber domain and, therefore, that the domain is inherently escalatory. However, as a whole we have not seen the widespread devastation that some pundits have suggested, or even escalatory responses to cyber attacks, indicating that actors are indeed

---

[238] David E. Sanger, "Obama Strikes Back at Russia for Election Hacking," *The New York Times*, December 29, 2016, https://www.nytimes.com/2016/12/29/us/politics/russia-election-hacking-sanctions.html. As an interesting side note, Adam Segal, the Director of the Digital and Cyberspace Policy Program at the Council on Foreign Relations, noted that during a recent Cyber security dialogue between Center for Strategic and International Studies (CSIS) and China Institute of Contemporary International Relations (CICIR) that he was a participant at that several of the Chinese representatives did not see any deterrent value in the Obama's response to Russian interference in the election and questioned how it was proportionate for the harm that the administration claimed. For further information see, Adam Segal, "The Continued Importance of the U.S.-China Cyber Dialogue," *Council on Foreign Relations- Net Politics* (online blog), January 23, 2017, http://blogs.cfr.org/cyber/2017/01/23/the-continued-importance-of-the-u-s-china-cyber-dialogue/#more-4554.

[239] Neil MacFarquhar, "Vladimir Putin Won't Expel U.S. Diplomats as Russian Foreign Minister Urged," *The New York Times*, December 30, 2016, https://www.nytimes.com/2016/12/30/world/europe/russia-diplomats-us-hacking.html.

www.manaraa.com

moderating their actions in cyberspace.[240] This moderating behavior is indicative of the self-dampening nature of cyber operations.

An important implication, therefore, is that if a state is determined to escalate in response to a cyber attack, it may find a cross-domain response more appealing than a cyber one, just as cross-domain deterrence strategies have been appealing for policymakers. This is because the cyber weapon required to send the ideal proportionate response that a policymaker may want to send may not be waiting in her cyber arsenal. Indeed, most cyber operations require custom accesses and capabilities that are unrealistic to produce in an appropriate timeframe. Developing precision cyber weapons takes time and resources. As already noted, cyber weapon development, particularly against critical infrastructure requires extensive intelligence support and high research and development costs. This cost cannot be amortized over numerous targets because cyber weapons, particularly those against the most critical targets, lack universal lethality.

Additionally, there is nothing that guarantees that a cyber-pure response will be perceived or accurately attributed unless coupled within the context of a larger military operation where attribution can be assumed or declared through some sort of diplomatic message. Indeed, the traditional factors that hinder communication between state actors (e.g. different languages, lack of direct communication, etc.) are compounded in a domain that is prone to obfuscation.

Further, in a domain that is devoid of norms of appropriate responses, a tit-for-tat in cyberspace could set off a cross-domain escalatory spiral if the response is seen as a gross escalation or on the contrary, could be seen as weak if not enough damage is triggered. Indeed, there is an incentive to not create mutually agreed thresholds, or to cross what another state may consider a red line (such as attacks on identified critical infrastructure), for fear that the strategic

[240] Libicki, *Cyberspace in Peace and War*, 273; Libicki, *Crisis and Escalation in Cyberspace*, 73.

www.manaraa.com

ambiguity that currently exists amongst some states actors in this space may become rigidly defined.[241] Unlike cross-domain escalatory responses, escalation solely within the cyber domain is self-dampening because a tit-for-tat exchange in cyberspace may only be iterated in an increasingly escalatory fashion a limited number of times given access limitations, capability constraints, and evolving defenses. Waging an effective long term cyber campaign is difficult and extraordinarily costly under a condition of no universal lethality in cyberspace. Over time, actors may exhaust their arsenals or have them rendered inert through vulnerability patching the longer they are exposed.

Finally, the anonymizing nature of the cyber domain means that attribution of any cyber attack takes time. The temporal requirements of achieving attribution creates breathing space and resistance to escalatory behavior across all domains. There is no concept of State A's nuclear armed ICBMs will hit their targets in X minutes, and that therefore the defending state must launch theirs in Y minutes to avoid annihilation. This is because knowing who to respond against following a cyber attack takes time and, at least for the near future, the independent use of cyber power is not capable of creating enough devastation to annihilate, let alone cripple, the warfighting ability of another state.[242]

Overall, this analysis suggests reasons for both optimism and pessimism. On the one hand, the cyber domain is inherently self-dampening, which means that we are unlikely to observe dangerous cyber escalatory spirals. Additionally, deterrence by denial is feasible under

[241] Indeed, states may elect to maintain situations of strategic ambiguity because 1) they face incentives to exploit the strategic ambiguity themselves and 2) uncertainty of the decisiveness of cyber warfare and not wanting to be at a the tactical or strategic disadvantage in the future.

[242] For further reference on the destructive limitations of cyber power as an independent tool of coercion see, Borghard and Lonergan, "The Logic of Coercion in Cyberspace." Also, the authors note that technical attribution may be forgone when the attack is coupled with a diplomatic message or placed in the context of a larger conflict.

www.manaraa.com

certain conditions, which suggests that escalation could potentially be contained at a given rung of the cyber escalation ladder if states have developed denial capabilities to contain the ratcheting up of an interaction in the domain. On the other hand, if a state seeks to deliberately escalate as part of a bargaining strategy, or feels domestic political pressure to respond to a cyber attack, it may turn to kinetic means to do so. The fact that the cyber domain is multipolar means that the risk of great power conflict stemming from crises that arises in cyberspace is nontrivial. Therefore, it is imperative to consider the extent to which arms control and confidence building measures are possible in the cyber domain.

www.manaraa.com

Arms Control and Confidence Building Measures for the Cyber Domain

Shawn W. Lonergan

This paper assesses how, if at all, arms control, which seeks to institutionalize constraints on offensive military technology and guard against inadvertent conflict and escalation, can be used to produce stability to mitigate the dangers posed by cyber operations. Though arms control has direct implications for crisis management, arms control, as addressed in this paper, seeks to identify means to shape the incentives that lead to crises and thereby foster stability between cyber powers. While this paper argues that cyber arms control is not a realistic endeavor at this time, I submit that confidence building measures (CBMs), a step short of arms control, are not only necessary, but are also viable vehicles to mitigate some of the contributing factors leading to instability posed by cyber operations. In assessing current cyber CBM development efforts, this paper creates a novel framework to better understand these efforts and to demonstrate areas that are not being addressed and remain as potential flashpoints that could exacerbate tensions and spark conflict.

The first section of this article reviews arms control theory as refined during the Cold War and assesses what relevance this scholarship has on the control of cyber arms. My analysis finds that arms control in cyberspace is difficult because of the ambiguity surrounding the strategic balance of cyber weapons, an inability to monitor for compliance, the dynamic nature of the methods and means of cyber operations, and issues of assigning and enforcing culpability. However, an incentive exists, at a minimum, to avoid inadvertent conflict and stabilize disruptions to international security stemming from cyber operations. Similar issues existed in the early stages of nuclear weaponry where mistrust prevented arms control agreements. My

www.manaraa.com

research notes these fears were mitigated over time through the development of confidence building measures, mutually agreed-upon notification requirements that signaled the intent behind military activities that might otherwise promote fear or indicate a pending attack. The paper proceeds by reviewing the general theory of confidence building measures and their incorporation into the Helsinki Final Act, the benchmark for all CBM agreements. My analysis concludes by building on the existing scholarship of CBMs in the cyber domain, noting how the domain creates new categories of measures that did not apply to nuclear weapons, and proposes novel measures that could be implemented in light of this research into Cold War CBMs. Indeed, my research suggests that that there are additional steps that can be taken to enhance mutual security and guard against inadvertent conflict stemming from cyber operations.

Toward Cyber Arms Control?

One tried and tested method to alter the military incentives for the use of offensive technologies is to reshape them through arms control. Indeed, states seek arms control regimes when doing so can improve mutual security, often this takes the form of increased transparency of national security policies between states.[243] There is a robust canon that normatively addresses how arms control should occur in areas that are prone to increase the likelihood of war, with some authors, such as Schelling and Halperin, going so far as to suggest that it is irresponsible for states to not create measures that avoid false alarms and mistaken intentions.[244] Indeed, arms

[243] Hedley Bull describes the objectives of arms control across economic, moral, and the international security domains in Chapter 1 of his seminal work, *The Control of the Arms Race: Disarmament and Arms Control in the Missile Age*, vol. 2 (London: Praeger for the Institute for Strategic Studies, 1961). Bull's objectives of arms control that relate to inadvertent conflict escalation are of special relevance to this article.

[244] For instance, see Thomas C. Schelling and Morton H. Halperin, *Strategy and Arms Control*, (Washington: Pergamon-Brassey, 1985), ix-6., and, David W. Kearn Jr., *Great Power Security Cooperation: Arms Control and the Challenge of Technological Change* (Lexington, MA: Lexington Books, 2014), Chapter 1.

www.manaraa.com

control agreements typically follow when the capabilities in question are destabilizing to international security.[245] However, it is puzzling that cyber operations, which both upset traditional balances of power and create the conditions for misperceptions that could lead to inadvertent conflict, have not prompted states to pursue arms control regimes to secure traditional power asymmetries and guard against undesired outcomes stemming from uncertainty. As Schelling and Halperin note, arms control can support security policy and occurs when it is in each side's strategic interest.[246] In an extension of this logic, Jervis notes that the fundamental postulate of arms control that "…hostile states almost always have important interests of military policy in common," which he noted often could be crisis stability.[247] Indeed, several countries, such as the United States and Russia, created bilateral agreements to establish hotlines to guard against misunderstandings stemming from cyber operations in a crisis. However, these efforts have stopped short of the creation of targeting norms, limits on the production of cyber weapons, or the development of more robust mechanisms to promote transparency of the domain. This hesitance is due not to a lack of an incentive, but to the inability to satisfy the basic requirements for arms control. There are three theories of security cooperation that identify when arms control agreements are most likely to form in light of technological change. Offense-Defense Theory, Technological Opportunism, and Military Expectation Theory all make predictions as to when security cooperation is likely to occur and

[245] Examples include the 1963 "Hot Line" Agreement, the 1971 "Accidents Measures" Agreement, and the 1972 Antiballistic Missile (ABM) Treaty.

[246] Schelling and Halperin, *Strategy and Arms Control*, 141-142.

[247] Robert Jervis, "Arms Control, Stability, and Causes of War," *Political Science Quarterly* 108, no. 2 (1993): 239-253, 241, 248.

www.manaraa.com

lead to arms control efforts.[248] The following section outlines each theory and identifies why cyber arms control formation may not have been pursued despite an incentive for states to seek greater stability.

Offense-Defense Theory

Offense-Defense Theory seeks to explain the propensity of states to go to war.[249] The theory rests under the umbrella of defensive realism and thus assumes that states inherently value the status quo and avoiding armed conflict.[250] The theory focuses on two key variables: technology and geography.[251] When technology and geography favor the offense, the cost of mounting an offensive is low. In other words, when offense is dominant the likelihood of states

[248] All three theories sit within the Realist school of thought which envisions states, under a condition of international anarchy, in a constant security competition with other state actors.

[249] For a further discussion on the offense-defense theory start with Robert Jervis, "Cooperation Under the Security Dilemma," *World Politics* 30, no. 2 (1978): 167-214. Continue with Charles L. Glaser and Chaim Kaufmann, "What is the Offense-Defense Balance and How Can We Measure It?" *International Security* 22, no. 4 (1998): 44-82; Stephen Van Evera, "Offense, Defense, and the Causes of War," *International Security* 22, no. 4 (1998): 5-43; Stephen Van Evera, *Causes of War: Power and the Roots of Conflict* (Ithaca, NY: Cornell University Press, 2013); Ted Hopf, "Polarity, the Offense-Defense Balance, and War," *American Political Science Review* 85, no. 2 (1991): 475-493. Also, see the following critiques, Karen Ruth Adams, "Attack and Conquer? International Anarchy and the Offense-Defense-Deterrence Balance," *International Security* 28, no. 3 (2003/04): 45–83; Richard K. Betts, "Must War Find a Way?: A Review Essay," *International Security* 24, no. 2 (1999): 166-198; James W. Davis, et al., "Taking Offense at Offense-Defense Theory," *International Security* 23, no. 3 (1998/99): 179-206; and Sean M. Lynn-Jones, "Offense-Defense Theory and its Critics," *Security Studies* 4, no. 4 (Summer 1995): 660-691.

[250] For a further discussion on defensive realism and its relation to offensive realism see: Eric J. Labs, "Beyond Victory: Offensive Realism and the Expansion of War Aims," *Security Studies* 6, no. 4 (1997): 1-49; Stephen G. Brooks, "Dueling Realisms," *International Organization* 51, no. 3 (Summer 1997): 445-77; John J. Mearsheimer, "The False Promise of International Institutions," *International Security* 19, no. 3 (1994/95); Jack Snyder, *Myths of Empire,* (Ithaca: Cornell University Press, 1991), 12; Fareed Zakaria, "Realism and Domestic Politics," *International Security* 17, no. 1 (1992): 177-98; and Benjamin Frankel, "Restating the Realist Case," *Security Studies* 5, no. 3 (Spring 1996): 1-12.

[251] Jervis, "Cooperation Under the Security Dilemma," 194-199.

www.manaraa.com

pursuing revisionist aspirations through conquest becomes more likely. Likewise, when technology and geography favor the defense conquest is less likely and states are more apt to accept the status quo. Yet this leads to one of the basic paradoxes of arms control, when it is most needed it is often unsuccessful, but when the need for it is less, it easier to achieve.[252] Indeed, under conditions of defense dominance, states feeling secure, are more likely to cooperate and lock in the status quo through arms control agreements.[253] However, Robert Jervis notes what he calls the "security dilemma," a condition where by one state's efforts to bolster its defenses actually may be self-defeating if it is perceived as threatening by another actor, causing it to bolster its own capabilities in turn.[254] This can lead to a state inadvertently undermining the status quo and their own security by an action that they intended to do the opposite. According to Jervis this could cause states to tragically spiral to war.[255] When the risk of inadvertent conflict exists between actors, arms control measures that create transparency of the intent behind national security policies and actions are more likely to form.

Where geography is a static variable as opposed to technology which is dynamic, scholars and decision makers naturally tend to focus on the effect changes in armaments and the supporting doctrine has on the balance. However, a danger exists that leaders may misperceive the true balance at any given time, resulting in a disastrous outcome. Indeed, Jervis notes that during World War I, belligerents thought that the weapons of the day favored the offense,

---

[252] The author would like to thanks Robert Jervis for this point.

[253] Andrew Kydd, "Sheep in Sheep's Clothing: Why Security Seekers Do Not Fight Each Other," *Security Studies* 7, no.1 (Autumn 1997): 114-155, 119-120.

[254] Jervis, "Cooperation Under the Security Dilemma."

[255] Robert Jervis, *Perception and Misperception in International Politics* (Princeton, NJ: Princeton University Press, 1976), 62-78.

www.manaraa.com

however, "If they had known the power of the defense beforehand, they would have rushed for their own trenches rather than for the enemy's territory."[256] This ambiguity can lead to a false sense of military advantage for the offense over defense and may lead to conflict.[257]

Stemming from this ambiguity, offense-defense proponents have devised two means to measure the balance. The first method assesses the force ratios necessary to take and hold land.[258] In other words, how many offensive troops are needed to overwhelm defensive forces. By tradition, for every defender, three offenders are necessary to win an engagement. However, technological advancements and supporting doctrine can shift the ratio and award an advantage to one side. The second method envisions assessing the offense-defense balance through a cost-exchange comparison.[259] Jervis, who first proposed the method, argued that offense is dominant when $1 spent on offense can defeat $1 spent on defense.

With cyber warfare, geography remains constant with each state sharing a virtual border. However, technology is constantly changing with accesses being gained and lost and new capabilities being developed while others expire. This has given rise to a budding debate about the offense-defense balance in cyberspace, with most scholars noting challenges of measuring the balance in the domain.[260] The reality is that the balance is exceedingly difficult to measure at

---

[256] Jervis, "Cooperation Under the Security Dilemma," 191.

[257] Ibid., 199-214.

[258] Jack S. Levy, "The Offensive/Defensive Balance of Military Technology: A Theoretical and Historical Analysis," *International Studies Quarterly* 28, no. 2 (1984): 219-238, 227-228.

[259] Jervis, "Cooperation Under the Security Dilemma," 188; Levy, "The Offensive/Defensive Balance of Military Technology," 227.

[260] Concerning the offense-defense balance in cyberspace see Adam P. Liff, "Cyberwar: A New 'Absolute Weapon'? The Proliferation of Cyberwarfare Capabilities and Interstate War," *Journal of Strategic Studies* 35, no. 3 (June 2012): 401-428; Ilai Saltzman, "Cyber Posturing and the Offense-Defense Balance," *Contemporary Security Policy* 34, no. 1 (April 2013): 40-63;

www.manaraa.com

any given time given the necessary secrecy surrounding cyber arsenals. Furthermore, in a point that will be elaborated shortly, even if there was complete transparency it would be challenging to assess relative cyber power between actors, since there is no universal lethality of cyber weapons. The cost-exchange method is the most viable means to assess the balance, though it is not without its challenges. First, most states keep their cyber budgets classified as to not tip off to an adversary the technologies that they are investing capabilities against. Second, it is difficult to differentiate between offensive and defense expenditures given that, at least from a military perspective, the forces that conduct offensive cyber operations may very well be the same as those that conduct defensive cyber operations. This suggests that it may not be possible to clearly divide expenditures into offensive versus defensive baskets. It may be the case that one is more likely to be able to assess the systemic balance in general terms, as opposed to empirically measuring the dyadic balance at any moment unless a researcher is willing to develop a restrictive proxy measure. Such a measure may be parsimonious, but may carry limited explanatory power. Further research is needed on assessing the balance and is beyond the scope of this paper. Yet it is important to note that an implication of this measurement difficulty is that the Offense-Defense Theory criterion of defense dominance being necessary for arms control formation may not be sufficiently satisfied because of this inability to measure the balance at any given time. The theory suggests that this inability by all parties to accurately measure the balance will confound security cooperation, thus indicating that until an accepted means to assess the balance is devised, cyber arms control formation is unlikely.

Rebecca Slayton, "What Is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment," *International Security* 41, no. 3 (Winter 2016/17): 72-109; Jon R. Lindsay, "Stuxnet and The Limits of Cyber Warfare," *Security Studies* 22, no. 3 (2013): 365-404.

www.manaraa.com

Technological Opportunism

In the book *War and the Engineers*, Keir Lieber argues that Offense-Defense Theory is flawed because it views shifts in technology as the cause of war or peace.[261] Rather, Lieber purports in his "Technological Opportunism" theory, changes in technology are filtered through the politics of states and are molded to help the state pursue a strategic end. The theory, which is nested in the offensive realism tradition, pictures states as security maximizers that use technology to improve their relative position in the world. By implication, uncertainty over the offense-defense balance has no bearing on security competitions because states seek to escape the effects of the security dilemma by maximizing their power. Though Lieber argues that arms control efforts to prevent misperceptions are misguided because they do not make war less likely, he does not find them as fruitless endeavors.[262] Lieber conceives arms control as a means for states to coordinate to more efficiently allocate material resources. Indeed, coordination can be a means to eliminate material requirements, such as having to build a wall against a nonexistent threat, which can be a distraction from states' being able to otherwise pursue their strategies.

Though there are significant shortcomings in Lieber's theory, of relevance to this analysis is that the theory does not provide a robust framework to assess the conditions under which security cooperation is likely.[263] Furthermore, though the theory assesses the role of uncertainty, it does not account for the role of secrecy, a deliberate condition, and its effect on cooperation. In

[261] Keir A. Lieber, *War and the Engineers: The Primacy of Politics over Technology* (Ithaca: Cornell University Press, 2005), 2. For other critiques of Offense-Defense Theory see footnote 7.

[262] Lieber, *War and the Engineers,* 155.

[263] For further critiques of Technological Opportunism see Damon Coletta's review of "War and the Engineers: The Primacy of Politics over Technology," *Air & Space Power Journal* 25, no. 3 (2011): 111-113; and Kearn, *Great Power Security Cooperation*, 20-23.

www.manaraa.com

its application to cyber warfare, the greatest contribution of the theory is that it may explain, at least in some cases, why states have used cyber operations to pursue their strategic ends—often to challenge traditional power asymmetries.

Military Expectation Theory

David Kearn puts forth a theory of when states are most likely to pursue arms control agreements in his book, *Great Power Security Cooperation: Arms Control and The Challenge of Technological Change*. Kearn's "Military Expectation Theory" attempts to bridge the offense-defense and technological opportunism camps by presenting military technology as both a means for states to pursue their strategies and also as a cause of deliberate or inadvertent fear.[264] Kearn submits that there are two variables that shape the likelihood of arms control formation: 1) the expectation of the military effect of the technological change; and 2) the perceived effectiveness of existing modes of surveillance.

To Kearn, if a weapon system is perceived as being decisive—that is, "…one[s] that offer [to] dramatically improve capabilities over comparable existing systems"—then states that possess them may be willing to limit their proliferation. States without are motivated to pursue them, thus confounding cooperation.[265] However, when weapons are not decisive and offer only *incremental* benefits to existing systems, security cooperation is achievable. Furthermore, arms control formation is more likely to occur if states have confidence in the abilities of their intelligence apparatus to detect cheating.

In the application of Military Expectation Theory to cyber operations, it is important to consider if cyber is, indeed, a decisive weapon. As an independent tool of coercion, cyber

[264] Kearn, *Great Power Security Cooperation*, 23-24.

[265] Ibid., 24-27.

www.manaraa.com

weaponry is not decisive, at least until societies become reliant on automation to support their essential services (assuming they hold resiliency and redundancy constant).[266] However, when coupled with traditional elements of national power, cyber weapons can drastically improve existing systems by making their employment more effective. For example, it is alleged that Israel's 2007 aerial bombing of a Syrian nuclear enrichment faculty was only made possible by a cyber attack against Syria's integrated air defense systems which enabled the uncontested bombing run.[267] In this sense, cyber power can serve as a decisive military shaping operation.

This incentive for states to hedge the future of cyber capabilities is not unjustified, though is perhaps more psychological than rational.[268] Indeed, state actors live in fear of cyber war and thus are hesitant to limit the development, stockpiling, and use of cyber weaponry for fear of finding themselves at a strategic disadvantage. With warnings of "World War 3.0"[269] and a "cyber-Pearl Harbor"[270] in mainstream media, this fear is not without merit.[271] States do not want to be at a military disadvantage in this domain, particularly given the presumed offensive

---

[266] Erica D. Borghard and Shawn W. Lonergan, "The Logic of Coercion in Cyberspace," *Security Studies* 26, no. 4 (2017): 452-481, 477-480.

