Coercion through Cyberspace: The Stability-Instability Paradox Revisited

Jon Lindsay
University of Toronto
jon.lindsay@utoronto.ca

Erik Gartzke
University of California San Diego
egartzke@ucsd.edu

25 August 2016

In Kelly M. Greenhill and Peter J. P. Krause, eds., *The Power to Hurt: Coercion in Theory and in Practice* (Oxford University Press, Forthcoming)

[Final copyedited text may differ]

Introduction

Protecting and exploiting computing infrastructure has become a policy priority for governments and other actors around the world. Critical infrastructure for banking, energy, transportation, and manufacturing increasingly relies on embedded computers connected to the internet. Firms and citizens entrust their personal, medical, and financial data to distant servers in return for more convenient and efficient services. Military command and control depends on digital networks to connect pervasive surveillance to power projection capabilities. The same interconnectivity that improves efficiency and control, however, also facilitates new modes of crime, protest, espionage, and warfare. The U.S. Defense Department accordingly describes cyberspace as a new "war fighting domain" alongside the physical land, sea, air, and space domains.[1] Any military forces that drive, sail, fly, or orbit rely on computers for intelligence, communications,

[1] William J. Lynn III, "Defending a New Domain," *Foreign Affairs*, 2010.

logistics, and administration; any intrusions that disrupt, confuse, or deceive these systems may undermine tactical performance, potentially with strategic consequences (i.e., if forces cannot get to the war in time or, cannot target the enemy, and only the enemy, if they do). The cyber domain becomes attractive as a way to shape conflict in other domains, or to bypass military conflict altogether by exploiting civilian infrastructure for political or intelligence advantage.

Widespread belief that offense is easier than defense in cyberspace, that stronger states are increasingly vulnerable while weaker actors are increasingly empowered, and that the anonymity of cyber operations precludes effective deterrence leads many to argue that cyberspace brims with unprecedented, even revolutionary dangers.[2] Yet national security officials, defense firms, media pundits, and a burgeoning private cybersecurity industry all have incentives to exaggerate the threat, while the extreme secrecy of cyber operations complicates sober assessment.[3] Critics of the cyber revolution argue that most actors lack the capacity to overcome significant barriers to weaponization in cyberspace, while those that have the capacity lack the motivation to use it, choosing instead to explore digital variations on traditional espionage and covert action.[4] Nevertheless, even if breathless scenarios of a "digital Pearl

---

[2] Richard A Clarke and Robert K Knake, *Cyber War: The next Threat to National Security and What to Do about It* (New York: Ecco, 2010); Joel Brenner, *America the Vulnerable: Inside the New Threat Matrix of Digital Espionage, Crime, and Warfare* (New York: Penguin Press, 2011); Lucas Kello, "The Meaning of the Cyber Revolution: Perils to Theory and Statecraft," *International Security* 38, no. 2 (2013): 7–40. For counterarguments to these three particular claims, cf. Jon R. Lindsay, "Stuxnet and the Limits of Cyber Warfare," *Security Studies* 22, no. 3 (2013): 365–404.

[3] Myriam Dunn Cavelty, "Cyber-Terror—Looming Threat or Phantom Menace? The Framing of the US Cyber-Threat Debate," *Journal of Information Technology & Politics* 4, no. 1 (2008): 19–36; Jerry Brito and Tate Watkins, "Loving the Cyber Bomb: The Dangers of Threat Inflation in Cybersecurity Policy," *Harvard National Security Journal* 3, no. 1 (2011): 39–84; Sean Lawson, "Beyond Cyber-Doom: Assessing the Limits of Hypothetical Scenarios in the Framing of Cyber-Threats," *Journal of Information Technology & Politics* 10, no. 1 (2013): 86–103.