[267] Richard A. Clark and Robert K. Knake, *Cyber War: The Next Threat to National Security and What to Do About It* (New York: Ecco, 2010), 1-8.

[268] It is the author's contention that many policymaker's and military planners perceive that cyber weapons are indeed decisive or, at least, potentially decisive weapons. Though cyber weapons can produce decisive shaping effects as part of a military operation, there independent offensive use is not militarily decisive as defined in this article. For more on this discussion see, Borghard and Lonergan, "The Logic of Coercion in Cyberspace."

[269] Michael Joseph Gross, "World War 3.0," *Vanity Fair,* May 2012.

[270] Leon Panetta, "Remarks by Secretary Panetta on Cybersecurity to the Business Executives for National Security," U.S. Department of Defense, New York City, October 11, 2012.

[271] For a rebuttal see, Thomas Rid, *Cyber War Will Not Take Place* (Oxford University Press, 2013).

www.manaraa.com

parity that exists between some states with offensive cyber capabilities.[272] Indeed, while serving as Chairman of the Joint Chiefs of Staff, General Martin Dempsey noted that the cyber domain is the only domain where the United States has peer competitors.[273] Historically, when uncertainty is created concerning the effect of a technological change, extensive strategic analysis has followed. However, Schelling notes that it may be impossible to build a common understanding within a single government, let alone between states, as to what the full effect is.[274] Therefore, the perceived potential decisiveness of cyber weaponry, coupled with an uncertain future concerning cyber conflict, suggests that a state's dominant strategy could be to build extensive and diverse offensive cyber arms capabilities and direct espionage efforts to identify means to thwart a foreign adversary's cyber warfighting capability under Kearn's logic.[275]

---

[272] As previously noted, there is no effective measure of relative offensive cyber power between states. However, on a subjective scale, certain states (e.g., United States, China, Russia, France, Great Britain, and Israel) are capable of wreaking widespread havoc against another actor via the offensive use of cyber power.

[273] Chris Wallace, "Gen. Dempsey reacts to Paris attacks; Sens. Hoeven, Coons talk Keystone showdown," *Fox News Sunday*, January 11, 2015, http://www.foxnews.com/on-air/fox-news-sunday-chris-wallace/2015/01/11/gen-dempsey-reacts-paris-attacks-sens-hoeven-coons-talk-keystone-showdown#p//v/3983017948001.

[274] Thomas C. Schelling, "Reciprocal Measures for Arms Stabilization," *Daedalus* 89, no. 4 (1960): 892-914, 898.

[275] Though beyond the focus of this article, an interesting counterpoint to the importance that Kearn places on the decisiveness of a weapon and the likelihood of arms control is the relationship between the lethality (or strategic utility) of a weapon and the likelihood of arms control agreements. Weapons that are catastrophically lethal (e.g. nukes) prompt arms control agreements because their anticipated destructive costs are so high and weapons that are not particularly useful but harmful (e.g. chemical weapons and lasers) prompt arms control. Arguably these agreements occurred in these examples because great powers did not see much utility in their employment and therefore were willing to pursue mechanisms to constrain their use. However, one may be unlikely to see arms control formation on weaponry that falls in the spectrum between these two camps because of the perceived utility these capabilities may possess in war. Offensive cyber capabilities presumably fall somewhere between these two bounds.

www.manaraa.com

Kearn's second variable, monitoring for defection, is not a new criterion of arms control agreements. Long before Kearn, scholars discussed the verification problem in the context of arms control agreements and noted the importance of verifying what is in a current state's arsenal and also monitoring future develop efforts for compliance with any accord.[276] This issue has also been noted in the literature on cyber conflict with scholars pointing to the obfuscating nature of the domain as confounding verification efforts.[277] In the traditional sense,

> Monitoring entails the gathering of data on treaty-related activities, whether by national technical means, on-site inspections, data exchanges, or intelligence. Based on that monitoring data, verification entails a judgement concerning parties; compliance (or non-compliance) with the limits set by the arms control agreement.[278]

The requisites for establishing this level of compliance in the cyber domain would require affected parties to agree to extremely intrusive access to a government's networks. In other words, if a state agreed to stop developing a type of cyber weaponry and agreed to allow verification, it would essentially have to open its networks to the other state, or a third party, to inspect for compliance. Not only would it be technically impossible to scour every government network for evidence of cyber weapons—assuming a state would be foolish enough to construct cyber weapons in violation of a treaty and store them on an inspectable network—but it also

---

[276] For instance, see William R. Frye, "Characteristics of Recent Arms-Control Proposals and Agreements," *Daedalus,* 89, no. 4 (1960): 723-743, 725, 735; Karl Pieragostini. "Arms Control Verification: Cooperating to Reduce Uncertainty," *Journal of Conflict Resolution* 30, no. 3 (September 1986): 420-444; William C. Potter, *Verification and Arms Control* (Lexington, MA: Lexington Books, 1985); and Donald Wittman, "Arms Control Verification and Other Games Involving Imperfect Detection," *American Political Science Review* 83, no. 3 (1989): 923-945.

[277] Phillip A. Johnson, "Is It Time for a Treaty on Information Warfare?" *International Law Studies* 76, no.1 (2002); and Neil C. Rowe et al., "Challenges in Monitoring Cyberarms Compliance," in *Conflict and Cooperation in Cyberspace: The Challenge to National Security*, eds. Panayotis A. Yannakogeorgos and Adam B. Lowther (Boca Raton, FL: Taylor & Francis, 2014), 81-100.

[278] Lewis A. Dunn, "Arms Control Verification: Living with Uncertainty," *International Security* 14, no. 4 (Spring 1990): 165-175, 165, footnote 1.

www.manaraa.com

would be unfathomable for one state to allow another, or any outside actor, to have unfettered access to its networks because doing so would increase the risk of cyber exploitation and attack, precisely what it was seeking to avoid in the first place.[279] Indeed, the only way that a state can have full confidence that another state is not violating the agreement is if it is monitoring the potential belligerent's every network. Besides being untenable from a resource perspective, this type of persistent access could be a violation of the agreement it is attempting to enforce.[280]

Another means to ensure compliance would be to have monitoring occur through national technical means of intelligence. During the Cold War this often took the form of analyzing imagery collected from satellites that monitored the nuclear posture of another state. The equivalent in cyberspace would be using cyber espionage to collect on internal networks. However, where satellite collection is entirely passive, gaining access to and absconding with data from sensitive government networks is invasive and, if detected, can be destabilizing. Indeed, the critical systems that would likely be intelligence targets also would likely be prized targets for attack in a conflict. This inability to perceive the intent behind such access, whether it is for compliance monitoring or preparation for an attack, could be a source of conflict.[281]

In sum, Military Expectation Theory predicts that cyber arms control formation is unlikely for two reasons. First, the debate over the decisiveness of cyber power will most likely leave leaders erroring on the safe side and resisting any efforts to control a capability that they

---

[279] For further discussion on the logic behind this risk see Ben Buchanan, *The Cyber Security Dilemma* (Oxford: Oxford University Press, 2016); and Shawn W. Lonergan, "Cooperation under the Cybersecurity Dilemma," in *Confronting Inequality: Wealth, Rights, and Power*, ed. Hugh Liebert, Thomas Sherlock, and Cole Pinheiro (New York: Sloan, 2016).

[280] The author would like to thank Robert Jervis for this point.

[281] Indeed, as already noted in the previous discussion of the security dilemma, misperceptions of intent can lead to inadvertent conflict.

www.manaraa.com

are still understanding the true character of. Second, there are currently no adequate means to monitor for compliance without leading to greater vulnerability that would be acceptable to all signatories.

Additional Obstacles Unique to Cyber Arms Control

In addition to the issues identified by the preceding theories, three additional problems confound cyber arms control efforts. First, arms control seeks to achieve *mutual* security through either a preservation or mutual adjustment of the status quo that contributes to greater security of all signatories. In other words, arms control either maintains or forms a new strategic balance between potential adversaries.[282] Indeed, Hedley Bull noted that the balance of military power is a central tenet of any arms control agreement for these regimes do not abolish military power, but seek to stabilize a military situation.[283] Therefore, the prototypical arms control regime places limits on capabilities, be it their numbers or employment. Yet as will be further discussed shortly, this view of arms control is overly reductionist for arms control, as forming a strategic balance requires stability not only of arms racing behavior but also during a crisis. Underlying this balance is some mutual concept of relative strength between potential belligerents. However, in cyberspace this has no bearing as there is no measure of relative strength. Unlike traditional weaponry, nuclear weapons, or even chemical ordinances, where one can count the number of warheads or pounds of a virulent gas that a state possesses, a balance of cyber capabilities between states is nonsensical because there is no measure of relative cyber power between states. Cyber weapons lack universal lethality, indeed targets of cyber weapons are often are key pieces of critical infrastructure, which requires the development of custom exploits not only to gain

[282] Richard Burt, "Reassessing the Strategic Balance," *International Security* 5, no. 1 (Summer 1980): 37-52.

[283] Bull, *The Control of the Arms Race,* 67-68, 75.

www.manaraa.com

access, but also to manipulate the system as an actor desires.[284] These capabilities are often cloaked in secrecy so that states cannot tailor defenses against these niche capabilities.[285] Therefore, cyber arms control cannot lead to a strategic balance between state actors because the balance defies quantification.

Second, the attack vectors and offensive capabilities at the tactical level of the cyber domain are continuously evolving. In the nuclear arena, where innovations that challenged arms control agreements such as the introduction of multiple independently targetable reentry vehicles (MIRVs) had long development timelines and deliberate and often observable fielding which enabled breathing space for arms control agreements to adjust or for states to develop other means, such as tailored intelligence or their own complimentary programs, to mitigate the fear the advances posed. However, the open-ended promise of innovation coupled with the constant changing tradecraft that can emerge with little to no warning in the cyber domain will challenge the creation of any agreement. Indeed, new information technology vulnerabilities are discovered every day and the median time between the public reporting of the vulnerability and a vendor marketing a custom exploit is only 30 days.[286] With fields that are prone to technological change, such as cyber arms, one runs the risk that a formalized agreement could be outdated or restrictive in some unanticipated way before the ink has had time to dry. Furthermore, even *if* interested actors could overcome these hurdles, the agreement may *decrease* security in yet to be realized ways. As has been seen in other types of arms control agreements, arms accords have often

[284] For more on this point see Borghard and Lonergan, "The Logic of Coercion in Cyberspace," 465-466.

[285] Though writing about strategic forces during the Cold War, Schelling and Halperin, *Strategy and Arms Control*, 37, note that secrecy of such forces not only hinders arms control formation, but it also creates a risky situation.

[286] Verizon, "2016 Data Breach Investigations Report," 13-14.

www.manaraa.com

encouraged innovation in ways that are often against the spirit of preserving the status quo.[287] This risk is magnified in the cyber domain as it offers countless opportunities for a would-be aggressor. For instance, a cyber agreement that sought to ban a specific attack methodology could encourage a state to redirect those resources to pushing the scientific frontier and developing a new means of attack for which the other signatories are even more vulnerable.

Third, even if the preceding obstacles could be overcome, punishing defection would be difficult to implement because of the difficulty of attribution and divining a proportionate punishment for several reasons. First, in the event of a violation, states would have to attribute the source of the defection and develop a metric of confidence in the attribution that they feel comfortable with using to pursue reciprocity. In the context of a multilateral treaty, the state that detected the violation will need to have compelling evidence that is enough to convince other signatories of the defection. These states may use differing metrics for confidence in attribution than the party that detected the defection and if the violation was attributed through cyber espionage, the state that attributed may have to expose national intelligence collection means and methods. Thus, the value of the intelligence loss must be weighed against the necessity to punish. Indeed, losing an access into a critical system that is necessary for intelligence collection may very well undermine the stability that the arms agreement was seeking to create. Indeed, exposing the access will most likely lead to the loss of their ability to monitor for future violations and collect other intelligence that the state deems is important to their national security. This risk would be compounded if the access is critical to collect information to guard against a surprise attack.

---

[287] For an excellent article discussing how arms control has encouraged military innovation see Charles H. Fairbanks Jr, and Abram N. Shulsky, "From 'Arms Control' to Arms Reductions: The Historical Experience," *The Washington Quarterly* 10, no. 3, (Summer 1987): 59-73.

www.manaraa.com

Second, a state would most likely need to use traditional elements of national power to punish the defection in cyberspace because responding with a timely and proportionate cyber driven tit-for-tat response may be difficult due to resource and access requirements.[288] Punishing a cyber arms agreement defection can create a dilemma because crafting an effective but also proportionate response that relies on physical elements of national power may be difficult to formulate if the defection only caused virtual damage. There are no norms quantifying virtual damage which means that finding an appropriate non-cyber response to punishing defection, such as economic or diplomatic sanctions or a military action, may be difficult to gauge. Theoretically, this could encourage any cyber arms control agreement to establish a punishment schedule for defection as part of the drafting process.

Traditional arms control regimes that mandate bans on capabilities or rely on monitoring for compliance are unrealistic constraints for cyber weaponry as they do not enhance the security of signatories nor would they adequately address issues within the context of technological change. There are three alternative forms of arms control that could be considered to overcome the issues of monitoring and compliance. One approach would be to seek complete offensive disarmament, as some scholars first argued with nuclear weapons, and was pursued with chemical and biological weapons.[289] However, given the leveling effect that cyber weaponry can

[288] Though speaking about coercion in cyberspace, Borghard and Lonergan, "The Logic of Coercion in Cyberspace," note the complexities of generating sufficiently costly signals in cyberspace. As applicable to this case, the temporal aspects of generating a punishment using cyber power may encourage a response in the physical domains of state interaction.

[289] For further discussion on disarmament during the Cold War, see Erich Fromm, "The Case for Unilateral Disarmament," in *Arms Control, Disarmament, and National Security*, ed. Donald G. Brennan, (New York: G. Braziller, 1961); Bull, *The Control of the Arms Race,* chapters 5-8; and Thomas C. Schelling, "Surprise Attack and Disarmament," *Bulletin of the Atomic Scientists* 15, no. 10 (1959): 413-418. With respect to the abolishment of chemical and biological weapons see, the Protocol for the Prohibition of the Use in War of Asphyxiating, Poisonous or other Gases, and of Bacteriological Methods of Warfare (also known as the "Geneva Protocol") and the

www.manaraa.com

bring to a weak state's warfighting ability, the perceived decisive effect that it may afford to all parties, and the widespread proliferation of offensive cyber weaponry that has already occurred, this idea seems unlikely to take hold. Another approach could entail the banning of specific means and methods of operating in the domain, but in a rapidly changing technological environment this potential solution runs the risk of being quickly outdated and encouraging innovation which, as previously noted, may lead to unanticipated sources of vulnerability. Another method that could be marginally successful is to ban attacks on certain categories of targets and to place limits on the overall damage a cyber attack can legitimately achieve during a war, but this approach follows the commencement of hostilities, is voluntary, and does little to promote mutual security in peacetime.[290]

The above analysis demonstrates the hurdles associated with arms control agreements in cyberspace. Nevertheless, this does not preclude alternative means to foster stability. Thus, rather than banning specific capabilities, or seeking an agreement that depends on verification, a more realistic goal would be to create voluntary mechanisms that promote clarity of the domain and enable effective crisis management.

Confidence Building Measures, Stability, and the European Experience

When arms control seems to be a bridge too far between adversaries that hold too many points of disagreement and mistrust, yet both acknowledge the potential for inadvertent conflict,

---

Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on their Destruction (also known as the "Biological Weapons Convention"), respectively.

[290] The implications of cyber warfare on International Humanitarian Law is a well-researched area. For further insight on the application of the law of armed conflict to the cyber domain see David E. Graham, "Threats and the Law of War," *Journal of National Security Law & Policy* 4 (2010): 87-102; and *Tallinn Manual on the International Law Applicable to Cyber Warfare*, ed. Michael N. Schmitt, (Cambridge: Cambridge University Press, 2013).

www.manaraa.com

decision makers have employed confidence building measures in lieu of establishing robust arms control regimes. Like arms control, CBMs may be uni-, bi-, or multilateral agreements. As trust is built between parties they may give way to more formalized arms control agreements because of the role they have in reassuring a potential adversary—though they often do not. By this logic, CBMs are a form of reassurance that seek to demonstrate intentions to a potential adversary, therefore (ideally) conveying a desire to maintain the status quo and foster a sense of security in an otherwise threatened state.[291] Indeed, they are designed to ensure crisis situations, routine tensions, or localized conflicts between states do not become inadvertent lighting rods that spark a general war.[292] Since CBMs are only intended to signal intent behind military activities, they do not change the overall power dynamics between two or more adversaries. Rather, CBMs are simply designed to preserve a fragile stability resulting from intense security competitions between states.

Confidence building measures provide reassurance through three mechanisms. First, they seek to demonstrate non-aggressive postures by increasing transparency of military actions through methods such as inviting designated observers or the public to witness events that otherwise could be construed as threatening.[293] Second, they place self-imposed limits on

---

[291] Jonathan Alford, "Confidence- Building Measures in Europe: The Military Aspects," *Adelphi Papers* 19, no. 149 (1979): 4-13, 5.

[292] Kevin N. Lewis and Mark A. Lorell, *The Utility of Confidence-Building Measures in Crisis Situations: Some Case Studies* (Santa Monica, CA: RAND Corporation, P-6947, January 1984), 2-7.

[293] In this sense, states act in accordance to Schelling and Halperin's "positive-evidence principle" which notes that states are motivated to provide evidence that they are not violating the understanding; Schelling and Halperin, *Strategy and Arms Control*, 97-98. However, providing the level of transparency that can completely mitigate the fears of defection is unlikely in cyberspace due to the necessary secrecy that surrounds these operations. CBMs that rely on inspection for compliance are unrealistic.

www.manaraa.com

security activities, such as military exercises, that could cause another state to feel threatened. Third, CBMs often operate in a time of crisis by enabling a vital communications link between adversaries. In other words, CBMs contribute to stability and détente by helping convey intent behind one's unilateral security policies and actions that would otherwise be cloaked in uncertainty.[294] Furthermore, CBMs also create predictability in a potential adversary's actions and make it easier for another state to detect a deviation from a norm and therefore mitigate the vulnerability of a surprise attack by offering an assurance of early warning.[295] Though CBMs do not replace the vital role of national technical means of intelligence in assessing another actor's capabilities and intent, they supplement it by enabling a fuller picture of the meaning behind a military policy or action than otherwise would have been available.[296] However, CBMs do not decrease mistrust of an adversary or limit its capabilities; that would require an arms control regime that addresses specific security concerns from two or more parties.[297] Though there has been a noted worry amongst scholars and policy makers that CBMs could be used to mask a surprise attack, this was a concern that practitioners overcame during the Cold War due to the

[294] Johan Jørgen Holst and Karen Alette Melander, "European Security and Confidence- Building Measures," *Survival* 19, no. 4 (July/August 1977): 146-154, 148.

[295] Johan Jørgen Holst, "Confidence-Building Measures a Conceptual Framework," *Survival* 25, no. 1 (January/February 1983): 2-15, 2; and Rolf Berg, "Military Confidence-Building in Europe," *Building Security in Europe: Confidence-Building Measures and the CSCE* (New York: East-West Monograph Series 2, 1986): 13-68, 51-54.

[296] Holst, "Confidence- Building Measures a Conceptual Framework," 3.

[297] Richard E. Darilek, "Reducing the Risks of Miscalculation: The Promise of the Helsinki CBMS," *Confidence-building Measures in Europe,* eds. F. Stephen Larrabee, and Dietrich Stobbe (New York: East-West Monograph Number 1, 1983): 59-90, 59.

www.manaraa.com

mutually paramount interest of avoiding inadvertent conflict.[298] One way that states historically have been able to overcome these concerns has been by increasing the degree of confidence that a potential adversary might hold of the intent about a military action or policy by voluntarily shedding light on such matters. In 1979, the representative of the Federal Republic of Germany noted this point before the UN General Assembly,

> A higher degree of confidence can only be achieved when the amount of information which governments command enables them to foresee satisfactorily and to calculate actions and reactions of other Governments within their political environment. In other words, the degree of confidence primarily depends on the degree of openness and transparency with which states are prepared to conduct their political and military affairs.[299]

Though CBMs can be unilaterally implemented, they often take the form of agreements so that all parties can understand the level of transparency necessary for one another to develop confidence in the intent behind another actor's security policy or action.

The Helsinki Final Act

As noted, CBMs need not be formalized in international law or codified in another type of formal agreement to be effective, however they may become institutionalized over time as they evolve from state practice.[300] The touchstone for all confidence building measures was codified in the Final Act of the Conference on Security and Co-operation in Europe that took

[298] See Richard K. Betts, "Hedging Against Surprise Attack," *Survival* 23, no. 4 (1981): 146-156; and Thomas C. Schelling, "Confidence in Crisis," *International Security* 8, no. 4 (Spring 1984): 55-66, 64.

[299] United Nations, General Assembly, *Confidence-Building Measures: Report of the Secretary-General*, A/34/416, October 5, 1979, 19.

[300] Holst and Melander, "European Security and Confidence- Building Measures," 148.

www.manaraa.com

place in Helsinki, Finland in 1975.[301] Broadly speaking, the conference had the goal of creating stability, noting,

> …the need to contribute to reducing the dangers of armed conflict and of misunderstanding of military activities which could give rise to apprehension, particularly in a situation where the participating States lack clear and timely information about the nature of such activities.[302]

The Helsinki Final Act, initially signed by 35 states, sought to create stability by addressing issues that strained East-West relations on topics ranging from sovereignty to freedom of the press and cultural exchanges.[303] Arguably no part of the agreement has been as closely scrutinized as the establishment of confidence building measures between all signatories. To wit, the original act stipulated voluntary reporting with at least a 21-day prior notification of military maneuvers that exceed over 25,000 troops and that occurred within 250 kilometers from a state's

---

[301] In addition to Holst, "Confidence-Building Measures a Conceptual Framework," 2-15, see James Macintosh, "Confidence-Building Measures- A Conceptual Exploration," *Confidence Building Measures and International Security*, eds. R.B. Byers, F. Stephen Larrabee, and Allen Lynch (New York: East-West Monograph Series 4, 1987): 9-29, for a conceptual overview of confidence building measures that a state may choose to enact.

[302] U.S. State Department, *Conference on Security and Co-operation in Europe Final Act*, Department of State Publication 8829, August 1975, 84.

[303] The Helsinki Final Act solely dealt with cooperation between the East and West during the Cold War. However, they have been used as the benchmark for other regional security competitions. Similar agreements can be seen in Ariel E. Levite, and Emily B. Landau, "Confidence and Security Building Measures in the Middle East," *The Journal of Strategic Studies* 20, no. 1 (1997): 143-171; Laurie Nathan, ""With Open Arms" Confidence- and Security- Building Measures in Southern Africa," *South African Journal of International Affairs,* Vol. 1, no. 2 (1994): 110-126; Ralph A. Cossa, *Asia Pacific Confidence and Security Building Measures*, (Center for Strategic & International Studies, 1995); Michael Krepon, Dominique M. McCoy, and Matthew CJ Rudolph, *A Handbook of Confidence-Building Measures for Regional Security*, (The Henry L. Stimson Center, 1993); and United Nations, Department for Disarmament Affairs, *Confidence and Security-Building Measures - From Europe to Other Regions,* (New York: United Nations, 1991).

www.manaraa.com

border.[304] The provision also enabled the exchange of observers for these maneuvers as well as the hosting of military delegations.[305]

The Helsinki Final Act noted that "…the experience gained from the implementation of the provisions…together with further efforts, could lead to developing and enlarging measures aimed at strengthening confidence," and as such created a framework for follow on meetings. The first of these occurred in Belgrade in 1977, followed by Madrid in 1980, Stockholm in 1984, and Vienna in 1986.[306] Each of these conferences comprised multi-year efforts that endeavored to innovate new and creative means to demonstrate intent and promote transparency in response to changing security policies and technology. By the time the 2011 Vienna Document was finalized, CBMs had expanded to include the annual exchange of military information such as organizational charts, manning and equipment numbers, unit locations, defense budgets, and information relating to the employment of new weapon's systems.[307] Furthermore, additional CBMs included the development of more robust communication regimes that could operate in a time of a crisis as well as for routine exchanges of officers and demonstrations of new major weapon systems. The original provisions for troop notifications were also refined to require at

[304] With only one exception between 1975 and 1982, the Soviet Zapad-81 military exercise, all signatories observed the reporting requirements of the original Helsinki Final Act. For a more detailed analysis of the exercises conducted during this period see, Holst, "Confidence- Building Measures a Conceptual Framework," 7-11.

[305] For the exact language for the requirements of this provision see U.S State Department, "Conference on Security and Co-operation in Europe Final Act," 85-86.

[306] For a concise history of CBMs up to the 1992 Vienna document see James Macintosh, "Confidence Building Measures in Europe: 1975 to the Present," *Encyclopedia of Arms Control and Disarmament,* ed. Richard Dean Burns (New York: Charles Scribners Sons, 1993).

[307] The 2011 Vienna document builds on previous agreements, specifically the 1975 Helsinki Final Act, the Document of the Stockholm Conference of 1986, the 1992 Helsinki Document, and the Vienna Documents of 1990, 1992, 1994, and 1999.

www.manaraa.com

least a 42-day warning of exercises of at least 9,000 troops or 250 battle tanks. There were also controls addressing the number of major exercises that a state could perform per year and restrictions on the number of short notice inspections of another signatories' military maneuvers and other troubling sites that a state could annually perform.[308]

In summary, when formalized arms control agreements that seek to change the incentives for military action are not feasible, confidence building measures are an alternative means to mitigate the risk of inadvertent conflict by enabling increased transparency and openness surrounding a state's security policies and operations. Over time these measures may evolve to become formalized arms control regimes as confidence in each other's willingness to adhere to an agreement is manifested. However, changes in security requirements, polices, and technology suggest that for confidence building measures to promote lasting stability they must be reassessed and amended on an iterative basis, as was seen throughout the duration of the Cold War and in the ensuing years.

Confidence Building Measures for the Newest Domain

The current scholarship on cyber CBMs is in a nascent stage. Though multiple scholars have noted their need to avoid inadvertent conflict, few had postulated finite measures that states may implement.[309] Herbert Lin attributes this dearth of measures to the revolutionary nature of the domain. In Lin's words,

---

[308] For a more detailed overview of the 2011 Vienna Document see, Organization for Security and Co-operation in Europe, *Vienna Document 2011 on Confidence- and Security-building Measures*, November 30, 2011, http://www.osce.org/fsc/86597.

[309] For instance, the need for cyber CBMs have been noted by, Jason Healey, "The Five Futures of Cyber Conflict and Cooperation," *Georgetown Journal of International Affairs* (2011): 110-117; and James Andrew Lewis, "Confidence-Building and International Agreement in Cybersecurity," *Disarmament Forum: Confronting Cyberconflict*, vol. 4 (2011).

www.manaraa.com

> Meaningful analogs to… [confidence building] measures in cyberspace are difficult to find. For example, there is no analog to large-scale troop movements—cyber forces can be deployed for attack with few visible indicators. Agreed conventions for behavior, such as "rules of the road," do not cover intent and in cyberspace, intent may be the difference between a possibly prohibited act, such as certain kinds of cyberattack, and an allowed one such as cyber espionage.[310]

Tughral Yamin notes this dilemma, but argues that, "A necessary precondition for developing cyberspace CBMs is to have good national cyber security policies and practices, particularly for the protection of critical infrastructure."[311] Though Yamin fails to quantify the requisite level of policy creation necessary for the effective formation of CBMs, he does make an important contribution by noting that institutional development of cyber security organizations within a state are necessary because of their role in information realization and transference in a domain that is not necessarily readily conceptualized. Indeed, without institutions that assist in information sharing of vulnerabilities, known threats, remediation strategies, and national policies and attitudes for approaching the cyber domain then it is unlikely that actors within and external to a state will understand the risks posed by cyber operations.

Despite academia's lack of extensive research on cyber CBMs, both the United Nations and the Organization for Security and Co-operation in Europe (which sponsored the original Helsinki Final Act), have made efforts towards developing Cyber CBMs. Specifically, the United Nations' Group of Governmental Experts (GGE) on Developments in the Field of Information and Telecommunications in the Context of International Security was convened in

[310] Herbert Lin, "Arms Control in Cyberspace: Challenges and Opportunities," *World Politics Review* (2012): 14-19.

[311] Tughral Yamin, *Cyberspace CBMs between Pakistan and India* (Islamabad: National University of Sciences and Technology, 2014), 102.

www.manaraa.com

2004 to discusses potential areas of cooperation.[312] A year later it failed to reach a consensus and no report was submitted. A second GGE was convened in 2009 and after four meetings over two years it devised the first set of CBMs.[313] This list was expanded by a third and fourth round of GGE panels that concluded in 2013 and 2015, respectively.[314] Relatedly, the Permanent Council of the OSCE directed efforts in 2012 to begin drafting CBMs specific for cyberspace, noting that CBMs were necessary to, "enhance interstate co-operation, transparency, predictability, and stability, and to reduce the risks of misperception, escalation, and conflict that may stem from the use of ICTs."[315] These efforts lead to the drafting of additional CBMs in 2013 and a more comprehensive list in 2016.[316] See Appendix A for a listing of these CBMs.

It should be noted that both the UN and the OSCE lists avoid using the word "cyber" unless in reference to the concept of "cybersecurity." Indeed, they instead rely on "ICT," which

---

[312] United Nations, *General Assembly Resolution 58/32*, A/RES/58/32, December 8, 2003, 47, http://undocs.org/A/RES/58/32.

[313] United Nations, General Assembly, *Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, A/65/201, July 30, 2010, http://undocs.org/A/65/201.

[314] United Nations, General Assembly, *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, A/68/98, June 24, 2013, reissued for technical reasons on July 30 2013, http://undocs.org/A/68/98; and United Nations, General Assembly, *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, A/70/174, July 22, 2015, http://undocs.org/A/70/174.

[315] Organization for Security and Co-operation in Europe, *Permanent Council Decision no. 1039,* April 26, 2012, http://www.osce.org/pc/90169.

[316] Organization for Security and Co-operation in Europe, *Permanent Council Decision no. 1106,* December 3, 2013, http://www.osce.org/pc/109168.; Organization for Security and Co-operation in Europe, *Permanent Council Decision no. 1202,* March 10, 2016, http://www.osce.org/pc/227281. However, also see: USCE Secretariat, Transnational Threat Department, "Cyber/ICT Security," http://www.osce.org/secretariat/256071?download=true., and Lamberto Zannier, "Cyber/ICT Security: Building Confidence," *Security Community- The OSCE Magazine*, Is. 2, 2014, http://www.osce.org/magazine/122525.

www.manaraa.com

is a catch all term for "Information and Communications Technology."[317] In a thoughtful legal review of the measures, Katharina Ziolkowski notes that adopting the ICT label, as opposed to "cyberspace" or "information space" (as some states use in their doctrine), avoids potential political hurdles.[318] Indeed, many states have competing definitions and terminology for the same concepts and by employing the much more neutral ICT moniker the measures were able to avoid opposition at the onset. As noted in Appendix A, OSCE Measure 9 speaks to this point and voluntarily encourages all states to publish their national ICT terminology and calls for all states to build a consensus behind a common lexicon.

Johan Holst argued in 1983 that CBMs come in four varieties: *information*, *notification*, *observation*, and *stabilization* and noted that some measures may encompass several of the categories.[319] *Information* measures involved the sharing of defense related information such as budgets and organizational structures between interested parties. *Notification* pertained to the advanced warning of major military activities within a geographic concentration, such as a military exercise or a major change in force distribution. *Observation* measures included activities such as inviting potential adversaries to physically observe military exercises, the fielding of new weapon systems, or other related military activities first-hand. However, as Holst notes, *stabilization* measures were multifaceted and encompass three dimensions,

> …crisis stability (relative absence of pressures to take early military action to forestall moves by the adversary); arms-race stability (relative absence of inducement to expand

[317] Though not commonly used amongst academics, in keeping with the spirit of the lexicon used the UN and OSCE, this paper will now adopt the "ICT" CBM convention.

[318] Katharina Ziolkowski, "Confidence Building Measures for Cyberspace," *Peacetime Regime for State Activities in Cyberspace. International Law, International Relations and Diplomacy,* (Tallinn: NATO CCD COE Publication, 2013): 533-563, 549-550.

[319] Holst, "Confidence- Building Measures a Conceptual Framework," 4-5.

www.manaraa.com

military forces); and political stability (relative absence of pressures for breakdown of the international order).[320]

However, ICT CBMs are more diverse.[321] Indeed, whereas Holst justly envisioned his *information* categorization as an exchange of defense related data, given the diversity of threat actors in the cyber domain and difficulty of defense, ICT information CBMs have three components. First, the sharing of *threat information* that identifies emerging methods and means for exploitation and attack. This can include the sharing of threat information that relates to specific online personas, country profiles, as well as law enforcement information about non-state actors. Second, *security information* which pertains to the dissemination of system vulnerability reports as well as instructions for remediation. Third, *use information* which includes Holst's conceptualization of the sharing of state-level defense related materials, such as doctrine and national policies, but for ICT is broadened to incorporate private sector use given that they are significant stakeholders in cyberspace. The recognized influence and role of the private sector is evident in both the UNGA and OSCE CBMs that address the sharing of ICT information relating to "national attitudes" and views from both public and private sources.[322] Indeed, there is a recognition in these CBMs that the actors in this space are not solely states and the information about the uses of the cyberspace must extend beyond traditional state actors. Furthermore, several of the measures are designed to explain what norms the state is attempting to establish within its borders, such as desires for a free and open Internet and what ICT infrastructure it considers to be critical. These ICT CBMs are not only designed to avoid

[320] Ibid., 4.

[321] The CBMs in Appendix A are classified by an updated version of John Holst's CBM classification guidelines.

[322] For example, see UNGA Measure 3 and OSCE Measure 7, respectively, in Appendix A.

www.manaraa.com

inadvertent conflict, as was the case with CBMs during the Cold War; several of these CBMs are designed to track norm emergence and evolution. Indeed, the strategic culture with which a state operates in the domain is vital to interpreting intent behind actions relating to and in the cyber domain. This culture affects not only conflict avoidance as states try to create a framework to interpret another actor's intent and actions, but also how to interpret and respond to coercion and conflict escalation emanating from the operational employment of cyber power.[323]

Unlike what was addressed in Holst's framework, several modern-day measures pertain to the maintenance of the CBMs. These *administrative* measures are designed to enable the preservation and continued relevance of the measures as well as the conservation of the respective organization that facilitated its creation. These measures demonstrate advances in the creation of CBMs in the last 30 years since Holst's article. One common element to both the UN and the OSCE list of measures is a reliance on Computer Emergency Readiness Teams (CERTs) for the dissemination of threat and security information. Since the first team was created at

[323] Though beyond the scope of this paper, there are signs of states developing dramatically different strategies for operating militarily in the cyber domain that I argue reflect the strategic culture that exists within those states. For instance, Russian interference in the 2016 US presidential election is in keeping with a common tactic the USSR employed throughout the Cold War (Dov H. Levin, "Partisan Electoral Interventions by the Great Powers: Introducing the PEIG Dataset," *Conflict Management and Peace Science* 33, no. 4, (September 2016)). North Korea has reportedly benefited from financially motivated cybercrime to generate capital and get around the effects of economic sanctions (Paul Mozur and Choe Sang-Hun, "North Korea's Rising Ambition Seen in Bid to Breach Global Banks," *The New York Times*, March 25, 2017, https://www.nytimes.com/2017/03/25/technology/north-korea-hackers-global-banks.html). China has used cyber espionage to steal industrial and intellectual property to grow their economy ("Foreign Spies Stealing US Economic Secrets In Cyberspace," *Office of the National Counter Intelligence Executive*, October 2011). Finally, the United States has historically pressed for the "rules based order" that exists in the physical domains of strategic interaction to apply to cyberspace (for instance see Dan Seifert, "President Obama wants to Prevent a Cyber Weapon 'Arms Race'," *The Verge*, September 5, 2016, https://www.theverge.com/2016/9/5/12798836/president-obama-prevent-cyber-weapon-arms-race, and Hongju Koh, "International Law in Cyberspace," Faculty Scholarship Series, Paper 4854, (2012)).

www.manaraa.com

Carnegie Mellon University in 1989, the concept has expanded to include over 360 teams operating in over 80 countries who mutually promote security cooperation by sharing technical vulnerability and remediation information.[324]

Holst's measures that are not observed in either the UNGA or OSCE ICT measures include: the *notification*, *observation*, and *political* categories. This is not surprising given the prior discussion on obstacles confronting cyber arms control agreements. Indeed, notification of a cyber event or an exercise does not readily make logical sense. However, Paul Meyer, one of the few scholars that attempted to postulate cyber CBMs, called for exchanges of personnel to observe "cyber-security exercises" between potential adversaries.[325] Though Meyer does not clarify what a "cyber-security exercise" is, it could be beneficial to all parties if there was coordination of a cyber defensive exercise if the intent was to build capacity. Indeed, if the exercise demonstrates how an actor intends to respond to and remediate a cyber-attack, inviting allies to observe can help them grow their cyber defensive infrastructure, to include national authorities necessary to respond to a crisis, and help identify points of inject where they could augment efforts and enable a unified cyber defense. However, inviting an adversary to take part may have a different effect. When discussing a "cyber security" exercise the assumption is that this is an entirely defensive mission with all cyber effects being delivered solely on that state's internal network. Indeed, most of these exercises are used to identify both technical and

[324] Forum of Incident Response and Security Teams, "Alphabetical list of FIRST Members," accessed November 14, 2016, https://www.first.org/members/teams.

[325] Paul Meyer, "Cyber-Security through Arms Control: An Approach to International Co-operation," *RUSI Journal* 156, no. 2 (2011): 22-27, 26-27. Note that Yamin, "Cyberspace CBMs between Pakistan and India," 108-112, also takes a rudimentary step towards identifying cyber CBMs that could be established in a bilateral context, however, many of the measures fall more under the convention of norm creation, such as "avoiding hostile propaganda" and have limited applicability beyond the India-Pakistan relationship.

www.manaraa.com

procedural vulnerabilities on internal networks.[326] These exercises do not showcase the units or the capabilities that would conduct offensive operations. Such an exchange is not in keeping with the spirit of the interchange of observers of military exercises as envisioned in the Helsinki Final Act. In other words, the effect that this would have on establishing confidence in both the command and control of and the efficacy of another state's offensive cyber forces is highly dubious. Noting this concern, states could try to build offensive operations into the exercise. It is plausible to imagine building a counterstrike that targets an infected server that is commanding the attack into such a scenario, however, any capability for access and attack that is used will most likely be limited to an already publically available open source tool or will be fictionalized as to not give away to the adversary the specific vulnerability in the target system they are exploiting. For if a state used real cyber weapons from their arsenal, it is plausible that any observing state would develop hardware and software upgrades to render the demonstrated capability inert. Yet, exercising the planning process and the command and control of such capabilities increases the organization's efficiency and control of such operations. However,

[326] For example, *Cyber Guard* is an annual defensive exercise cohosted by the US Cyber Command, the Federal Bureau of Investigation, and the Department of Homeland Security that incorporates the private sector and numerous governmental organizations to respond to a cyber attack against US public and private critical infrastructure. The exercise is focused on identifying existing technical vulnerabilities, developing and testing response actions, and improving defense of DoD information systems. For further information see, Mark Pomerleau, "Cyber Forces Prepare for Attack on a Scale," *Defense Systems*, June 20, 2016, https://defensesystems.com/articles/2016/06/20/cyber-guard-dod-civilian-industry-exercise.aspx. However, US Cyber Command does have exercises, such as their annual *Cyber Flag*, which incorporate offensive capabilities and involve multi-national coalition partners. During this exercise the offensive response actions may be fictionalized or involve real capabilities. However, despite the multinational representation at the exercise, the cyber weaponry used are employed on isolated test networks and only those with appropriate security clearances are aware of the specific capabilities involved. In addition to promoting a common understanding of response actions between partners, the exercise, as is the case with Cyber Guard, is widely used by the military to validate the proficiency of their cyber teams. Discussion with Chief Warrant Officer Judy Esquibel, *Army Cyber Institute*, April 30, 2017.

www.manaraa.com

public notification of the successful *execution* of such exercises could increase the adversaries' confidence in the actor's ability to command and control cyber capabilities. Notification of such security related exercises is a possible avenue for future ICT CBMs, however, any reporting mandate that accounts for both public and private cyber security events could become administratively difficult to manage if the government is not a directly involved party to the exercise.

Furthermore, the lack of measures that seek political stability in cyberspace is not surprising. Though assurances to avoid interfering in the internal affairs of another state occurred during the Cold War, the Internet creates an avenue to undermine regimes that some states have a moral or strategic incentive to take advantage of. Indeed, both authoritarian and democratic regimes view the Internet as a medium to influence not only their own, but each other's citizenry. However, there are sharp divisions between states on their view of both the Internet within their borders which directly informs the Internet norms and policy they press for external to the state.[327] Figure 1 highlights the differences of views of the Internet by regime type which may confound the development of political CBMs across dyads of unlike regime types.

[327] Though beyond the scope of this paper, there is a growing body of literature addressing the differences in Internet policy across regime type. Most authoritarian states have instituted hierarchies in their intrastate Internet infrastructure to prevent their citizens from accessing prohibited material. However, the West (for the most part) has pursued a free and open Internet that is largely devoid of state censorship. These conflicting visions for the Internet were evident during the 2012 breakdown of the United Nation's International Telecommunications Union's World Conference on International Communication (WCIT) when China and Russia used the wake of Arab Spring to get support from many Middle Eastern countries to push for treaty that limited the openness of the Internet and created restriction on free speech. In response, most Western democracies refused to ratify the treaty. This divide has given rise to extensive debates about Internet governance, state sovereignty in cyberspace, and the "Balkanization" of the Internet. See James D. Fielder, "The Internet and Dissent in Authoritarian State," in *Conflict and Cooperation in Cyberspace;* Daniel W. Drezner, "The Global Governance of the Internet: Bringing the State Back In," *Political Science Quarterly* 119, no. 3 (Fall 2004): 477–498; Stephen K. Gourley, "Cyber Sovereignty," in *Conflict and Cooperation in Cyberspace*; Schmitt,

www.manaraa.com

Figure 1: Contrasting Approaches to the Internet by State Type

	View of the Internet *internal* to their borders	View of the Internet *external* to their borders
Authoritarian Regimes	• Internet censorship and monitoring is necessary for state security • States link allowing access to open Internet as undermining regime stability	• There is a need for a rigidly defined concept of cyber sovereignty • Internet affords a means to influence an external actor by spreading defamatory information and propaganda
Democratic Regimes	• Limited censorship across most democratic regimes; most restrictions deal directly with illicit activities • Monitoring of online activity limited by civil liberty protections • Free and open access to the Internet is in keeping with democratic ideals	• Access to a free and open Internet may be a human right • Internet can be used to propagate democratic values and facilitate government transparency and accountability

Most Western actors view access to a free and open Internet as in keeping with their democratic principles, with some actors going so far as to view such access as a human right, and therefore a moral requirement for states to safeguard, as was suggested in a 2011 UN Special Rapporteur report, and reaffirmed in a 2016 UN Resolution that condemned intentional disruption of Internet access by governments.[328] However, despite these reports, many

---

ed., *Tallinn Manual on the International Law Applicable to Cyber Warfare*; Dana Polatin-Reuben and Joss Wright, "An Internet with Brics Characteristics: Data Sovereignty and the Balkanisation of the Internet," (paper presented at the 4th USENIX Workshop on Free and Open Communications on the Internet: San Diego, CA, August 18, 2014). For a comprehensive report on state censorship by country see, Sanja Kelly et al., "Privatizing Censorship, Eroding Privacy - Freedom on the Net 2015," *Freedom House,* October 2015, https://freedomhouse.org/sites/default/files/FOTN%202015%20Full%20Report.pdf.