[4] Thomas Rid, "Cyber War Will Not Take Place," *The Journal of Strategic Studies* 35, no. 5 (2012): 5–32; Adam P Liff, "Cyberwar: A New 'Absolute Weapon'? The Proliferation of Cyberwarfare Capabilities and Interstate War," *The Journal of Strategic Studies* 35, no. 3 (2012): 401–28; Erik Gartzke, "The Myth of Cyberwar: Bringing War in Cyberspace Back Down to Earth," *International Security* 38, no. 2 (2013): 41–73; Jon R. Lindsay, "The Impact of China on Cybersecurity: Fiction and Friction," *International Security* 39, no. 3 (Winter 2014): 7–47; Brandon

Harbor" or "cyber 9/11" are overblown, cyberspace poses real challenges for international relations in theory and practice. As Austin Long argues in chapter 2, intelligence and coercion are increasingly linked, and cyberspace is increasingly valuable for intelligence. Recent events demonstrate that strategic actors are willing to use cyber operations as a tool of statecraft, even as the strategic results have proved ambiguous at best: Russian denial of service attacks and information operations in Estonia, Georgia, and Ukraine; relentless Chinese espionage campaigns and intrusive internet censorship; U.S.-Israeli sabotage of Iranian nuclear enrichment infrastructure; Iranian retaliation against Saudi Aramco and American banks; American cooptation of major internet firms for global signals intelligence collection revealed by Edward Snowden; criminal breeches of leading firms and government agencies exposing the private data of millions of citizens and government employees; North Korean harassment of Sony in Hollywood to protest a satirical movie; Russian attempts to influence the 2016 U.S. presidential election, and the list goes on.

To paraphrase Clausewitz, cyberwar is politics by other means. Understanding the dynamics, magnitude, and likelihood of aggression online requires an assessment of the operational requirements for staging various types of cyber operations, the strategic benefits actors hope to gain through them, and the risks of unintended consequences. In this chapter we lay out a typology of cyber operations that combines the logic of technological possibility with the logic of strategic utility. We distinguish a number of myths that assume cyber attacks can provide high rewards at low cost from more realistic options that deliver variable rewards at variable costs. There is no free lunch in cyberspace. As a result of technical and political constraints on secret operations that depend on interconnections between adversaries, the

---

Valeriano and Ryan C. Maness, *Cyber War versus Cyber Realities: Cyber Conflict in the International System* (New York: Oxford University Press, 2015).

coercive potential of cyberspace is more limited than generally appreciated. Because voluntary connections to the internet make cyber harms possible in the first place, aggressors must be careful not to provoke their victims to disconnect. The social and economic value of the internet both expands and constrains the scope for minor aggression like espionage, covert influence, and symbolic protest. Moreover, the availability of military instruments beyond the cyber domain creates potential for retaliation for unacceptable harms. There are diminishing incentives to "go big" with cyber warfare, even as an adjunct to battlefield operations, because victims have incentives to mount major investigations and shift domains to punish cyber aggression. Coercion still has an important role in cyberspace, nonetheless, especially when exploited in conjunction with other forms power such as military force. We thus delineate the ways in which the cyber domain can be used alone or in conjunction with other domains for deterrence or compellence.

Strategic logic helps to explain the highly skewed distribution of cyber harms we observe historically. While information technology creates the *possibility* for harm, it is political and economic incentives that determine the *probability* of harm. Small-scale aggression online and computer crime is relatively appealing and thus more abundant; large-scale cyber attacks are more difficult and less desirable for initiators and thus far less likely to occur. This argument extends the logic of the "stability-instability paradox" pioneered in the 1960s. Mutually assured destruction may have restrained the superpowers from engaging in direct confrontations during the Cold War, but nuclear threats could not credibly prevent the exercise of proxy wars throughout the Third World. The mechanisms of restraint in the cyber domain are slightly different than in the nuclear world insofar as actors look to maintain connectivity and avoid military retaliation vs. mutual Armageddon, but the results are similar: we see little to none of the most dangerous behavior but a great deal of provocative friction. It turns out that cyber

revolutionaries and cyber skeptics are both partially correct. We should expect to see a lot more creative exploitation of global information infrastructure, but threat actors have strong incentives to restrain the intensity of their exploitation.