[328] United Nations, General Assembly, *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, Frank La Rue*, A/HRC/17/27, May 16, 2011, http://undocs.org/A/HRC/17/27; United Nations General Assembly, *The Promotion, Protection and Enjoyment of Human Rights on the Internet, Human Rights Council*, A/HRC/32/L.20, June 27, 2016, http://undocs.org/A/HRC/32/L.20. However, view that access to the Internet is a human right remains hotly contested. For further reference see: Daniel Joyce,

www.manaraa.com

authoritarian states feel that an open Internet encroaches on their sovereignty and is threatening to regime survival.[329] An implication of this is that political CBMs, in line with the definition that Holst put forward, seems unrealistic between dyads of differing state type. Given the potential far-ranging effect online posts can have, a political CBM would most likely attempt to require states to censor their citizen's online posts if the material can be perceived as undermining the ideology or government of another state; an unfathomable obstacle for many Western democracies. Furthermore, many Western states may be reluctant to create political stability CBMs because of a perceived strategic benefit the current ambiguity affords. For democracies this includes the spread of democratic principles and globalization that the Internet enables.[330] This would explain why the United States government invests heavily in the development of anonymity technology through the US State Department's *Bureau of*

"Internet Freedom and Human Rights," *European Journal of International Law* 26, no. 2 (2015): 493-514; Vinton G. Cerf, "Internet Access is Not a Human Right," *The New York Times*, January 4, 2012; and Michael L. Best, "Can the Internet be a Human Right," *Human Rights & Human Welfare* 4, no. 1 (2004): 23-31.

[329] For instance, see Gary King, Jennifer Pan, and Margaret E. Roberts, "How Censorship in China Allows Government Criticism but Silences Collective Expression," *American Political Science Rev*iew 107, no. 2 (2013): 326-343; Chris C. Demchak, and Peter Dombrowski, "Rise of a Cybered Westphalian Age," *Strategic Studies Quarterly* (Spring 2011): 32-61; and Fielder, "The Internet and Dissent in Authoritarian State," in *Conflict and Cooperation in Cyberspace.*

[330] Though beyond the scope of this paper, this view is in line with the literature on the Democratic Peace Theory that holds that democracies are more likely to be peaceful with one another because of shared culture, norms, and structural mechanisms that promote peaceful conflict resolution. Additionally, scholarship has shown that common ideologies encourage alliance formation whereas divergent ideologies can have a positive effect on threat perception and domestic stability. Indeed, authoritarian regimes view access to the unfettered Internet as a venue that can encourages civil unrest because it exposes their citizenry to differing ideological thought and serves as a venue for likeminded individuals to assemble in relative safety. This view is rational given the role social media played during Arab Spring which resulted in multiple regime changes. Some Western policy makers and strategists may hold that by keeping a free and open Internet, democracies can spread democratic values in hopes that it will encourage democratic revolutions in authoritarian regimes. Thus, resulting in a strategic victory and an international environment where they face fewer threats and potentially new allies.

www.manaraa.com

Democracy, Human Rights and Labor, which seeks annual grants for the development of software that contributes to Internet freedom.[331] Furthermore, it would explain why the US government has spent "approximately $2 million annually during the past decade to help enable Internet users in China and other Internet restricting countries to access its websites, such as Voice of America and Radio Free Asia."[332]

Many authoritarian regimes, however, which have pushed for the development of a concept of cyber sovereignty might be hesitant to see them realized. The 2016 US presidential election exposed the effect Russian meddling in a campaign could have. With the public release of hacked emails, to include the falsification of some of the leaked materials, trolling, and the fake news, the US election may have very well been influenced.[333] This same type of activity has occurred and in ongoing in Europe though the full extent of these operations is still being realized.[334] That said, the point is that the Internet affords a means to directly reach the citizenry

---

[331] U.S. State Department, "Bureau of Democracy, Human Rights and Labor Internet Freedom Annual Program Statement," June 2, 2014, http://www.state.gov/j/drl/p/227048.htm. As an interesting side note, the development of the most commonly used anonymity software available, The Onion Router (commonly known as "Tor"), was sponsored by the Defense Advanced Research Projects Agency (DARPA) and US Navy's Office of Naval Research (ONR).

[332] Thomas Lum, Patricia Moloney Figliola, and Matthew C. Weed, "China, Internet Freedom, and U.S. Foreign Policy," *Congressional Research Service*, July 13, 2012.

[333] A comprehensive report on Russian interference into the 2016 US Presidential election has yet to be written. However, given that there are multiple ongoing US government investigations across multiple agencies, as well as investigative journalists that are doing the same, this case is still unfolding. The most comprehensive assessment to date is the following- Office of the Director of National Intelligence- National Intelligence Council, "Assessing Russian Activities and Intentions in Recent US Elections," ICA 2017-01D, January 6, 2017, https://www.dni.gov/files/documents/ICA_2017_01.pdf.

[334] For instance, see Andrew Higgins, "It's France's Turn to Worry About Election Meddling by Russia," *The New York Times*, April 17, 2017, https://www.nytimes.com/2017/04/17/world/europe/french-election-russia.html; Danny Hakim and Christopher F. Schuetze, "Before Elections, Dutch Fear Russian Meddling, but Also U.S. Cash," *The New York Times*, March 7, 2017,

www.manaraa.com

of another state in a way that has not been possible before. Many authoritarian regimes have moved to block this access through censorship and many Western states are struggling with finding a way to block this interference without sacrificing their democratic ideals. Though a political CBM that amounts to a promise to avoid interfering in the internal governance of another state seems like all parties have an incentive to pursue, it is unlikely, particularly the wake of Russian interference in the 2016 US presidential election. Indeed, this influence operation exposed the decisive effect cyber operations can create, and just as arms control agreements are unlikely to take hold when a new technology is perceived to offer decisive effects, a CBM that seeks to block such activities is also unlikely to take hold.

One other notable absence from both the UNGA and OSCE ICT CBM lists is any measure dealing directly with cooperation on cybercrime. Though both lists include sharing of threat information, to include technical signatures and threat actor data, there is no provision that encourages cooperation on the prosecution of cyber criminals or the syncing of domestic cyber-criminal statutes. The framework in Figure 1 provides some insights into why such a provision is conspicuously absent. Most Western states only restrict online activity for illicit activities. Keeping with this view, much of the West has agreed to cooperate on the prevention of cybercrime by becoming signatories to the Budapest Convention on Cybercrime.[335] Where this type of cooperation may be possible between dyads of similar regime type, it is unlikely across non-like dyads. Indeed, the Russian Federation is the only member of the Council of Europe that

---

https://www.nytimes.com/2017/03/07/world/europe/before-elections-dutch-fear-russian-meddling-but-also-us-cash.html.

[335] Council of Europe, "European Treaty Series no. 185- Convention on Cybercrime," (Budapest: November 23, 2001), https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185.

www.manaraa.com

has not signed the treaty.[336] This resistance to cooperation on cybercrime exists for several reasons. First, some authoritarian regimes benefit from being directly or indirectly complicit in online illicit activity as they receive the monetary proceeds or intellectual property gained from cybercrime. Indeed, North Korea has been said to net millions of dollars from cybercrime to get around the crippling effects of economic sanctions.[337] Second, the sheltering of cyber criminals enables the state to have a future cyber proxy actor that can conduct plausibly deniable cyber operations at the behest of the state.[338] Though Western actors could receive the same benefits from cybercrime, most of these states have domestic statutes that limit their citizenry and the government from engaging directly in this type of activity.

ICT CBMs Beyond Europe

In practice, the most notable ICT CBMs outside Europe take the form of bilateral agreements. Specifically, the bilateral CBMs between the United States, China, and Russia. With respect to China and the US, some progress has been made in developing mechanisms that promote transparency and cooperation both during peacetime as well as during a crisis. In 2015 Presidents Obama and Xi signed an agreement to abstain from cyber-enabled intellectual property theft that enabled a commercial competitive advantage, to exchange vulnerability and law enforcement information, and to create a working group to further discuss the UN Group of

[336] "Chart of Signatures and Ratifications of Treaty 185," *Council of Europe*, Status as of June 6, 2017, https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185/signatures. However, it is important to note that in April 2017 Russia put forward their own convention on cyber crime to the United Nations as a direct replacement to the Budapest Convention. For further information see "Russia Prepares New UN Anti-cybercrime Convention- Report," *RT*, April 14, 2017, https://www.rt.com/politics/384728-russia-has-prepared-new-international/.

[337] Mozur and Sang-Hun, "North Korea's Rising Ambition Seen in Bid to Breach Global Banks."

[338] For an in-depth discussion on the motivation of authoritarian states to employ cyber proxies see Erica D. Borghard, and Shawn W. Lonergan, "Can States Calculate the Risks of Using Cyber Proxies?" *Orbis* 60, no. 3 (Summer 2016): 395-416.

www.manaraa.com

Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security 2015 Report.[339] Though this agreement did not specifically create mechanisms to avoid inadvertent conflict or contribute to crisis stability, it did clarify how each state intends to use the domain. Furthermore, it made it apparent when the other state was defecting from a pattern of compliance and potentially engaging in revisionist behavior.

In a separate landmark agreement, in 2014 the US DoD and the Chinese PLA allowed the exchange of observers for major military activities created a military crisis notification system utilizing the Defense Telephone Link between the two countries that was first established in 2008.[340] Though neither document mentioned the word "cyber" or "ICT," it was understood at the signing that the driving catalyst was uncertainty stemming from the potential for inadvertent cyber conflict and escalation during a crisis.[341]

In June 2013, the United States and Russia created a working group within the context of the Bilateral Presidential Commission that sought to "promote transparency and reduce the possibility that an incident related to the use of ICTs could unintentionally cause instability or

---

[339] The White House, "Fact Sheet: President Xi Jinping's State Visit to the United States," Office of the Press Secretary, September 25, 2015, https://www.whitehouse.gov/the-press-office/2015/09/25/fact-sheet-president-xi-jinpings-state-visit-united-states. Note that intellectual property theft that supports national security objectives is still permissible for espionage for it considered a necessary state practice by Customary International Law.

[340] "Memorandum of Understanding Between the United States of America Department of Defense and the People's Republic of China Ministry of National Defense on Notification of Major Military Activities Confidence-Building Measures Mechanism," November 4, 2014, http://archive.defense.gov/pubs/141112_MemorandumOfUnderstandingOnNotification.pdf. "Military Crisis Notification Mechanism for Use of the Defense Telephone Link," September 2015, http://www.defense.gov/portals/1/documents/pubs/us-china_crisis_communications_annex_sep_2015.pdf.

[341] Discussion with a senior official at the United States Department of Homeland Security, July 14, 2016.

www.manaraa.com

escalation."[342] Though the United States suspended its participation in the Bilateral Commission following Russia's invasion of Ukraine, the agreement mentioned three measures of note.[343] First, the continuous sharing of ICT threat information between the US CERT located at the Department of Homeland Security and its Russian equivalent. Second, an agreement to utilize the Nuclear Risk Reduction Center (NRRC), first established in 1987, to facilitate inquiries about cybersecurity incidents. In the closing days of the 2016 presidential election, it was reported that the United States used the NRRC to compel Russia to avoid interfering with US voting systems.[344] What is unique about this case is not that the hotline was used, but that it was used for deterrence rather than for détente. Finally, the commission also created a direct line between the White House's Cybersecurity Coordinator and the Kremlin's Deputy Secretary of the Security Council integrated into the Direct Secure Communications System which, like the NRRC, was first developed to manage nuclear crises during the Cold War. Though these bilateral agreements are examples between major state powers, there are other agreements between regional powers, particularly on cyber security coordination efforts, that are beyond the scope of this paper. That said, these examples highlight confidence building measures that have been taken to avoid escalation and prevent inadvertent conflict between countries that are potential adversaries, thus promoting both peacetime and crisis stability.

---

[342] U.S. State Department, *U.S.-Russia Bilateral Presidential Commission: 2013 Joint Annual Report*, (2013), http://www.state.gov/p/eur/ci/rs/usrussiabilat/219086.htm#8.

[343] The White House, "Fact Sheet: U.S.-Russian Cooperation on Information and Communications Technology Security," Office of the Press Secretary, June 17, 2013, https://www.whitehouse.gov/the-press-office/2013/06/17/fact-sheet-us-russian-cooperation-information-and-communications-technol.

[344] David Ignatius, "In our new Cold War, deterrence should come before détente," *The Washington Post*, November 15, 2016.

www.manaraa.com

Conclusion

Cyber operations can be destabilizing to interstate relations and lead to increased tensions and inadvertent conflict because their secretive nature promotes uncertainty between actors. However, states have an incentive to avoid unwanted tensions, particularly when it runs the risk of escalating into a general war. Therefore, confidence building measures that facilitate a dialogue between states is a first step toward mitigating the destabilizing effects posed by the cyber domain. The existing Computer Emergency Response Team infrastructure, which enables a common threat picture for state and non-state actors alike, reduces some of the uncertainty and vulnerability that exists between actors in this volatile space. However, the CERTs do not help manage crisis escalation nor do they promote transparency beyond the technical threat and security information that they freely share. To those ends, states have utilized, to differing degrees, both bilateral and multilateral agreements to create mechanisms to share information about their intended uses of cyberspace, law enforcement information concerning nefarious actors, as well as to share information in a crisis.

Though it is impossible to eliminate the incentives for actors to misrepresent or disguise their aggressive cyber actions, an incentive exists to avoid misunderstandings that can escalate to inadvertent conflict. Considering this research, a CBM that seeks a voluntary commitment between two or more states that they will avoid developing new accesses and enduring backdoors on systems within the sovereign territory of the other state or any allies it is obligated to defend during a time of crisis may be warranted. In other words, during a crisis, states should avoid hacking another state's critical systems, because gaining access to them for intelligence collection runs the risk of being misinterpreted as gaining a foothold through which a cyber attack can be launched. Such a measure would address the acute risk these types of operations

www.manaraa.com

pose for inadvertent conflict creation and escalation. As previously discussed, an intelligence gathering cyber operation may be seen by the recipient as a precursor to a cyber attack, particularly during a period of increased hostilities.

A second step would be a commitment to limit the number of non-state offensive cyber actors in the domain. The traditional move would be a CBM that amounts to a commitment for states to keep the operational capabilities in the hands of the military, but oversight and launch authority in the hands of policymakers, as many states have done with nuclear arms. However, the lack of complete military control of offensive cyber capabilities is already commonplace as many states must rely on civilian industry for expertise and development, thus rendering the consolidation of offensive cyber power by the military a costly and often unrealistic proposition. This occurs for several reasons. First, the line between governmental and non-governmental actors is often blurred, with some states operating parastatals that depend heavily on cyber espionage for economic growth. Therefore, limiting non-state actors that engage in cyber espionage and offensive operations may not be acceptable to some states due to economic concerns. Second, some states lack robust indigenous cyber capabilities, personnel, and the resources to produce them and are thus forced to employ cyber proxies to fulfill national security objectives.[345] Further adoption of international institutions that seek to standardize laws between states for prosecution of cyber-crime and other types of nefarious cyber related activity, as has been seen in the Budapest Convention, are a good first step, but as already noted, have had limited traction amongst states outside the West.

[345] In addition to hiring cyber proxies due to a dearth of localized talent, states may also employ them for plausible deniability. For further reference to the state-proxy exchange see Borghard, and Lonergan, "Can States Calculate the Risks of Using Cyber Proxies." Note that the authors make a strong argument for the risks states face when they chose to employ a cyber proxy.

www.manaraa.com

In a related logic, limiting not only actors, but also the proliferation of technology is an admirable goal. However, many of these tools are publically available via online forums or for sale via the Dark Web, a section of the Internet that is accessible via most web browsers and is known to facilitate illicit transactions.[346] For instance, the source code for Stuxnet as well as NSA capabilities for surpassing firewalls have been compromised and a tech savvy actor can learn to morph them into something more advanced.[347] Second, efforts have been made to control the export of ICT technology that can support offensive operations using The Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies. However, the 2013 amendments which tried to apply this to cyber quickly met industry opposition because the technology that supports offensive operations is also necessary to discover vulnerabilities that are in need of being patched, thus highlighting the offensive and defensive dual use nature of many cyber security tools.[348] Indeed, this provision triggered significant resistance from private industry which felt it would inevitably lead to greater insecurity by placing restrictions on cyber security-related technology and activities, such as penetration testing technology, the sharing of threat information, and the use of multinational

[346] For further information on the Dark Web marketplace see Paul Stockton and Michele Golabek-Goldman, "Curbing the Market for Cyber Weapons," *Yale Law & Policy Review* 32, no. 1 (2013): 239-266.

[347] For instance, see *The Hacker News*, "Stuxnet Source Code Released Online- Download Now," July 2, 2011, http://thehackernews.com/2011/07/stuxnet-source-code-released-online.html; and Ellen Nakashima, "Powerful NSA hacking tools have been revealed online," *The Washington Post*, August 16, 2016.

[348] See testimony on "Wassenaar: Cybersecurity and Export Control," U.S. House of Representatives- Subcommittee on Information Technology, January 12, 2016, https://oversight.house.gov/hearing/wassenaar-cybersecurity-and-export-control/; and Trey Herr, "Malware Counter-Proliferation and the Wassenaar Arrangement," Proceedings of the 8th International Conference on Cyber Conflict, January 4, 2016.

www.manaraa.com

computer bug bounty programs.[349] To date, the provisions of ICT technology on cyber security capabilities of the Wassenaar Arrangement are still being refined both collectively by the Wassenaar Plenary and by member countries as they nest domestic regulation with their obligations under the Arrangement. For instance, in response to the public feedback, the specific 2013 Wassenaar amendments that covered the training and employment of vulnerability detection systems were never implemented in the United States.[350]

Finally, care should be given to craft a measure that speaks not only to how states intend to employ cyber power to achieve objectives, but also the delegation of authorities that each state mandates for the approval of various types of cyber operations. This would assist in understanding what organizations and individuals are behind specific operations, thus adding clarity to attribution efforts. Furthermore, such a measure would assist in building confidence between states that these operations are maintained through a rigid command and control structure.

The further creation of ICT confidence building measures and the continued maintenance of those already in existence is a necessary step towards mitigating the security dilemmas and

---

[349] Russell Brandom, "Google says controversial exports proposal would make the world 'less secure,'" *The Verge*, July 20, 2015, http://www.theverge.com/2015/7/20/9005351/google-wassenaar-arrangement-proposal-comments; and Chris Bream, "Wassenaar rules are not the right direction," *Facebook U.S. Public Policy*, July 28, 2015, https://www.facebook.com/uspublicpolicy/posts/1047027321981746. For a more comprehensive list of the public feedback the US Department of Commerce received regarding this regulation see the nearly 1000-page report titled- "Public comments for Wassenaar Arrangement 2013 Plenary Agreements Implementation: Intrusion and Surveillance Items proposed rule (published May 20, 2015)," https://efoia.bis.doc.gov/index.php/documents/public-comments/public-comments-2015/1027-bis-2015-0011-comment-report/file.

[350] Discussion with a senior official in the US Department of Commerce- Bureau of Industry and Security- November 28, 2016. Specifically, Wassenaar Arrangement Category 4 rules: 4.A.5, 4.D.4, and 4.E.1.C, were never added to the Commerce Control List as elucidated in the Department's Export Administration Regulations (EAR) in either 2014 or 2015 following the 2013 amendments.

www.manaraa.com

inadvertent risks that states confront in cyberspace. Despite traditional arms control regimes being unrealistic and ill-suited for controlling cyber operations, an incentive exists for states to press for the further adoption of ICT confidence building measures that mitigate uncertainty and facilitate crisis management, thereby promoting stability between states.

www.manaraa.com

APPENDIX A

THE UNGA RECOMMENDED THE FOLLOWING CBMS ON JULY 22, 2015	CLASSIFICATION OF CBM
1. The identification of appropriate points of contact at the policy and technical levels to address serious ICT incidents and the creation of a directory of such contacts;	*Stability- Crisis*
2. The development of and support for mechanisms and processes for bilateral, regional, subregional and multilateral consultations, as appropriate, to enhance inter-state confidence-building and to reduce the risk of misperception, escalation and conflict that may stem from ICT incidents;	*Information- Use*
3. Encouraging, on a voluntary basis, transparency at the bilateral, subregional, regional and multilateral levels, as appropriate, to increase confidence and inform future work. This could include the voluntary sharing of national views and information on various aspects of national and transnational threats to and in the use of ICTs; vulnerabilities and identified harmful hidden functions in ICT products; best practices for ICT security; confidence-building measures developed in regional and multilateral forums; and national organizations, strategies, policies and programmes relevant to ICT security;	*Information- Use* *Information- Threat* *Information- Security*
4. The voluntary provision by states of their national views of categories of infrastructure that they consider critical and national efforts to protect them, including information on national laws and policies for the protection of data and ICT-enabled infrastructure. States should seek to facilitate cross-border cooperation to address critical infrastructure vulnerabilities that transcend national borders. These measures could include:	*Information- Use*
– A repository of national laws and policies for the protection of data and ICT-enabled infrastructure and the publication of materials deemed appropriate for distribution on these national laws and policies;	
–The development of mechanisms and processes for bilateral, subregional, regional and multilateral consultations on the protection of ICT-enabled critical infrastructure;	
– The development on a bilateral, subregional, regional and multilateral basis of technical, legal and diplomatic mechanisms to address ICT-related requests;	
– The adoption of voluntary national arrangements to classify ICT incidents in terms of the scale and seriousness of the incident, for the purpose of facilitating the exchange of information on incidents.	

www.manaraa.com

THE UNGA RECOMMENDED THE ADDITIONAL CBMS ON BILATERAL, SUBREGIONAL, REGIONAL AND MULTILATERAL BASIS	CLASSIFICATION OF CBM
A. Strengthen cooperative mechanisms between relevant agencies to address ICT security incidents and develop additional technical, legal and diplomatic mechanisms to address ICT infrastructure-related requests, including the consideration of exchanges of personnel in areas such as incident response and law enforcement, as appropriate, and encouraging exchanges between research and academic institutions;	*Information-Security*
B. Enhance cooperation, including the development of focal points for the exchange of information on malicious ICT use and the provision of assistance in investigations;	*Information-Security*
C. Establish a national computer emergency response team and/or cybersecurity incident response team or officially designate an organization to fulfil this role. States may wish to consider such bodies within their definition of critical infrastructure. States should support and facilitate the functioning of and cooperation among such national response teams and other authorized bodies;	*Information-Threat* *Information-Security* *Stability- Crisis*
D. Expand and support practices in computer emergency response team and cybersecurity incident response team cooperation, as appropriate, such as information exchange about vulnerabilities, attack patterns and best practices for mitigating attacks, including coordinating responses, organizing exercises, supporting the handling of ICT-related incidents and enhancing regional and sector-based cooperation;	*Information-Threat* *Information-Security* *Stability- Crisis*
E. Cooperate, in a manner consistent with national and international law, with requests from other States in investigating ICT-related crime or the use of ICTs for terrorist purposes or to mitigate malicious ICT activity emanating from their territory.	*Information- Threat* *Information-Security*

www.manaraa.com

THE FOLLOWING CBMS WERE ADOPTED THROUGH OSCE PERMANENT COUNCIL DECISION NO. 1106 ON DECEMBER 3, 2013	CLASSIFICATION OF CBM
1. Participating States will voluntarily provide their national views on various aspects of national and transnational threats to and in the use of ICTs. The extent of such information will be determined by the providing Parties.	*Information- Security*
2. Participating States will voluntarily facilitate co-operation among the competent national bodies and exchange of information in relation with security of and in the use of ICTs.	*Information- Threat*
3. Participating States will on a voluntary basis and at the appropriate level hold consultations in order to reduce the risks of misperception, and of possible emergence of political or military tension or conflict that may stem from the use of ICTs, and to protect critical national and international ICT infrastructures including their integrity.	*Stability- Arms-race*
4. Participating States will voluntarily share information on measures that they have taken to ensure an open, interoperable, secure, and reliable Internet.	*Information- Use*
5. The participating States will use the OSCE as a platform for dialogue, exchange of best practices, awareness-raising and information on capacity-building regarding security of and in the use of ICTs, including effective responses to related threats. The participating States will explore further developing the OSCE role in this regard.	*Administrative*
6. Participating States are encouraged to have in place modern and effective national legislation to facilitate on a voluntary basis bilateral co-operation and effective, time-sensitive information exchange between competent authorities, including law enforcement agencies, of the participating States in order to counter terrorist or criminal use of ICTs. The OSCE participating States agree that the OSCE shall not duplicate the efforts of existing law enforcement channels.	*Information- Use*
7. Participating States will voluntarily share information on their national organization; strategies; policies and programmes – including on co-operation between the public and the private sector; relevant to the security of and in the use of ICTs; the extent to be determined by the providing parties.	*Information- Use*
8. Participating States will nominate a contact point to facilitate pertinent communications and dialogue on security of and in the use of ICTs. Participating States will voluntarily provide contact data of existing official national structures that manage ICT-related incidents and co-ordinate responses to enable a direct dialogue and to facilitate interaction among responsible national bodies and experts. Participating States will update contact information annually and notify changes no later than thirty days	*Stability- Crisis* *Information- Use*

www.manaraa.com

after a change has occurred. Participating States will voluntarily establish measures to ensure rapid communication at policy levels of authority, to permit concerns to be raised at the national security level.	
9. In order to reduce the risk of misunderstandings in the absence of agreed terminology and to further a continuing dialogue, participating states will, as a first step, voluntarily provide a list of national terminology related to security of and in the use of ICTs accompanied by an explanation or definition of each term. Each participating state will voluntarily select those terms it deems most relevant for sharing. In the longer term, participating states will endeavor to produce a consensus glossary.	*Information- Use*
10. Participating states will voluntarily exchange views using OSCE platforms and mechanisms inter alia, the OSCE communications network, maintained by the OSCE Secretariat's conflict prevention centre, subject to the relevant OSCE decision, to facilitate communications regarding the CBMs.	*Administrative*
11. Participating states will, at the level of designated national experts, meet at least three times each year, within the framework of the security committee and its informal working group established by Permanent Council Decision No. 1039 to discuss information exchanged and explore appropriate development of CBMs. Candidates for future consideration by the IWG may include inter alia proposals from the consolidated list circulated by the chairmanship of the IWG under PC.DEL/682/12 on 9 July 2012, subject to discussion and consensus agreement prior to adoption.	*Administrative*

www.manaraa.com

THE FOLLOWING CBMS WERE ADOPTED THROUGH OSCE PERMANENT COUNCIL DECISION NO. 1202 ON MARCH 10, 2016	CLASSIFICATION OF CBM
12. Participating States will, on a voluntary basis, share information and facilitate inter-State exchanges in different formats, including workshops, seminars, and roundtables, including on the regional and/or subregional level; this is to investigate the spectrum of co-operative measures as well as other processes and mechanisms that could enable participating States to reduce the risk of conflict stemming from the use of ICTs. Such activities should be aimed at preventing conflicts stemming from the use of ICTs and at maintaining peaceful use of ICTs.	*Information-Security*
With respect to such activities participating States are encouraged, inter alia, to:	
– Conduct such activities in the spirit of enhancing inter-State co-operation, transparency, predictability and stability;	
– Complement, through such activities, UN efforts and avoid duplicating work done by other fora; and	
– Take into account the needs and requirements of participating States taking part in such activities.	
Participating States are encouraged to invite and engage representatives of the private sector, academia, centres of excellence and civil society in such activities.	
13. Participating States will, on a voluntary basis, conduct activities for officials and experts to support the facilitation of authorized and protected communication channels to prevent and reduce the risks of misperception, escalation, and conflict; and to clarify technical, legal and diplomatic mechanisms to address ICT-related requests. This does not exclude the use of the channels of communication mentioned in Permanent Council Decision No. 1106.	*Stability- Arms-race* *Stability- Crisis*
14. Participating states will, on a voluntary basis and consistent with national legislation, promote public-private partnerships and develop mechanisms to exchange best practices of responses to common security challenges stemming from the use of ICTs.	*Information-Security*
15. Participating states, on a voluntary basis, will encourage, facilitate and/or participate in regional and subregional collaboration between legally-authorized authorities responsible for securing critical infrastructures to discuss opportunities and address challenges to national as well as trans-border ICT networks, upon which such critical infrastructure relies.	*Information-Security* *Information- Threat*
Collaboration may, inter alia, include:	
– Sharing information on ICT threats;	
– Exchanging best practices;	
– Developing, where appropriate, shared responses to common	

www.manaraa.com

challenges including crisis management procedures in case of widespread or transnational disruption of ICT-enabled critical infrastructure;	
– Adopting voluntary national arrangements to classify ICT incidents in terms of the scale and seriousness of the incident;	
– Sharing national views of categories of ICT-enabled infrastructure states consider critical;	
– Improving the security of national and transnational ICT-enabled critical infrastructure including their integrity at the regional and subregional levels; and	
– Raising awareness about the importance of protecting industrial control systems and about issues related to their ICT-related security, and the necessity of developing processes and mechanisms to respond to those issues.	
16. Participating states will, on a voluntary basis, encourage responsible reporting of vulnerabilities affecting the security of and in the use of ICTs and share associated information on available remedies to such vulnerabilities, including with relevant segments of the ICT business and industry, with the goal of increasing co-operation and transparency within the OSCE region. OSCE participating states agree that such information exchange, when occurring between states, should use appropriately authorized and protected communication channels, including the contact points designated in line with CBM 8 of Permanent Council Decision No. 1106, with a view to avoiding duplication.	*Information-Security*

www.manaraa.com

SECURITY STUDIES
2017, VOL. 26, NO. 3, 452–481
http://dx.doi.org/10.1080/09636412.2017.1306396

Routledge
Taylor & Francis Group

The Logic of Coercion in Cyberspace

Erica D. Borghard and Shawn W. Lonergan

ABSTRACT

What are the dynamics of coercion in cyberspace? Can states use cyber means as independent tools of coercion to influence the behavior of adversaries? This article critically assesses traditional coercion theory in light of cyberspace's emergence as a domain in which states use force, or its threat, to achieve political objectives. First, we review the core tenets of coercion theory and identify the requisites of successful coercion: clearly communicated threats; a cost–benefit calculus; credibility; and reassurance. We subsequently explore the extent to which each of these is feasible for or applicable to the cyber domain, highlighting how the dynamics of coercion in cyberspace mimic versus diverge from traditional domains of warfare. We demonstrate that cyber power alone has limited effectiveness as a tool of coercion, although it has significant utility when coupled with other elements of national power. Second, this article assesses the viability and effectiveness of six prominent warfighting strategies in the traditional coercion literature as applied to the cyber domain: attrition, denial, decapitation, intimidation, punishment, and risk. We conclude that, based on the current technological state of the field, states are only likely to achieve desired objectives employing attrition, denial, or decapitation strategies. Our analysis also has unique implications for the conduct of warfare in cyberspace. Perhaps counterintuitively, the obstacles to coercion that our analysis identifies may prompt states to reevaluate norms against targeting civilian infrastructure.

Cyberspace has definitively emerged as the latest frontier of militarized interactions between nation-states. Governments, as they are wont to do in an anarchic international system, have already invested considerable resources to develop offensive and defensive military capabilities in cyberspace. It remains to be seen, however, how and to what extent these tools can be employed to achieve desired political objectives. Put simply, what is the logic of coercion in cyberspace? Can governments use cyber power to deter state adversaries from taking undesirable actions

Erica D. Borghard is an assistant professor in the Department of Social Sciences and Executive Director of the Grand Strategy Program at the United States Military Academy at West Point. Shawn W. Lonergan is an assistant professor and research scientist at the Army Cyber Institute at the United States Military Academy at West Point.

© 2017 Taylor & Francis Group, LLC

المنارة للاستشارات

www.m

or compel them to bend to their wills and, if so, how and under what conditions?[1] This analysis draws on the large corpus of coercion theory to assess the extent to which existing frameworks can shed light on the dynamics of coercion in cyberspace. The article proceeds as follows. First, we outline the theoretical logic of coercion theory and identify the factors necessary for successful coercion. Each element of coercion is immediately followed by a discussion of how it applies to the cyber domain and an assessment of how the particularities of the domain reflect on the requirements of successful coercion. We demonstrate that, based on current capabilities, cyber power has limited effectiveness as an independent tool of coercion. Second, we explore the extent to which cyber power could be used as part of a warfighting strategy to target an adversary's ability or willingness to resist and suggest which strategies are likely to be more versus less effective.[2] We assert that, based on current capabilities, attrition, denial, and decapitation strategies are most likely to be effective in cyberspace. Finally, we conclude with recommendations for policymaking and further research.

Coercion Theory

As Thomas C. Schelling so eloquently articulated, coercion is fundamentally about affecting an adversary's behavior using the threat or limited application of military force; "[i]t is the *threat* of damage, or of more damage to come, that can make someone yield or comply."[3] Coercion involves producing a desired behavior or outcome on the part of an adversary by forcing her to confront a cost–benefit calculus, such that the adversary believes it is less costly to concede to the threatener's preferred course of (in)action than to defy the latter's demands.[4] Coercion is distinct from brute force. In the latter case, one state defeats another militarily and then imposes a political settlement on the defeated power; in the former case, the target of coercion retains the military capacity to resist or concede, and the coercer seeks to achieve a political settlement short of full-scale

[1]Thomas C. Schelling makes the important distinction between compellence and deterrence. The former involves the threat or limited application of force to change an adversary's behavior, while the latter involves the threat of force (or pain, in Schelling's parlance), to preserve the status quo. See Thomas C. Schelling, *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1960) and idem., *Arms and Influence* (New Haven, CT: Yale University Press, 2008), 69–86. Robert J. Art elaborates on this concept. See Robert J. Art, "To What Ends Military Power?" *International Security* 4, no. 4 (Spring 1980): 3–35.

[2]The authors are grateful to Jack Snyder for pointing out the distinction between coercion and warfighting strategies.

[3]Schelling, *Arms and Influence*, 3. Emphasis in the original. Alexander L. George et. al. also emphasize that coercion can involve both the threat or limited application of military power. See Alexander L. George, David K. Hall, and William R. Simons, eds., *The Limits of Coercive Diplomacy: Laos, Cuba, Vietnam* (Boston: Little, Brown and Company, 1971), 2, 18. Lawrence Freedman distinguishes between coercion, as defined by Schelling, and "strategic coercion," which is "the deliberate and purposive use of overt threats to influence another's strategic choices." See Lawrence Freedman, "Strategic Coercion," in *Strategic Coercion: Concepts and Cases*, ed. Lawrence Freedman (Oxford: Oxford University Press, 1998), 15.

[4]Robert A. Pape, *Bombing to Win: Air Power and Coercion in War* (Ithaca, NY: Cornell University Press, 1996), 4. It is important to note, however, that Pape's reference to coercion in this context is distinct from deterrence; Schelling uses the umbrella term "coercion" to refer to both compellence and deterrence. See also Daniel L. Byman and Matthew Waxman, *The Dynamics of Coercion: American Foreign Policy and the Limits of Military Might* (Cambridge: Cambridge University Press, 2002), 3.

www.m

war by manipulating the cost–benefit calculus of the target state.[5] While coercion has always been a fundamental element of the exercise of state power, the advent of nuclear weapons and mutual assured destruction has made coercion even more critical. As Schelling explains, the prospect of civilization-ending nuclear warfare, coupled with advances in technology making it possible to target an enemy's population centers and hold its society at risk without first defeating its armed forces, has turned statecraft into the diplomacy of violence.[6] The significance of coercion for interstate relations has not decreased with the advent of cyber warfare; if anything, it has increased. Indeed, like nuclear weapons, cyber weapons enable governments to target adversary populations while bypassing the latter's military forces. For example, cyber weapons could be employed to target a state's critical infrastructure to render key pieces of a state's military systems inoperable at decisive times, as was allegedly the case when Syrian air defense systems failed to respond to an Israeli bombing operation against a purported Syrian nuclear enrichment facility in 2007.[7]

Notwithstanding the central role coercion plays in states' strategies, successful coercion—both its deterrent and compellent varieties—is difficult to achieve.[8] There is a large body of empirical literature that assesses the reasons for failed coercion, particularly focusing on examples of failed coercion in American foreign policy during the Vietnam War and through the use of air power in the post-Cold War international system.[9] In general, using the threat or limited application of military force to affect an adversary's behavior is difficult to accomplish because there are many factors that are necessary conditions for successful coercion, some of which are in tension with others. Moreover, if coercion is difficult to achieve through the threat or use of conventional military power, we argue that it is even more challenging in cyberspace. The literature on coercion suggests that four fundamental conditions must be met for coercion to succeed: the coercive threat must be clearly communicated; it must be linked to a cost–benefit calculus such that the

[5]Schelling, *Arms and Influence*, 2–6. Pape, *Bombing to Win*, 13.

[6]Schelling, *Arms and Influence*, 18–34. Schelling links technological advances in the power to hurt with the increased "importance of war and threats of war as techniques of influence, not of destruction; of coercion and deterrence, not of conquest and defense; of bargaining and intimidation," 33.

[7]Richard A. Clarke and Robert K. Knake, *Cyber War: The Next Threat to National Security and What to Do About It* (New York: Ecco, 2012), 1–8.

[8]It is widely accepted that deterrence may be easier to achieve, but harder for social scientists to observe due its negative object (for example, we only observe deterrence failures). Conversely, compellence is easy to observe but more difficult to achieve for precisely the same reason—there are reputational costs associated with being seen to back down and concede to an adversary's demands. Leaders who are successfully deterred could point to a variety of reasons they chose to not alter the status quo without losing face. See the discussion in Schelling, *Arms and Influence*, 74–75; Robert J. Art, "Coercive Diplomacy: What Do We Know?" in *The United States and Coercive Diplomacy*, eds. Robert J. Art and Patrick M. Cronin (Washington, DC: United States Institute of Peace Press, 2003), 361–62.

[9]See, for example, Pape, *Bombing to Win*; Todd S. Sechser, "Goliath's Curse: Coercive Threats and Asymmetric Power, *International Organization* 64, no. 4 (October 2010): 627–60; Wallace J. Thies, *When Governments Collide: Coercion and Diplomacy in the Vietnam Conflict, 1964–1968* (Berkeley: University of California Press, 1980); Alexander L. George, *Forceful Persuasion: Coercive Diplomacy as an Alternative to War* (Washington, D.C.: United States Institute of Peace Press, 1991); Art, "Coercive Diplomacy;" Alexander L. George and William E. Simons, eds., *The Limits of Coercive Diplomacy*, 2nd ed. (Boulder, CO: Westview Press, 1990); Byman and Waxman, *Dynamics of Coercion*; Thomas J. Christensen, *Worse Than a Monolith: Alliance Politics and Problems of Coercive Diplomacy in Asia* (Princeton, NJ: Princeton University Press, 2011).

www.m

target's costs of conceding are less than the costs of not complying; it must be credible; and there must be an element of reassurance.[10]

Communication

The essence of successful coercion is clear communication.[11] The target of a coercive threat has to know precisely the behavior in which the coercing state wants the target state to engage (or refrain from engaging), the timeframe in which the coercing state expects the target to comply, and the costs associated with cooperation versus defection. The target state must understand "what behavior of his will cause the violence to be inflicted and what will cause it to be withheld."[12] Ideally, coercion takes the form of an ultimatum: if State B does not do action X within timeframe Y, State A will take specified action Z. However, in the vast majority of international crises, political leaders default to ambiguity, rather than clarity, of threats; leaders often prefer to retain flexibility to escape from costly or imprudent commitments or be adaptive in their responses to an adversary's behavior, especially if they lack domestic political support.[13] The fundamental fact of anarchy complicates clear communication because it leads to poor, fragmentary information and creates incentives to misrepresent private information—indeed, this is a cause of war.[14] Beyond incentives for strategic ambiguity, clear signaling is complicated by misperceptions stemming from both cultural differences and cognitive limitations.[15] Insights from cognitive psychology have demonstrated that recipients of a signal tend to fit incoming information into preexisting beliefs, interpret signals based on implicit theories about their meaning, prefer simplicity over complexity, and are influenced by motivated biases.[16] Put simply, signaling often fails "because the perceiver does not understand what message the actor is trying to communicate."[17] Communication is

[10]Of course, this is not an exhaustive list of all of the factors that contribute to successful coercion. For example, George and Simons identify nine conditions that favor coercive diplomacy: clarity of objective, strong motivation, asymmetry of motivation, sense of urgency, strong leadership, domestic support, international support, fear of unacceptable escalation, and clarity of terms. See George and Simons, eds., *Limits of Coercive Diplomacy*, 279–91. However, we propose that these various lists and factors could be grouped into the four main conditions identified above.

[11]However, it is important to note a caveat that, in some instances, sending ambiguous signals can be advantageous for the purposes of coercion. Particularly in the context of nuclear bargaining, the threat that leaves something to chance—precisely because the risk of nuclear war generates extraordinary costs—may help a coercing state. See Schelling, *Strategy of Conflict*, chap. 8.

[12]Idem., *Arms and Influence*, 3–4.

[13]Jack Snyder and Erica D. Borghard, "The Cost of Empty Threats: A Penny, Not a Pound," *American Political Science Review* 105, no. 3 (August 2011): 429; Robert Jervis, "Deterrence Theory Revisited," *World Politics* 31, no. 2 (January 1979): 303; Richard Ned Lebow, *Between Peace and War: The Nature of International Crises* (Baltimore, MD: Johns Hopkins University Press, 1981), 29–27; Glen H. Synder and Paul Diesing, *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises* (Princeton, NJ: Princeton University Press, 1977), 213–15, 220. Even Schelling acknowledges that "most commitments are ultimately ambiguous in detail," *Arms and Influence*, 67.

[14]Freedman, "Strategic Coercion," 18; James D. Fearon, "Rationalist Explanations for War," *International Organization* 49, no. 3 (Summer 1995): 379–414.

[15]Robert Jervis, *Perception and Misperception in International Politics* (Princeton, NJ: Princeton University Press, 1976).

[16]Ibid., 117–202. Robert Jervis, "Signaling and Perception: Drawing Inferences and Projecting Images," in *Political Psychology*, ed. Kristen Renwick Monroe (Mahwah, NJ: Lawrence Erlbaum Associates, 2002), 306–8. See also Robert Jervis, *The Logic of Images in International Relations* (New York: Columbia University Press, 1970).

[17]Jervis, "Signaling and Perception," 304.

www.m

facilitated when actors can agree on a shared meaning of a particular type or vehicle of signaling (such as diplomatic language). In the case of diplomatic language, for example, clarity is easier to achieve because "both the signaler and the perceiver agree as to the message that the former is trying to convey."[18]

Communication in Cyberspace

Understanding intent is exceptionally difficult in the cyber domain. Many scholars and US government-sponsored studies have noted that cyber operations create a high probability of misunderstanding the coercing state's intentions.[19] Unlike diplomatic channels, in cyberspace there is no agreed-upon language that guides policymakers to a common understanding that helps divine the meaning behind a cyber signal. Moreover, in cyberspace, most operations are interactions between humans and machines facilitated by code for which there are few, if any, norms governing the exchange.[20] That many high-level decision makers lack even a basic understanding of the cyber domain and, therefore, are likely to be intellectually unprepared during a time of crisis, compounds this uncertainty. Furthermore, the signaler may be uncertain about what kind of cyber tool she should select to communicate in cyberspace because the actual effects of a cyber attack may be unpredictable ex ante—even to the signaler.[21]

Signaling in cyberspace is the most problematic of all the domains (land, sea, air, space, and cyber) because the signal may go unrealized. In other words, in cyberspace only the initiator may perceive the engagement.[22] Moreover, even if a target state realizes it has been attacked, it is difficult to infer the intent behind a cyber signal based solely on an observed incursion. This ambiguity has the potential to trigger unintended escalation because it is difficult to distinguish between hostile and benign intentions when an outside actor is perceived to have accessed a critical system.[23] In a

[18]Ibid., 300.

[19]For further reference, see Andru E. Wall, "Demystifying the Title 10-Title 50 Debate: Distinguishing Military Operations, Intelligence Activities & Covert Action," *Harvard National Security Journal* 3 (December 2011): 85–142; Robert Chesney, "Military-Intelligence Convergence and the Law of the Title 10/Title 50 Debate," *Journal of National Security Law and Policy* 5 (October 2011): 539–629; William A. Owens, Kenneth W. Dam, and Herbert S. Lin, eds., *Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities* (Washington, DC: National Academies Press, 2009).

[20]For more on norms and international law as they pertain to the cyber domain, see Catherine Lotrionte, "A Better Defense: Examining the United States' New Norms-Based Approach to Cyber Deterrence," *Georgetown Journal of International Affairs* 8, no. 10 (April 2014): 75–88; Martha Finnemore, "Cultivating International Cyber Norms," in *America's Cyber Future: Security and Prosperity in the Information Age*, ed. Kristin M. Lord and Travis Sharp (Washington, DC: Center for a New American Security, 2011); Tim Maurer, "Cyber Norm Emergence at the United Nations—An Analysis of the UN's Activities Regarding Cyber-Security," *Belfer Center for Science and International Affairs* (September 2011); Oona A. Hathaway, et al., "The Law of Cyber-Attack," *California Law Review* 100, no. 4 (August 2012): 817–85; David E. Graham, "Cyber Threats and the Law of War," *Journal of National Security Law and Policy* 4 (2010): 87–102; Jack Goldsmith, "How Cyber Changes the Laws of War," *European Journal of International Law* 24, no. 1 (2013): 129–38.

[21]The authors are grateful to Robert Jervis for illustrating this.

[22]For instance, it is easy to imagine how a single signal could get lost in the over eighty-eight thousand petabytes of IP traffic that are estimated to transverse the Internet per month.

[23]For further discussion of risks surrounding the ambiguity of intent in cyberspace, see Shawn W. Lonergan, "Cooperation under the Cybersecurity Dilemma," in *Confronting Inequality: Wealth, Rights, and Power*, ed. Hugh Liebert, Thomas Sherlock, and Cole Pinheiro (New York: Sloan, 2016). Also see Robert Jervis, "Some Thoughts on Deterrence in the Cyber Era," *Journal of Information Warfare* (forthcoming): 8–9.

www.m

hypothetical example, Japan may have an intelligence requirement to monitor the uranium enrichment efforts of North Korea. However, Japan's access to a network at a North Korean enrichment facility does not necessarily suggest that it intends to destroy North Korea's nuclear ambitions through cyber means; Japan could simply be monitoring the program to meet its own defensive requirements, which is widely accepted by international convention to be a necessary state practice.[24] Actors could exploit this uncertainty to their advantage, but it may also lead to unintended conflict.[25] Herbert Lin notes this ambiguity in cyberspace and concludes that the cyber domain presents an increased risk of accidental escalation: "In the absence of direct contact with those conducting such operations—sometimes even in the presence of such contact—determining intent is likely to be difficult and may rest heavily on inferences made on the basis of whatever attribution is possible. Thus, attempts to send signals to an adversary through limited and constrained military actions—problematic even in kinetic warfare—are likely to be even more problematic when cyber attacks are involved."[26]

Attribution problems complicate effective communication in cyberspace because they create problems for both target and initiator. From the perspective of the target state, a fundamental impediment to deciphering the intent behind a cyber signal is the difficulty of identifying the actor who sent it. This presents a challenge to policymakers because, if a cyber action is uncovered, the true meaning of the signal may not be ascertained without attribution. While some actions in themselves may send a clear signal without attribution, typically the identity of the signaling state is critical for coercion to succeed. For instance, in the prior scenario we assumed North Korea attributed the cyber incursion to Japan. However, what if North Korea were unable to attribute the access to Japan and had to surmise the intent of the incursion devoid of attribution? The spectrum of possible motivations of such an incursion ranges from a preparation for a preemptive attack from a rival such as Japan or the United States on one end of the spectrum, to a benign case of espionage from an ally such as China on the other. In this hypothetical case, not only is the signal obfuscated because intent cannot be deduced without attribution, but North Korea also does not know what is an appropriate response and against whom to respond. If these conditions are not met, coercion is by definition not possible.

There are, however, several methods to assign attribution following a cyber incursion.[27] The easiest ascription approach is when the perpetrator publically accepts responsibility for the action and the target state believes that the self-identified attacker possessed both the capability and motivation to carry it out.

[24]Geoffrey B. Demarest, "Espionage in International Law," *Denver Journal of International Law and Policy* 24 (1995): 321–48.
[25]Jervis, *Logic of Images in International Relations*, 86–87.
[26]Herbert Lin, "Escalation Dynamics and Conflict Termination in Cyberspace," *Strategic Studies Quarterly* 47 (2012): 57.
[27]For further reference on attribution, see Thomas Rid and Ben Buchanan, "Attributing Cyber Attacks," *Journal of Strategic Studies* 38, no. 1–2 (2015): 4–37; Jon R. Lindsay, "Tipping the Scales: The Attribution Problem and the Feasibility of Deterrence against Cyberattack," *Journal of Cybersecurity* 1, no. 1 (2015): 1–15.

www.m

Another attribution technique mandates that the target state had access to the attacker's network from which the incursion originated and either witnessed the operation in real time or recorded it. This second method is difficult because it requires that the target state had access to the specific network from which the aggressor initiated an attack; that they observed the onslaught developing in real time and intentionally refrained from establishing tailored defenses or engaging in a preemptive attack to block the assault; or that that they had complete intelligence collection of all cyber operations from the adversary's network, which is typically technically difficult to consistently collect. However, in some instances governments may decide that the intelligence value of maintaining access outweighs the likely damage from the attack. Another attribution method is when sensors placed either at Internet service providers or key nodes in the Internet run algorithms that analyze raw data flows and scan for anomalies and variants of known attack signatures. However, the real-time use of such technology is still in a nascent stage and there is currently no guarantee that, once detected, the source of the malware could be traced back to the true originator.[28] The final method of assigning attribution occurs when the signature of the attack (the coding) is so unique that it could be traced to a specific actor or threat network. Yet, this method heightens the risk of falling victim to deceptive techniques, such as embedding remarks in a foreign language of a noninvolved party, which may confound forensic experts seeking to assign attribution. Recently, however, there have been advances in signature recognition software designed to scour millions of lines of code to compile unique profiles of the developers.[29] Moreover, from the target's perspective, even if she is able to successfully attribute an attack to a particular actor, she may be hesitant to reveal her ability to do so because it would likely require going public with valuable information that could compromise her own capabilities and accesses. For instance, the United States' decision to quickly attribute the Sony hack in late 2014 to North Korea likely revealed and compromised American accesses to other governments' cyber infrastructure.[30]

[28]Gerhard Munz and Georg Carle, "Real-Time Analysis of Flow Data for Network Attack Detection" (paper presented at the 10th IFIP/IEEE International Symposium on Integrated Network Management, Munich, Germany, 21 May 2007).

[29]Developing unique signatures of attackers and code developers is becoming more common as exploits become increasingly sophisticated and threat data is shared among cyber security practitioners. Stuxnet, for instance, was nearly fifteen thousand lines of code that comprised over five hundred thousand bytes. That is the equivalent amount of digital data as a large textbook. Governments have been keen to invest in digital forensic technology, as evident in then Secretary of Defense Leon Panetta's 11 October 2012 address from the deck of USS *Intrepid*, where he noted: "The department has made significant advances in solving a problem that makes deterring cyber adversaries more complex: the difficulty of identifying the origins of that attack. Over the last two years, [the] DoD has made significant investments in forensics to address this problem of attribution and we're seeing the returns on that investment. Potential aggressors should be aware that the United States has the capacity to locate them and to hold them accountable for their actions that may try to harm America," see Leon Panetta, "Remarks by Secretary Panetta on Cybersecurity to the Business Executives for National Security, New York City," 11 October 2012, http://archive.defense.gov/transcripts/transcript.aspx?transcriptid=5136. For a review of Panetta's speech, see Jack Goldsmith, "The Significance of Panetta's Cyber Speech and the Persistent Difficulty of Deterring Cyberattacks," *Lawfare*, 15 October 2012.

[30]The authors thank Robert Jervis for this comment.

www.m

Attribution issues create problems not only for the target state attempting to infer the intent behind a signal, but also for the coercing state seeking to send a clear signal. The conventional wisdom on cyber operations posits that states typically seek to avoid attribution when conducting cyber exploitation and espionage operations. However, coercion in cyberspace requires attribution to be effective. A coercing state may employ several methods to ensure attribution. First, a state could couple the action in cyberspace with a formal diplomatic message, elucidating the meaning the signal (the cyber attack) was intended to convey.[31] Coupling a cyber operation with a diplomatic message may be the least costly method to ensure ascription for the coercer, but it must also be credible. This technique was observed in March 2016 when Secretary of Defense Ashton Carter, in a formal public statement, acknowledged that the United States conducted a cyber attack against the Islamic State of Iraq and Syria's command and control systems in Mosul, Iraq.[32] However, a coercing state must ensure that the target believes its self-declared attribution. This could present a problem for the coercing state if, in order to demonstrate that it was the one sending a signal, it had to reveal capabilities and accesses that it may prefer to keep private. Second, if coupling is not an available avenue, some have postulated several technical methods to ensure attribution, such as embedding unique signatures in code.[33] This type of ascription technique demands that some trace of the cyber operation remain on the target's machines.

This suggests that simply gaining access to a network and conducting cyber espionage is not sufficient to send a coercive signal in cyberspace—even if such accesses may be necessary to support a coercive signal.[34] While much of the discussion in the public domain conflates cyber espionage and cyber military operations, these are in fact distinct, just as they are in conventional domains. All forms of espionage, whether conducted in cyberspace or elsewhere, are fundamentally about collecting private information against another actor. Conversely, to be coercive, a cyber signal must be attributable and aim to disrupt, deny, degrade, and/or destroy data resident on computers and computer networks, or the systems themselves.

[31] Jervis, *The Logic of Images in International Relations*, 139–44.

[32] Damian Paletta and Felicia Schwartz, "Pentagon Deploys Cyberweapons against Islamic State," *Wall Street Journal*, 29 February 2016.

[33] Goldsmith, "Panetta's Cyber Speech."

[34] This, of course, creates something of a paradox for a coercing state because it may need to gain prior access to a system or network (which requires obfuscation and avoiding attribution) to send a subsequently attributable coercive signal. The one caveat to this is that cyber espionage could be used to conduct a data breach of sensitive information that can later be released to embarrass or otherwise intimidate some actor. Though cyber espionage may be a complex operation, depending on how well defended the network or computer the targeted information was resident on is, it is not a costly coercive cyber signal. Rather, cyber is being used as a vehicle to acquire information. It is not being employed as a signaling mechanism in itself. Due to this, the 2015 Sony hack is not a coercive cyber operation because it did not seek to destroy or disrupt and systems, but rather was used to steal embarrassing insider information. In this instance, if North Korea were to use the information gleaned from its alleged cyber attack on Sony's systems to coerce the company into refraining from releasing a movie, its cyber espionage would constitute a tool that was the component of a broader coercion strategy, but the fact that North Korea breached Sony's networks to steal information was not in itself sending a costly signal to the company. Cyber espionage is a routine aspect of interstate interactions in cyberspace and, in itself, does not meet the threshold of a costly signal. It is possible that cyber espionage activities could be coupled with other costly signals but, independently, it is not one.

www.m

Cost–Benefit Calculus

Coercion theory assumes that states are rational actors who make cost–benefit calculations when determining how to respond to threats and inducements posed by other actors in the international system. The benefit side of the calculus involves how much the adversary values a particular course of action, while the cost side entails the price she anticipates paying in order to carry it out.[35] Coercion, put simply, forces the target state to choose between "making concessions or suffering the consequences."[36] Therefore, to be effective, a coercing state must issue a threat such that the target perceives it to be more costly to suffer those consequences than to concede.[37] To succeed, the coercer must know what the target state values and, therefore, what it can hold at risk to get the target to comply; or, in Schelling's parlance, "[c]oercion requires finding a bargain, arranging for him to be better off doing what we want—worse off not doing what we want—when he takes the threatened penalty into account."[38] More important than an objective measure of costs versus benefits, however, is how the adversary perceives them, which stems from "the magnitude of the dangers and profits the adversary sees ahead for a given path and the probability of their occurrence."[39] At its core, therefore, coercion is the manipulation of the target's perceptions of the cost–benefit balance of a particular course of action.[40]

Affecting an adversary's cost–benefit calculus may seem deceptively simple; in practice, it could fail across multiple dimensions. Coercion could fail because the target does not understand what the adversary values and, therefore, does not know how to appropriately tilt the cost–benefit calculus. This could stem from poor intelligence or, more fundamentally, from the fact that leaders are not always rational, utility-maximizing economic individuals. It may be difficult to quantify what a target state values if it involves something intangible (such as prestige) and, therefore, hard to assign a numerical value to the cost a coercer must threaten to impose to achieve a desired behavior. Relatedly, coercion could fail because states are not unitary actors and, therefore, there may be domestic political or bureaucratic organizational considerations that factor into what a target state values, how it perceives costs versus benefits, and acceptable levels of risk that the coercing state does not take into account. Moreover, even if the coercer knew what the adversary values, it could be politically difficult to make a sufficiently costly threat. Finally, coercion could fail due to cognitive limitations on the part of the target. Insights gleaned from prospect theory, for instance, have illustrated that individuals often fail to make rational, cost–benefit calculations when assessing risk (such as being more averse to losses than gains) and misunderstand sunk costs.[41]

[35]Byman and Waxman, *Dynamics of Coercion*, 11.
[36]Pape, *Bombing to Win*, 12.
[37]Byman and Waxman, *Dynamics of Coercion*, 10. Pape, *Bombing to Win*, 15–16.
[38]Schelling, *Arms and Influence*, 4.
[39]Byman and Waxman, *Dynamics of Coercion*, 11.
[40]George, *Forceful Persuasion*, 11–14. George also points out that the more expansive or extreme the demands of the coercing state are, the costlier the threat must be to secure compliance.
[41]Byman and Waxman, *Dynamics of Coercion*, 10–14. Schelling, *Arms and Influence*, 86. Jack S. Levy, "Prospect Theory, Rational Choice, and International Relations," *International Studies Quarterly* 41, no. 1 (March 1997): 87–112.

المنارة للاستشارات
www.m

Cost–Benefit Calculus in Cyberspace

In cyberspace, the state issuing a coercive threat must calculate what target to go after and the effect it seeks to deliver against it. Similarly, the target must also calculate whether it can absorb the cost and, if so, whether the coercer can ratchet up the cost to the target while avoiding too much cost itself. There are several categories of targets a state may consider attacking in cyberspace to coerce another state. Generally speaking, the class of target that inflicts the highest level of cost, a state's critical infrastructure, is typically the hardest to gain access to due to the technical complexities stemming from custom, tailored uses and advanced physical and virtual defensive measures commonly emplaced around these vital capabilities. The United States Department of Homeland Security has defined these crucial nodes as" … systems and assets, whether physical or virtual, so vital that the incapacity or destruction of such may have a debilitating impact on the security, economy, public health or safety, environment, or any combination of these matters."[42] This category includes critical infrastructure that is essential for everything from the safeguarding of nuclear regulatory systems to gas pipelines, and control systems that enable communication systems to work.[43] In the United States, many of these critical systems are ran by private industry, but in states with parastatal enterprises (such as China), they remain centrally controlled by the government. Not all pieces of critical infrastructure, however, are universally valued across states. For instance, diverging state opinions over the ideal relationship between the citizen and the Internet has changed what states may consider critical infrastructure. Indeed, one accomplished Chinese academic with senior-level party connections noted to the authors that their "Great Fire Wall," which restricts citizen access to Western media sources, is considered part of China's critical infrastructure.[44] In this case, attacking a vital node that the state links to regime stability would be significantly costlier for China than the destruction of other types of critical infrastructure. Similarly, the recent hack of the US Democratic National Committee, allegedly committed by Russia or Russian-sponsored groups, is an example of how a state could target a critical component of a democratic regime—its electoral system.[45] This creates the potential for unintended escalation dynamics if the coercing state did not accurately calculate the extent to which the target values its electoral process.

Military capabilities may also be targeted by cyber attacks. These targets include everything from software running on advanced avionic platforms, to air defense

[42]"National Infrastructure Protection Plan: Partnering to Enhance Protection and Resiliency," *Department of Homeland Security*, 2009, https://www.dhs.gov/xlibrary/assets/NIPP_Plan.pdf.
[43]Control Systems are defined as, "Computer-based systems used within many infrastructure and industries to monitor and control sensitive processes and physical functions. These systems typically collect measurement and operational data from the field, process and display the information, and relay control commands to local or remote equipment or human-machine interfaces (operators)," ibid., 109.
[44]Professor at Peking University, Beijing, China, in discussion with the authors, 17 June 2015.
[45]David E. Sanger and Eric Schmitt, "Spy Agency Consensus Grows That Russia Hacked D.N.C.," *New York Times*, 26 July 2016.

www.m

assets, communication systems, and satellites tied into the Global Positioning System (GPS). Setting one's cyber sights on these military systems is similar in terms of costliness to targeting civilian critical infrastructure in that both are custom engineered and are typically difficult to gain access to and, therefore, mandate a highly tailored capability to exploit. Additionally, given that military systems are designed to be used during times of conflict, they tend to be more secure than civilian infrastructure because they are created with the expectation that they may be attacked via cyber means and, therefore, there is a greater emphasis placed on survivability and resilience early on in the development cycle.

From the target's perspective, attacks against critical national infrastructure and military capabilities are the most costly types of attacks, precisely because governments rely on these to survive in the international system and perform their basic functions. Coercing states may also choose to target the corporate sector of another state, depending on permissibility allowed by its own domestic legal regimes. Targets could include the online banking ability of a particular bank, the network of a leading defense contractor, or consumer information held by retailers. There is variation in terms of the cost to a coercer of targeting a particular company or sector of the economy, and this variation is largely a function of the resiliency and defenses that private actors choose to incorporate into their networks and systems. However, in terms of the perceived cost to the target state, generally speaking, cyber attacks against a private company are of a lower magnitude than attacks against critical national infrastructure and military capability, including command and control capabilities. Therefore, these kinds of attacks would only be useful to coerce a target state into conceding on relatively minor issues, if at all. This is analogous to conventional domains—dropping ordnance on a Walmart is fundamentally different from dropping ordnance on a communications node. However, there are two important caveats to this analysis. First, there may be some reputational costs a target may incur if attacks against certain private sectors actors are perceived to undermine the legitimacy of the regime. Second, there is likely to be important variation stemming from regime type, because some kleptocratic states may rely on the support of key industries or even companies to maintain regime stability. In these cases, attacks against business or industry may be comparable in terms of perceived cost to attacks against critical national infrastructure.

Regardless of the nature of the target, when sending a coercive signal in cyberspace, a policymaker must decide if she wants to produce a disruptive or destructive effect, the most salient distinction in the domain. Thus, a policymaker employing cyber attacks as a coercive instrument of state power must make a calculation of what effect is necessary to achieve the desired outcome. Destructive cyber attacks take two forms: the rare cyber attacks that generate an effect felt in the physical world, and the more common destruction of digital information, which can be almost as dire as a physical attack for many pieces of infrastructure. Disruptive attacks, conversely, seek to operationally diminish a system to the point that a user lacks confidence in its ability to perform some function. The latter may be more appealing to a coercing government

المنارة للاستشارات
www.m

because disruptive attacks enable functionality of the affected system to be restored once the attack is ceased and, thus, may aid in reassuring the target state, as will be discussed in a later section. Notwithstanding the above discussion, states may be unable to perfectly tailor a cyber signal to affect a target's cost–benefit calculus. In other words, the technical complexities of certain types of costly operations may force less capable states into sending less costly signals that don't sufficiently alter the target's calculations. Governments may find cheap, fast, and easy cyber operations appealing even when they are less effective for the purposes of coercion. Put simply, governments may hit what they can get, rather than the optimal target to coerce another state.

Credibility

Beyond being costly, a coercer's threat must be credible—the target must believe that the coercer will actually carry it out. A threat is credible if it is in a state's interests to carry it out and if that state has both the capability and the resolve, or political will, to do so.[46] Credibility is arguably one of the most difficult aspects of coercion, which is why Schelling devotes an entire chapter of *Arms and Influence* to "the threats that are hard to make, the ones that are not so inherently credible."[47] While it is challenging to assess a coercing state's ability to carry out a threat, it is even more difficult to discern and demonstrate resolve.[48] Furthermore, in games of chicken, testing resolve through the limited application of force only tends to harden resolve even more, because the very act of probing resolve engages a state's political legitimacy and reputation, making leaders more likely to dig in before they give in.[49] A target may doubt a coercer's resolve because it doesn't believe that it is in the latter's interests to carry out the threat (this was particularly important in the context of nuclear deterrence); or because it doubts that the leader has sufficient domestic political support to carry out the threat;[50] or because the coercer has not established a reputation for carrying out past threats.[51] How individuals actually assess credibility, however, is poorly understood.[52]

Because credibility is difficult to convey but essential for coercion, states attempt to enhance the credibility of their threats by making them costly—through sending costly

[46]Schelling, *Arms and Influence*, 36.
[47]Ibid. Of course, one of the reasons Schelling found credibility so confounding was the problem of coercion in the nuclear age, where carrying out a threat would mean immeasurable costs to oneself as well as one's adversary.
[48]Stephen Biddle, for example, discusses how assessing the raw, quantifiable capabilities of states' militaries is a poor predictor of battlefield outcomes because it does not take into account force employment. See Stephen Biddle, *Military Power: Explaining Victory and Defeat in Modern Battle* (Princeton, NJ: Princeton University Press, 2006). For difficulties ascertaining resolve, see Art, "Coercive Diplomacy," 365.
[49]Ibid., 365–66. See also Snyder and Diesing, *Conflict Among Nations*, 118–22.
[50]Randall L. Schweller, *Unanswered Threats: Political Constraints on the Balance of Power* (Princeton, NJ: Princeton University Press, 2008).
[51]Schelling asserts that a country's image—others' expectations about how it is likely to behave—is "one of the few things worth fighting for." This is due to what Schelling describes as the interdependence of threats. See Schelling, *Arms and Influence*, 124, 55–59. Also see Herman Kahn, *On Thermonuclear War* (Princeton, NJ: Princeton University Press, 1960), 566. For a critique of the importance of having a reputation for resolve, see Jonathan Mercer, *Reputation and International Politics* (Ithaca, NY: Cornell University Press, 1996).
[52]Robert Jervis, "Deterrence and Perception," *International Security* 7, no. 3 (Winter 1982–1983): 9.

www.m

signals. James D. Fearon asserts that, "to be credible, a threat must have some cost or risk attached to it that might discourage an unresolved state from making it."[53] That's because talk is cheap: "words are cheap, not inherently credible when they emanate from an adversary, and sometimes too intimate a mode of expression."[54] There are two mechanisms states can employ to generate costly and, therefore, credible signals. First, states can tie their hands, limiting their choices and increasing the costs of backing down in the event the target of coercion does not comply with the terms of the threat. Second, states can sink costs, taking actions that are costly up front, such as mobilizing troops.[55] Using a similar framework, Robert Jervis describes how states can use indices to generate costly signals. Indices are "behaviors (either verbal or nonverbal) that the perceiver believes are inextricably linked to a characteristic that helps predict what the actor will do in the future."[56] Democracies, it has been argued, may have an advantage in costly signaling because they can more easily tie their hands through incurring audience costs.[57] Generating costly signals does not come without risks—indeed, costly signaling, paradoxically, is designed to increase the risk of war through locking in coercers to the use of force in order to (hopefully) avoid it.[58] Furthermore, there are myriad reasons states may seek to avoid a perfectly committing threat through sending an unambiguous costly signal, as previously noted.

Credibility in Cyberspace

As the coercion literature has elucidated, making a credible threat requires both the capability to impose the threatened cost and the will to employ it if the party to be influenced does not comply with the issuer's demands. Credibility in cyberspace could be established via two mechanisms. First, establishing indices could create a venue for states to better communicate and demonstrate capability. However, indices of cyber power do not yet exist and are likely to be difficult to form. Therefore, at present, credibility is most likely to be inferred through costly signaling.

Cyber Power Indices

Establishing indices of cyber power contributes to the credibility of threats in cyberspace because it helps ascertain a state's capability.[59] Perfect information of another state's cyber capabilities does not exist; therefore, indices facilitate a state's

[53]James D. Fearon, "Signaling Foreign Policy Interests: Tying Hands Versus Sinking Costs," *Journal of Conflict Resolution* 41, no. 1 (February 1997): 69.
[54]Schelling, *Arms and Influence*, 150.
[55]Fearon, "Signaling Foreign Policy Interests," 70. Schelling also refers to these dynamics in his discussion of commitment through the use of bridge burning, trip wire forces, plate glass windows, and engaging a nation's honor and prestige through public commitments. Schelling, *Arms and Influence*, 44–49.
[56]Jervis, "Signaling and Perception," 300.
[57]The idea that democracies have an advantage in costly signaling has been the conventional wisdom in the literature, although Jessica L. Weeks argues that autocratic regimes are also capable of generating audience costs. Jessica L. Weeks, "Autocratic Audience Costs: Regime Type and Signaling Resolve," *International Organization* 62, no. 1 (Winter 2008): 35–64. For a different critique of audience costs logic, see Snyder and Borghard, "The Cost of Empty Threats."
[58]Fearon, "Signaling Foreign Policy Interests," 82–83.
[59]Robert Jervis, *The Logic of Images in International Relations*, 26–28.

www.m

assessment of another state's ability to carry out threats. In cyberspace, these indices include budgets, growing and training cyber forces, establishing commands, and advertising participation in major cyber exercises.[60] When assessing capabilities in cyberspace, it is also critical to analyze how the latter would be employed. In particular, states in this domain may feel less constrained by international laws and norms (or even the threat of assured retaliation because, as this analysis demonstrates, these threats are difficult to credibly convey). This is because actors in the cyber domain tend to prefer to obfuscate their identities, leading some state actors to be more willing to act in ways that they would not otherwise be willing to on a battlefield or via formal diplomatic channels.

Estimating the capability of a cyberspace actor is a conundrum that has challenged scholars because the opaque nature of the domain confounds measurement efforts.[61] In the nuclear and chemical warfare arenas, there are methods to estimate the stockpiles of arms a nation holds and for which there exist treaties, accords, and international oversight institutions that monitor and limit the quantities of these weapons. However, in the cyber world there is no measure of relative strength; one cannot simply count the number of cyber tools the way one can count the number of warheads or the pounds of poison gas a country possesses. This is because offensive cyber capabilities are not universally lethal. A shroud of secrecy surrounds a nation-state's cyber capability and, therefore, creates a situation of imperfect information from which a policymaker must judge another state's actions and intent. Unlike in the conventional or nuclear realms, where states can reveal their capabilities to bolster credibility (or where the technology necessitates public tests to assess their effectiveness, such as nuclear tests), in the cyber realm states typically prefer to—and can—keep capabilities secret because revealing them would enable adversaries to defend against them and render the capabilities impotent. In other words, it is harder for states to reveal private information in cyberspace to enhance the credibility of their threats.[62] Moreover, governments face unique difficulties deriving intent based on observed capabilities because many states in the cyber domain find themselves coercing with the weapons they have, rather than the ones they may want or need. In other words, there may be a large gap between capabilities and intent. A distinction should be made here between what a state can measure about its own cyber capabilities and what its adversaries can assess. Measuring a rival's military strength has always been more difficult than introspective assessment due to military secrecy. However, the difference in cyberspace is that self-assessments of cyber capabilities (at least currently) also happen to be much harder to conduct because effective metrics have yet to be

[60]It general sense, it may be easier for democracies to showcase their level of cyber power due to greater institutionalized transparency over military organizations and budgets compared to authoritarian regimes.

[61]For example, see H. J. Seo, Yoon-Cheol Choy, and SoonJa Hong, "A Study on the Methodology to Evaluate the Level of Nation's Capability for Cyber War" (paper presented at the 12th Annual International Workshop on Information Security Applications, Korea, August 2011).

[62]The authors are grateful to Robert Jervis for clarifying this point.

www.m

devised. This, in turn, makes assessing another state's cyber-military might even more difficult than for other domains and types of weapons.

Furthermore, measures of cyber power include factors beyond raw estimates of the size of cyber forces. While human capital and skill levels are important contributors to capability in the conventional domain, they are arguably even more vital in the cyber domain. Simply counting the number of cyber forces that a country may openly report as an assessment of cyber power does not take into account the differences in skill levels and a state's relative depth of cyber operations. A lack of homogeneity of material resources and technically proficient human capital across states means that one cannot precisely compare cyber capabilities between states. Comparing quantities of cyber forces is akin to comparing quantities of ships in a navy without distinguishing between tugboats and aircraft carriers. Regime type also factors into capabilities. Some states, such as Russia and China, place a greater emphasis on developing cyber forces to monitor their citizenry to detect unrest and preserve regime stability. From a technical standpoint, these operations are markedly different from conducting a destructive cyber attack against a state adversary. Democratic states have the advantage of devoting fewer cyber resources to population monitoring and, therefore, are freer to invest in adversary-centric capabilities.[63] Finally, what matters for capability in cyberspace is having the right operator, armed with the right capability, with access to a vulnerable target, rather than a numerical advantage. Capability and access imply that, regardless of how skilled an individual operator is, she will always be constrained by the cyber tools with which she has been equipped.

Cyber Operations as Costly Signals

In order to bolster the credibility of a threat, states often engage in costly signaling that ranges from national leaders' threats and troop mobilizations, to onshore trip wires, to the movement of aircraft carriers during times of crises.[64] All of these serve

[63]Authoritarian states have gone to extensive efforts to institute hierarchies in their Internet infrastructure so that they can keep their citizens from accessing material that they deem may threaten regime stability. However, the West has pursued a free and open Internet that is largely devoid of state censorship. These conflicting visions for the Internet was evident in the 2012 breakdown of the United Nation's International Telecommunications Union's World Conference on International Communication (WCIT) when, in the wake of Arab Spring, many Middle Eastern states joined a voting bloc led by China and Russia to press for a treaty that limited the openness of the Internet and removed protections on free speech and human rights. In response, Canada, the United States, and many European states refused to ratify the treaty. This divide has given rise to extensive debates about Internet governance, state sovereignty in cyberspace, and the "Balkanization" of the Internet. See James D. Fielder, "The Internet and Dissent in Authoritarian State," in *Conflict and Cooperation in Cyberspace: The Challenge to National Security*, ed. Panayotis A. Yannakogeorgos and Adam B. Lowther (Boca Raton, FL: Taylor and Francis, 2014); Daniel W. Drezner, "The Global Governance of the Internet: Bringing the State Back In," *Political Science Quarterly* 119, no. 3 (Fall 2004): 477–498; Stephen K. Gourley, "Cyber Sovereignty," in *Conflict and Cooperation in Cyberspace*; Michael N. Schmitt, ed., *Tallinn Manual on the International Law Applicable to Cyber Warfare* (Cambridge: Cambridge University Press, 2013); Dana Polatin-Reuben, and Joss Wright, "An Internet with Brics Characteristics: Data Sovereignty and the Balkanisation of the Internet" (paper presented at the 4th USENIX Workshop on Free and Open Communications on the Internet, San Diego, CA, 18 August 2014).

[64]Fearon, "Signaling Foreign Policy Interests." Schelling, *Arms and Influence*. Christian Le Mière, "The Return of Gunboat Diplomacy," *Survival* 53, no. 5 (October–November 2011): 53–68. Dr. Strangelove's Doomsday machine is perhaps the ideal embodiment of Schelling's perfect coercive capacity. We are grateful to an anonymous reviewer for making this analogy.

www.m

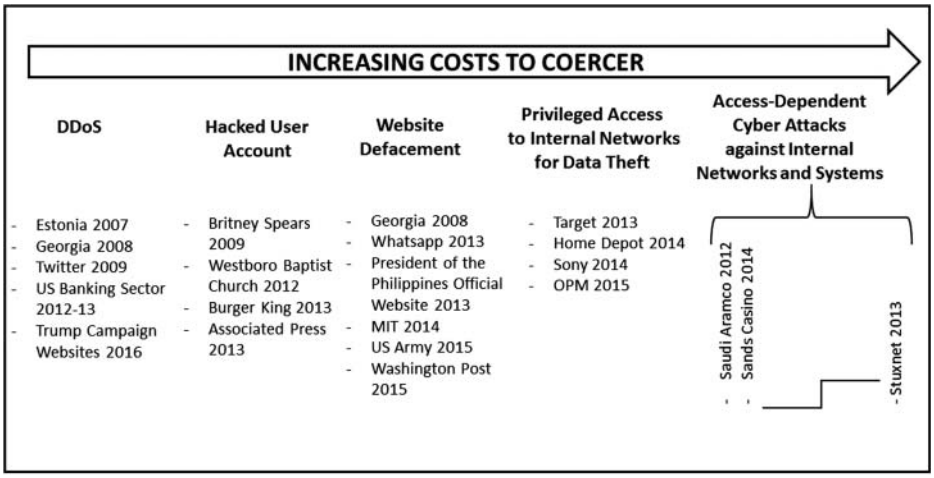

Figure 1. Spectrum of coercive cyber operations.[65]

to demonstrate a state's capability and willingness to follow through with the terms of a threat. The greater the cost to the initiating state of producing a given signal, ceteris paribus, the more effective the signal is as an indication of the initiating state's resolve. Therefore, leaders could use cyber operations to convey their commitment to a particular course of action if they are sufficiently costly to produce.[66] Not all cyber operations are equally costly for the coercing state, however. Some operations are resource intensive, whereas other types of operations, such as a Distributed Denial of Service (DDoS) attacks and website defacements, can be conducted using minimal resources. In this regard, it is helpful to conceptualize interstate cyber signaling as existing along a spectrum where the greater the resource requirements, the costlier the signal is to produce, and the more resolve it demonstrates.

As Figure 1 demonstrates, states could send signals via five broad categories of cyber attacks that are increasingly costly. The cheapest way to attack another entity is to conduct a DDoS attack. This is an operation where multiple compromised systems are

[65]Several of the examples provided in this graphic include cyber operations conducted by nonstate actors that were not necessarily coercive cyber operations. Though this study addresses state-initiated cyber operations, the listed examples are used to provide the reader with highly publicized cyber operations to assist with conceptualization. Furthermore, note the gap between the attacks against the Sands Casino and Saudi Aramco on the one hand, and Stuxnet on the other. There is a dramatic capability difference between the former examples and Stuxnet, which is assessed to have taken the work of thousands of individuals and millions of dollars over a several year time span. For further reference to the complexity of Stuxnet and the resources necessary to develop such a capability, see Ralph Langner, "Stuxnet's Secret Twin," *Foreign Policy*, 19 November 2013. Furthermore, it is important to note that there can be exceptions to the linear increase in costs depicted in this graphic. In other words, some types of attacks may be relatively more costly than how they are categorized in this graphic under unique conditions. For example, a DDoS attack against an extremely hardened target may be more costly to carry out than gaining access to a social media account.

[66]Though nonstate actors may engage in these activities, the scope of this article is limited to state-to-state exchanges. Furthermore, while Fearon discusses both tying hands and sinking costs as mechanisms for generating costly signals, we focus on cyber operations as sunk costs because the tying hands logic is a poor fit for the cyber domain. The only likely allegory to tying hands in cyberspace is the ability, in some instances, to create automaticity by initiating an autonomous offensive cyber response.

www.m

directed by a central computer to flood another computer with information requests. When enough compromised computers are connected together they act as one botnet (a network of enslaved information technology devices that can be centrally controlled) and, if the network is large enough, it may overwhelm the processing capabilities of the intended target and force it to shut down. Examples of this include the alleged Iranian-based DDoS attacks against the US financial sector in 2013, which took down the retail pages of over twenty-six corporations over a four-month time span.[67] These operations are on the far left of the spectrum because they are not inherently expensive to conduct (even though they may force the target to absorb high costs). The current going rate for a 24-hour DDoS attack is approximately $400–800 USD on the black market, depending on the size of the botnet being employed.[68] Furthermore, these operations are access agnostic in that, in order to conduct the operation, the attacker does not have to be pre-positioned with a back door into the target's network to facilitate the attack.

To send a costlier signal, a state could engage in operations designed to hack user accounts, including email and social media accounts. These are slightly costlier than DDoS attacks because they involve acquiring the credentials of an individual with access to the specific target (unless, in the unlikely scenario, the perpetrator can guess the target account's password). A well-known example of this is the 2013 hack of the Associated Press's Twitter feed, where hackers tweeted that there were two explosions in the White House and that the president was injured, prompting volatility in the stock market.[69]

Website defacement represents an additional level of cost for several reasons. First, it requires a minimal level of knowledge of webpage design coding. Second, website defacements involve delivering an effect to produce the observed defacement or redirection. Third, it is dependent on gaining access to the website administrator's account. Notable examples include the defacement of the United States Army's official website in 2015 and the Syrian Electronic Army's hack of a *Washington Post* website in 2015.[70]

Even more costly is gaining privileged access to internal networks for the purposes of data theft. This is more difficult than gaining access to a typical end user's account because it often relies on gaining access to internal systems and data repositories to which end users typically lack access. Most companies limit privileged accesses of this nature and compartmentalize this kind of information due to the potential consequences of a breach perpetrated against even a single actor with such extraordinary accesses (or by the actor herself). There is also an element of scale in these types of cases because attackers can acquire large amounts of private information, such as the contents of corporate email servers, billing records, personally identifiable

[67]Deloitte CIO Journal, "DDoS Attacks on U.S. Banks: Worst Yet to Come?" *Wall Street Journal*, 19 February 2013.
[68]Data comes from black markets accessed on the Dark Net on 17 March 2016. We are grateful to "BillyBear" for his assistance with this.
[69]David Jackson, "AP Twitter Feed Hacked; No Attack at White House," *USA Today*, 23 April 2013.
[70]Polly Mosendz, "Syrian Electronic Army Claims to Have Hacked US Army Website," *Newsweek*, 8 June 2015. Brian Fung, "The Syrian Electronic Army Just Hacked the *Washington Post*, Again," *Washington Post*, 14 May 2015.

www.m

information, and confidential information and documents pertaining to corporate strategy and development efforts. Two well-publicized examples of this kind of attack include the hack of the Department of Defense's Office of Personnel Management in 2015, allegedly committed by China, which compromised the personal information of nearly twenty-two million federal employees and their friends and family; and the 2014 attack against Sony Picture Entertainment, attributed by the US government to North Korea, which released embarrassing corporate communications, policies, and personally identifiable information of employees.[71]

The costliest type of signaling is a cyber attack that requires gaining access to well-defended or closed networks and seeks to disrupt or destroy key systems. Within this category, there is wide variation in the resources required to conduct these operations. The cost depends on the complexity of the attack and the relative difficulty of gaining access to the targeted systems. Since these types of operations disrupt or destroy data, they require customized tools that will produce the desired effect once inside the network. Furthermore, transacting in what is often a well-defended, restricted area is difficult not only because of the code-based language of exchange, but also because gaining access to closed and defended networks requires a significant investment of materiel resources and human capital. This investment includes not only the development of cyber tools to gain access to specific systems, but also the development of capabilities to exfiltrate information resident on the system and/or, more invasively, to completely subjugate the targeted machine. This investment extends beyond the development of cyber arms; it also requires extensive testing against a mockup of the intended target for both the developer and eventual cyber operator. The combination of technical knowhow with financial resources severely limits the number of states that can be called genuine cyber powers—particularly since such investments may be long-term commitments without guaranteed successful outcomes. Indeed, some cyber operations may take years from the time the concept is conceived until the operation is implemented. Operations that involve gaining access to hardened systems that use closed networks not connected to the open Internet, such as the Stuxnet attack against Iran to delay its uranium enrichment program, are significantly costly. In the case of Stuxnet, custom-engineered cyber capabilities containing over fifteen thousand lines of code were required to manipulate Iran's customized Supervisory Control and Data Acquisition (SCADA) systems; this would certainly be costlier than a cyber attack that simply deleted information from servers to which an actor had gained access.[72] In this example, the Stuxnet attack would be significantly more costly to conduct than the 2012 Saudi Aramco breach, which destroyed data

[71] Julie Hirschfeld Davis, "Hacking of Government Computers Exposed 21.5 Million People," *New York Times*, 9 July 2015. David E. Sanger and Nicole Perlroth, "U.S. Said to Find North Korea Ordered Cyberattack on Sony" *New York Times*, 17 December 2014.

[72] Langner, "Stuxnet's Secret Twin." Eric Oliver, "Stuxnet: A Case Study in Cyberwarfare," in *Conflict and Cooperation in Cyberspace*. Jon R. Lindsay, "Stuxnet and the Limits of Cyber Warfare," *Security Studies* 22, no. 3 (July–September 2013): 365–404.

www.m

resident on over thirty thousand corporate computers, due to the time, material, and personnel requirements that would be mandated by the former compared to the later. Finally, this category of cyber attacks could require incurring the additional cost of gaining physical access to a network, particularly if it is closed, through using human operators.[73]

Operating militarily in cyberspace requires a skill set that is not uniformly distributed across all states and takes years to develop. Moreover, unlike traditional means of signaling, sending a signal via cyberspace is uniquely costly because, once an attack capability is used, it often cannot be used again. While it may be possible to replicate a capability, as already noted, there is little universality of cyber capabilities. Most critical targets are unique, and potential victims can prevent exploitation once the threat signature has been identified and incorporated into their defenses, which also compounds the difficulty of a sustained assault. Furthermore, once these tools are deployed they have a limited lifespan as routine defensive techniques and vulnerability patching may render a tool that took years to develop obsolete within seconds of employment.

Governments can also generate costly signals through manipulating the shared risk of war. This concept was championed by Schelling, who submits that credibility can be enhanced by exhibiting risky behavior, particularly during times of crisis.[74] States can demonstrate resolve through acting in a manner that increases the risk of war and/or increases political costs to the party issuing the threat, but falls short of initiating an attack. For instance, a state can raise the alert status of its forces or move naval fleets into close proximity of an area of hostiles. Neither of these signals is inherently costly; however, during a time of increased tension, such maneuvers increase the likelihood of war due to the potential misperception of intent and miscalculation. Furthermore, leaders can generate political costs through tying hands. In other words, politicians that are subjected to electoral sanctioning may generate self-imposed reputational costs by committing themselves to a course of action, which could put their political future in jeopardy if they waiver from it.[75]

In cyberspace, risk generation occurs by acting in overt ways that ensure the receiver perceives the signal, but falls short of a cyber attack. These types of actions include actively scanning networks, pinging pieces of key infrastructure, and perhaps even deploying beacons on compromised infrastructure. These operations can increase the risk of war because their intent cannot be surmised and could be interpreted as a precursory step to offensive cyber operations. However, these operations generate tradeoffs between intelligence collection and coercion strategies that policymakers should take into account.

[73]Owens, Dam, and Lin, eds., *Cyberattack Capabilities*, 83–89.
[74]Schelling, *Arms and Influence*, chap. 3.
[75]Fearon, "Signaling Foreign Policy Interests." For an alternative point of view, see Snyder and Borghard, "The Cost of Empty Threats."

www.m

Reassurance

Finally, to succeed, a coercive threat must have an element of reassurance, such that the target is made to believe that compliance with the terms of the threat will ensure the coercer does not mete out the threatened punishment regardless.[76] In other words, "the pain and suffering have to appear contingent on his behavior; it is not alone the threat that is effective—the threat of pain or loss if he fails to comply—but the corresponding assurance, possibly an implicit one, that he can avoid the pain or loss if he does comply."[77] Related to reassurance, Schelling also describes the importance of saving face—leaving a backdoor that enables that adversary to back down without paying too high a price in its own reputation and integrity. Coercers should therefore deliver the threat in a way that "decouple[s] an adversary's prestige and reputation from a dispute."[78]

Reassurance is also a difficult aspect of coercion. Todd S. Sechser argues that great powers encounter problems reassuring weaker states that are the targets of compellent threats because the very military capability that enhances the credibility of the stronger state's coercive threat makes it more difficult for the target to believe that the stronger state won't simply make more demands following the former's compliance with the initial threat.[79] This sheds light on the inherent tension between the actions that enhance credibility versus those that buttress reassurance; the more a target believes the coercer will actually carry out a threat (credibility), the less likely the target believes the coercer will refrain from doing so in the event she complies (reassurance). Credibility is enhanced when the coercer is forced to inflict some punishment on a target in a way that makes it difficult for the coercer to back down or renege—but the more likely it is that the target believes the coercer will use force, the more insurmountable the task of simultaneously reassuring the target that the threat will be walked back, especially if the coercer's prestige and reputation are engaged in the process of enhancing credibility. In a similar vein, reassurance can be difficult for domestic political reasons; the leader of the coercing state may worry that sending a reassuring signal to an adversary in the context of a crisis situation makes her look weak and irresolute to her domestic public.[80]

Reassurance in Cyberspace

Assuring a target state that, once it capitulates to the aggressor's demands, the punishment will cease is perhaps the greatest obstacle to successful coercion in cyberspace. Effective command and control of a cyber attack are essential for reassurance. However, this is often exceedingly difficult in cyberspace depending

[76]For a broader discussion of assurance strategies, see Jeffrey W. Knopf, "Varieties of Assurance," *Journal of Strategic Studies* 35, no. 3 (April 2002): 375–399.
[77]Schelling, *Arms and Influence*, 4.
[78]Ibid., 125.
[79]Sechser, "Goliath's Curse.
[80]If this is the case, it would imply that democracies may encounter greater difficulties than autocracies when it comes to reassurance.

www.m

on how and by whom an attack is carried out. For instance, 128 distinct cyber attacks were recorded against Estonian websites during May 2007 in response to the Estonian government's decision to relocate a Soviet-era war monument.[81] Since these assaults lacked a centralized controller, it would have been difficult for a unitary actor to provide the Estonian government with a credible assurance that the attacks would cease if the statue were returned to its original location (if we can assume this was the objective of the attacks). Furthermore, many states choose to employ cyber proxies to conduct cyber operations because they may not have the means to conduct the operation themselves or desire plausible deniability. Proxies may not act in the way a government desires depending on the proxy's incentives for participating in the attack and a government's ability to incentivize good behavior.[82] Furthermore, once the attack tool is released, it may be exceedingly difficult to stop. For instance, the Stuxnet computer virus was presumably never intended to propagate beyond Iranian nuclear centrifuges, but it infected over 100,000 computers worldwide before it could be stopped.[83] Due to the technical complexities of cyber capabilities and the collective action issues that may surround command and control of a cyber attack, a rational actor would be wise to second guess a reassurance that an assault will stop in exchange for submission.

A unique paradox occurs as an implication of this analysis. The ideal means to reassure a target is to engage in a disruptive cyber attack. Disruptive attacks are easily reversible and can, therefore, be credibly revoked if a target complies with a coercer's demands. However, disruptive attacks are not particularly costly and, therefore, are less credible than a destructive attack. A destructive attack can deliver an immediate effect, and it also generates irreversible costs to the target that can increase over time. Together, this implies that a coercive cyber attack that both reassures and maximizes costs for the target may be unachievable.

Assessing Warfighting Strategies in Cyberspace[84]

As the above discussion illustrates, cyber power is not an ideal independent tool of coercion. Nevertheless, governments may still choose to use cyber power to pursue warfighting strategies aimed at eroding a target's ability or willingness to resist due to the perceived ease or cost effectiveness of conducting cyber operations as opposed to conventional ones, particularly under conditions of conventional asymmetry—as well as their destructive nature in many cases.[85] In this section, we

[81]Andreas Schmidt, "The Estonian Cyberattacks," in *A Fierce Domain: Conflict in Cyberspace, 1986 to 2012*, ed. Jason Healey (Vienna, VA: Cyber Conflict Studies Association, 2013), 182.
[82]For further reference, see Erica D. Borghard and Shawn W. Lonergan, "Can States Calculate the Risks of Using Cyber Proxies?" *Orbis* 60, no. 3 (Summer 2016): 395–416.
[83]Kim Zetter, "Report: Obama Ordered Stuxnet to Continue After Bug Caused It to Spread Wildly," *Wired*, 1 June 2012.
[84]For further discussion of the likely impact of cyber on strategy in general, see Joseph S. Nye, Jr., "Nuclear Lessons for Cyber Security?" *Strategic Studies Quarterly* (Winter 2011): 18–36.
[85]Anti-Access/Area-Denial strategies currently pursed by states to thwart the movement and maneuver of conventionally superior militaries in a theater of operations typically contain a strong element of cyber power. See, for example, Erica D. Borghard and Shawn W. Lonergan, "Will Air-Sea Battle Be 'Sunk' by Cyberwarriors?" *National Interest*, 8 December 2014.

www.m

Table 1. Assessing warfighting strategies in cyberspace.[86]

	Viable	Nonviable
Effective	Attrition Denial Decapitation	
Ineffective	Intimidation	Punishment Risk

analyze the viability and effectiveness of the six most prominent warfighting strategies that have been explored in the coercion literature and apply them to the cyber domain: attrition, denial, decapitation, intimidation, punishment, and risk. In particular, we discuss the extent to which these strategies are viable based on the current state of the field, as well as the extent to which they can generate sufficient costs to be theoretically effective.[87] We define viability in terms of whether the strategy is technically feasible based on known, existing capabilities.[88] We define effectiveness in terms of whether the strategy can generate sufficient costs to change state behavior. This is conditional on several things: on the target side, whether a target's systems are vulnerable, whether the systems themselves have resiliency, and whether the target as a whole has resiliency beyond the affected system(s) (that is, how dependent the target is on a particular set of cyber-enabled services or capabilities); and on the coercer's side, how costly the attack is to conduct in material, personnel, and political terms. Together, two dimensions of viability and effectiveness form a 2 × 2 matrix according to which the six strategies can be assessed, as depicted in Table 1.

Viable and Effective Strategies

Currently, we argue that there are three warfighting strategies that are likely to succeed using cyber power: attrition, denial, and decapitation. We claim that governments are most likely to achieve desired objectives using these strategies because the technical requirements and capabilities for carrying out these operations in cyberspace exist and because they can generate sufficient costs (in theory) to force a target government to concede. However, it is imperative to note that this discussion remains theoretical and its efficacy in practice is highly context dependent—whether a given government will concede to the demands of a coercing state will depend on the particular cost–benefit calculus it makes for a specific situation. While attrition, denial, and decapitation have different logics, what unites them is their discrete military application—these strategies are generally employed against military targets—and they are most likely to be successful when coupled with

[86]As is the case with conventional coercion, these capabilities may vary tremendously across states. In fact, this is especially likely in an emergent domain.
[87]Given the relative newness of these strategies, much of this discussion remains theoretical rather than applied.
[88]Of course, this is liable to change as the state of technology changes.

www.m

conventional military operations and/or diplomacy. In other words, the use of cyber power to undermine a government's ability or willingness to resist is not as effective in isolation from other instruments of state power.

Attrition

Attrition strategies seek to erode the adversary's military capability such that the target can no longer resist. Within the cyber domain, this strategy could include attacks that both degrade and destroy government or private networks and systems, depending on the latter's military utility. In cyberspace, the successful application of an attrition strategy would force a target to abandon a network or system through destroying it or building up a user's mistrust in it such that the target is forced to abandon its operation. A notable attribute of attrition strategies is that they seek to exhaust a target state's resources as it is forced to dedicate assets to protect or replicate capabilities in different and more secure manners. In particular, cyber raiding—targeting an enemy in its weakest areas—is a common tactic of attrition, where data is the equivalent of an enemy's provisions. Conventionally, raiding refers to stealing or destroying an enemy's provisions or equipment. These forays are commonly conducted behind an adversary's lines and are directed against their supply convoys and depots. In cyberspace, destroying or corrupting servers that handle military plans, air or ship tasking orders, or even defense developmental efforts, can prevent certain actions from occurring at the time they are urgently needed. More importantly, if they persist they will eventually erode a state's confidence in its networks and the data resident on them. It is difficult, if not impossible, to destroy a state's militarily capabilities through the exercise of cyber power alone. However, it is theoretically possible to force a state to suffer the gradual erosion of its capabilities—especially of its confidence in them—as vulnerable targets are attacked and as governments are forced to divert considerable resources to investigating and repairing them until the cost of continued resistance becomes unbearable.

Denial

A denial strategy involves increasing the costs to an adversary such that achieving a military objective—such as taking a piece of territory—becomes prohibitive or impossible.[89] As such, it could involve both a defensive component (increasing one's own defenses such that an adversary cannot go on the offense without incurring significant costs), as well as an offensive one (actively taking out enemy capabilities to deny the adversary the ability to achieve an objective).[90] In cyberspace, the targets of denial strategies mirror those of traditional domains of warfare, except that the effect achieved is delivered via a cyber operation. An adversary's

[89]Robert Pape defines coercion by denial as "using military means to prevent the target from attaining its political objectives or territorial goals." *Bombing to Win*, 13.
[90]Byman and Waxman, *Dynamics of Coercion*, 78–82.

www.m

Integrated Air Defense Systems (IADS), command and control apparatuses, and air traffic control systems are all examples of legitimate targets for a state pursuing a denial strategy. An example of using cyber means (coupled, in this case, with conventional military power) to target an adversary's air defense systems is the alleged 2007 Israeli air attack against Syria's nuclear facilities.[91] Unlike conventional approaches to denial, in cyberspace, due to the increasing reliance of embedded technology in many modern battlefield systems, the surface from which these systems can be attacked has significantly increased. For instance, in conventional warfare the only way to remove tanks from a battlefield is to destroy them piecemeal from the air or ground. However, theoretically, it may be possible in the not too distant future (if not already) to use a cyber attack to render entire fleets of weapon systems inert at a critical moment. This concern has already been realized by many policymakers and is evident in the discussion over Chinese cyber espionage of the research and development of the Joint Strike Fighter.[92] On the other hand, the length of the timeframe under consideration could affect assessments of the potential costliness of denial strategies in cyberspace. For instance, cyber instruments could be used to disable, rather than destroy, an adversary's weapons systems or command and control, rendering an attack costly in the short term but less costly than the ostensibly permanent destruction of those systems through conventional means.[93]

Decapitation

Decapitation strategies seek to achieve strategic paralysis by targeting command and control centers, leadership, critical economic nodes, and key weapons systems.[94] Currently, it is technically possible to use cyber attacks to shut down a command and control node. However, given that most states employ secondary and tertiary redundant systems (for example, analogue or even courier-based communication), as well as separate communication networks (for example, multiple classified and unclassified networks), the impact of this type of operation could be short lived. Nevertheless, successfully targeting a critical command and control node, such as the US government's Secure Internet Protocol Router Network (SIPRNET) or Joint Worldwide Intelligence Communications System (JWICS), would have immediate and significant material and psychological effects. Therefore, governments should either take into account temporal limitations when targeting command and control networks, or ensure that they also target all additional means of adversary communication. Conventional military operations that target command and control facilities can wipe out entire communications networks, for example, through dropping ordnance on a facility. In contrast, cyber

[91]Clarke and Knake, *Cyber War*, 1–8.
[92]David E. Sanger, "With Spy Charges, U.S. Draws a Line That Few Others Recognize," *New York Times*, 19 May 2014.
[93]The authors are grateful to Robert Jervis for pointing this out.
[94]Pape, *Bombing to Win*, 79.

www.m

operations can typically target a single or limited number of communications nodes or networks due to the compartmentalized nature of each network. This would therefore require multiple distinct cyber operations to achieve near-complete command and control paralysis. Furthermore, even if cyber attacks could be used to successfully target a government's primary communications networks, backup systems would likely need to be defeated through traditional forms of electronic warfare or conventional operations (for example, jamming transmissions, capturing curriers, or cutting telephone lines or undersea cables). Altogether, this analysis implies that one is more likely to observe decapitation strategies employed at lower echelons of command, such as troops in the field, where there are typically fewer redundant systems, or against less-capable state adversaries.

Viable and Ineffective Strategies

Warfighting strategies in cyberspace can be technically viable but ineffective because they cannot force the adversary to incur sufficiently high costs to prompt a change in her behavior. Much of the activity that currently occurs in cyberspace falls into this category—actors can harass, annoy, or otherwise inconvenience each other. Indeed, those who claim that the threat of a cyber Armageddon is exaggerated focus on these kinds of cyber attacks.[95]

Intimidation

An intimidation strategy is designed to directly address a state's domestic audiences and sometimes, national policymakers. Actions as part of an intimidation strategy do not cause significant damage and are typically tailored to undermine a government's legitimacy or convince domestic audiences that the government is powerless, prompting a loss of confidence by the public.[96] In cyberspace, intimidation typically takes the form of website defacement and email spamming campaigns. While these operations are technically easy to conduct because they involve fewer resources and a lower skill set compared to other types of operations, they cause minimal cost to the recipient. The effect these attacks produce is typically perceived as an annoyance, rather than a strategic message, because these types of attacks are fairly common and easy from which to recover. Therefore, they are unlikely to be sufficiently costly to force targeted governments to change their behavior. Indeed, observed intimidation strategies, such as the 2008 defacements of Georgian government websites portraying President Mikheil Saakashvili as Adolf Hitler, have had no real effect.[97]

[95]Thomas Rid, *Cyber War Will Not Take Place* (Oxford: Oxford University Press, 2013), xiv–xv.
[96]Andrew H. Kydd, and Barbara F. Walter, "The Strategies of Terrorism," *International Security* 31, no. 1 (Summer 2006): 66.
[97]Jeffrey Carr, *Inside Cyber Warfare*, 2nd ed. (Sebastopol, CA: O'Reilly, 2012), 183–84. Andreas Hagen, "The Russo-Georgian War 2008," *A Fierce Domain*, ed. Healy. Ronald J. Deibert, Rafal Rohozinski, and Masashi Crete–Nishihata, "Cyclones in Cyberspace: Information Shaping and Denial in the 2008 Russia–Georgia War," *Security Dialogue* 43, no. 1 (February 2012): 3–24.

www.m

Nonviable and Ineffective Strategies

The two paradigmatic strategies of traditional coercion that currently have the least utility in cyberspace are punishment and risk. While there has been considerable brouhaha in public and even government spheres regarding the potentially dire consequences of a "World War 3.0" or a "cyber Pearl Harbor," these are largely unrealistic given the current state of the domain.[98] However, as we will describe below, changes in modern societies' interconnectivity and reliance on automated systems, as well as advances in military cyber technologies, could change the value of these strategies.[99]

Punishment

Originally stemming from the work of Giulio Douhet, an Italian general and early proponent of the strategic use of air power, punishment strategies are designed to inflict sudden, large-sale pain and devastation on an adversary's civilian population until the panic-stricken citizenry demands an end to the war.[100] Indeed, Douhet envisioned that a single successful air raid on an enemy's population center could" … spread terror through the nation and quickly break down [a state's] material and moral resistance."[101] This concept was further refined by Schelling and modern coercion theorists, who applied it to the strategic use of nuclear weapons; holding an adversary's population at risk of extreme destruction is the foundation of modern deterrence theory.[102]

In theory, inflicting punishment in cyberspace would involve the use of cyber power to cause virtual and physical damage to civilian infrastructure and population centers. This could entail attacks against essential services, such as water treatment facilities, transportation, air traffic control systems, nuclear power plants, electrical grids, food safety systems, waste management systems, etc. However, in practice, there are two critical elements of punishment that cannot be sustained given the current nature of the cyber domain: first, the immediate and sudden nature of an attack; and second, the scale and scope of the pain. Put simply, governments cannot kill a lot of people in a very short period of time using cyber weapons; the magnitude of the pain states are currently capable of inflicting via the cyber domain alone is hardly comparable to the devastation wrought by conventional or nuclear attacks against cities. Access requirements and the customized nature of cyber capabilities render it nearly impossible to launch a time-dependent, highly coordinated cyber campaign of the scale required to inflict severe costs on enemy populations. The scope is also nearly impossible to achieve because it would require an extraordinarily

[98]See Michael Joseph Gross, "World War 3.0," *Vanity Fair*, 30 March 2012; Erik Gartzke, "The Myth of Cyberwar: Bringing War in Cyberspace Back Down to Earth," *International Security* 38, no. 2 (Fall 2013): 41–73.
[99]The utility of punishment or risk strategies in general is beyond the scope of this discussion.
[100]Giulio Douhet, *The Command of the Air*, trans. Dino Ferrari (New York: Coward-McCann, 1942), 57–58.
[101]Ibid., 57.
[102]Beyond Schelling, see, for instance, Patrick M. Morgan, *Deterrence: A Conceptual Analysis* (Beverly Hills, CA: Sage Publications, 1977); Robert Powell, *Nuclear Deterrence Theory: The Search for Credibility* (Cambridge: Cambridge University Press, 1990); Kahn, *On Thermonuclear War*.

www.m

large number of discrete and distinct cyber attacks. As discussed in prior sections, there is limited universal lethality of cyber weapons, which means that governments would have to develop unique accesses and distinct tools for each targeted system. Moreover, there is no guarantee that an effect can be delivered as planned. Additionally, it is difficult to envision a government entity being able to sustain a cyber assault against multiple key pieces of infrastructure in order to push a society to a breaking point before the target moves to mitigate the onslaught through preestablished redundant mechanisms and/or cyber or kinetic military operations.

Manipulation of Risk

Punishment and risk are fundamentally related—both involve targeting an adversary's population centers to force the government to concede to the coercer's demands. However, unlike punishment strategies that call for immediate and decisive destruction, risk strategies entail gradually escalating the intensity and scope of attacks against civilian targets.[103] There is a critical psychological element to the manipulation of risk in that what drives concessions is the threat and prospect of future pain. This requires that the coercing state can sustain and ratchet up an assault over time.

Like punishment, the manipulation of risk does not translate well into cyberspace. Carrying out a comprehensive, tiered cyber campaign plan to create the ratcheting effect of punishment that Schelling proscribes is exceedingly difficult for reasons already articulated. To wit, this would require a significant planning effort and mandate a costly access and capability development program. Furthermore, risk strategies do not rely upon the sudden and intense destruction that are envisioned by punishment strategies, but instead are designed to be employed over time. In order to be effective, the attack would have to be maintained against an adversary that would likely be active in trying to stop or mitigate the effects of the onslaught. Presumably, if a state is at the technical level where it is susceptible to large-scale cyber attacks, it also has the wherewithal to defend against them over time. Finally, the effective employment of a risk strategy in cyberspace would require an impossibly high level of control by the coercing government over the cyber tools it would employ against an adversary. According to Schelling, risk is most likely to succeed when an action, "once initiated, causes minimal harm if compliance is forthcoming and great harm if compliance is not forthcoming, is consistent with the time schedule of feasible compliance, is beyond recall once initiated, and cannot be stopped by the party that started it but *automatically* stops upon compliance, with all this fully understood by the adversary."[104] Indeed, the risks of using cyber power—effects getting beyond the control of the initiating state in unanticipated and potentially undesirable ways—are precisely the opposite of the calibrated manipulation of risk Schelling envisions.

[103]Schelling, *Arms and Influence*, 3. Also see Pape's discussion of manipulation of risk in *Bombing to Win*, 66–69.
[104]Schelling, *Arms and Influence*, 89. Italics in the original.

www.m

Future Trends in Viability and Effectiveness

The negligible utility of punishment and risk strategies rests on the current state of technology and the dependence (and, therefore, vulnerability) of modern societies on cyber-enabled essential services. Changes along either of these dimensions—technical viability and/or the costs that can be imposed on civilian populations—would alter the feasibility and effectiveness of these strategies. For instance, the dawn of the "Internet of Things" (a concept that depicts a not-too-distant future where everything from an individual's toaster and refrigerator to a city's garbage collection and other essential services are automated and connected to the Internet) could make it possible for governments to impose high and devastating costs on society through cyber means.[105] Moreover, it is conceivable that, as societies remove human redundancy through increased automation and become more dependent on interconnected networks of services, punishment and risk strategies could become more effective as the attack surface expands and more targets become vulnerable to a cyber attack.

Additionally, punishment and risk strategies could become more viable due to better investment in human capital, decreasing costs of planning and conducting large-scale cyber campaigns, increased government spending on developing cyber capabilities, and gaining and maintaining accesses to potential target sets, and the unknown unknowns of potentially disruptive technological innovations that make these attacks easier.

Strategic Implications of the Coercive Use of Cyber Power

This analysis explores the applicability of traditional theories of coercion to the cyber domain. We identify four key elements of coercion—communication, cost–benefit calculus, credibility, and reassurance—and assess how each manifests itself in cyberspace. We then analyze the utility of various warfighting strategies that seek to undermine an adversary's ability and/or willingness to resist and find that, based on the current state of the field, only three—attrition, denial, and decapitation—are likely to be useful for aspiring coercers in cyberspace. However, even these strategies are most useful in conjunction with conventional instruments of power and/or diplomacy; cyber power is rarely, if ever, independently decisive. For policymakers, this suggests that, especially if a coercing state has an asymmetrical advantage in other elements of national power, using cyber power to enable espionage, sabotage, and other shaping operations to support a cross-domain coercive strategy may be a more effective use of cyber capabilities than employing it as an independent instrument of state power.

This framework also highlights the importance of indices and developing an understanding of another state's intentions in cyberspace due to the high risk of misperception, which can lead to unintended outcomes and inadvertent escalation.

[105] Jayavardhana Gubbi et al., "Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions," *Future Generation Computer Systems* 29, no. 4 (September 2013): 1645–60.

www.m

A policy implication of this is that states should focus intelligence-collection efforts on developing an advanced understanding of the cyber capabilities and aspirations of potential adversaries. In addition to indices, states may also send signals through the use of cyber attacks. However, since neither of these signaling mechanisms is inherently clear, the most likely way to convey the intent behind an action in cyberspace is to ensure attribution and couple the event with a diplomatic message or place it within the context of a conventional military operation. Furthermore, for cyber power to be an effective coercive tool, the target needs to believe that an attack will cease once she complies with the coercer's demands. This would require assurances that would have to come via established means that often do not yet exist. Providing a credible reassurance is difficult because many types of cyber attacks, such as DDoS attacks, can come from numerous users and make it difficult for the threatening state to credibly demonstrate it exerts control over a decentralized network of attackers. This leads to a paradox in which the type of cyber attack that is most likely to aid in reassuring a victim may also not be able to generate the punishment that would be necessary for capitulation.

Cyber power can be used as a coercive instrument of state power but, once the theory of coercion meets the reality of cyber operations, many attractive targets may become too costly and out of reach for a state to attack in a timely manner. Therefore, governments are more likely to pursue coercive strategies that allow for a wide variety of targets that are more easily accessible than hardened critical infrastructure. In other words, long development timelines and access constraints often mean that policymakers cannot attack their ideal target(s) in a timely manner and, therefore, are more likely to pursue warfighting strategies that do not necessitate sudden and intense devastation but, rather, inflict costs against vulnerable public and private interests. Given current levels of dependency on technology, this type of attack would provide damaging, but limited, effects. This has unique and potentially troubling implications. Since the end of the Second World War, many states have sought to limit their coercive attacks to key pieces of government and military infrastructure out of ethical and legal concerns surrounding targeting civilian infrastructure (and due to the domestic and international political costs of doing so). However, given that in cyberspace much of the vulnerable infrastructure is owned by private industry, policymakers may reevaluate norms against targeting these systems as they pursue attrition, denial, or decapitation strategies. Cyber warfighting strategies that intentionally target civilian infrastructure, such as punishment and risk, are currently nonviable and ineffective. However, as technology evolves and the Internet of Things makes societies both more interconnected and vulnerable, states may find strategies that explicitly aim to wreak havoc on civilian populations more effective. Together, this suggests that, at the domestic level, governments should strive to continue to build resiliency into civilian networks and, at the international level, norms governing appropriate targeting in the cyber domain are urgently needed.

المنارة للاستشارات

www.m

Acknowledgments

We are grateful to the many individuals in the National Security Agency and the US Cyber Command who shared their candid thoughts with us on this sensitive topic. We would also like to thank Robert Jervis, Jack Snyder, Richard Betts, Thomas Walcott, and Brian Blankenship for their extensive and insightful feedback on our research, as well as to the anonymous reviewers. An earlier version of this article was presented at Columbia University's Symposium on Cyber Conflict, hosted by the Saltzman Institute of War and Peace Studies. We are grateful to Jason Healey and Austin Long for their comments during the conference, which helped guide our revisions and informed the current article. The views expressed in this article are personal and do not reflect the policy or position of the US Military Academy at West Point, Department of the Army, Department of Defense, or US Government.

Funding

The authors wish to thank the Carnegie Corporation of New York and Columbia University School of International and Public Affairs for the grant that made this research possible.

www.m