The Power to Hurt Online

It is common to hear alarming claims that the U.S. Department of Defense is attacked ten million times per day.[5] In reality most such "attacks" are routine automated port scans from cyber criminals trolling for low-hanging fruit. The majority of actual intrusions, including by sophisticated nation-state "advanced persistent threats" (APT), aim to steal data rather than disrupt systems. Cyber operations (also called "computer network operations") are conventionally divided into three functions: attack, exploitation, and defense.[6] Computer network exploitation (CNE) seeks to preserve the illusion of normal functioning in the target system while illicitly stealing data and using system resources. Attack (CNA), by contrast, may cause servers to shut down, alter important data, or create malfunctions in computer-controlled industrial equipment. CNE and CNA are so closely related that U.S. doctrine considers them together as offensive cyber operations (OCO). Both require the attacker to gain access to a target system through some combination of technical methods and malware (viruses, worms, Trojan horses, rootkits, *etc*.) that take advantage of design flaws or vulnerabilities, social engineering tricks to fool users into revealing sensitive data and passwords (phishing, baiting, water holing, candy drops, *etc*.), and a command and control network to coordinate the attack and obfuscate the attacker's identity. Only the malware payload differentiates exploitation from attack, which

---

[5] E.g., Zachary Fryer-Briggs, "U.S. Military Goes on Cyber Offensive," *Defense News*, March 24, 2012.
[6] A good primer is 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, D.C: National Academies Press, 2009). For U.S. doctrine see "Joint Publication 3-12 (R): Cyberspace Operations" (Washington, DC: U.S. Joint Chiefs of Staff, February 5, 2013).

creates a challenge for the defender trying to differentiate an intelligence penetration from something more sinister.[7] Moreover, most disruptive attacks invariably require supporting exploitation for preparatory reconnaissance and performance feedback. Ambiguity about the purpose and severity of intrusions contributes to something of a siege mentality in popular accounts of cybersecurity.

Cyber defense (CND) includes all measures taken to protect networks from adversarial attack and exploitation: the use of physical boundaries, network firewalls, and intrusion detection systems; bureaucratic procedures to keep software patches and antivirus definitions up to date and to educate users about operational security (or cyber "hygiene"); active monitoring of network activity and investigation of suspicious activity; and coordination with law enforcement and intelligence entities before and after an attack.[8] U.S. doctrine differentiates defensive cyber operations (DCO) that explicitly counter OCO from the routine maintenance of Department of Defense information networks (DODIN), but there is a defensive aspect inherent in any positive use of cyberspace that depends on the confidentiality, integrity, and availability of user data. The witting or unwitting insider threat remains an organization's weakest link: users routinely disregard prophylactic advice or get fooled by creative phishing scams. Sophisticated intruders especially prize "zero day" vulnerabilities (engineering flaws that have not yet been patched by vendors) for eluding detection during infiltration and exploitation since they are unguarded by definition. The use of multiple precious zero days in a single cyber campaign (as in Olympic Games by the United States or the so-called Elderwood group in Beijing) is often indicative of

[7] Ben Buchanan, *The Cybersecurity Dilemma* (London: Hurst, 2016).
[8] Ross Anderson, *Security Engineering: A Guide to Building Dependable Distributed Systems*, 2nd ed (Indianapolis, IN: Wiley Pub, 2008).

skilled and resourced state actor; a lucrative gray market has emerged to peddle zero days.[9] Cyber defense is often assumed to be relatively more difficult than attack or exploitation (i.e., cyberspace is supposedly "offense dominant"); however, it is not obvious that this is categorically true given the planning required for sophisticated penetrations and their vulnerability to quick remediation through patching or reconfiguration once compromised.[10]

The tactical contest in cyberspace, as in any other intelligence contest, is a battle of wits rather than brawn. OCO uses logical code rather than kinetic force to get into the target system, even as downstream effects of system malfunctions may result in physical damage. The hacker must find doors left open or con users into opening them; no amount of pounding on a closed door will cause it to open. As Martin Libicki rightly points out, "There is no forced entry in cyberspace."[11] All OCO depends on deception, and thus system developers and users become unwitting accomplices in their own exploitation. The ubiquity of deception in cyber operations also raises the possibility of using deception to reinforce defense and deter attacks.[12] An attacker who walks through an open door cannot be sure it does not lead to a trap. Defensive deception can undermine the attack through active counterintelligence measures like "honeypots" that draw intruders in for observation and quarantine, or even defensive counterattack (i.e., "active defense" or "hack back"). More complicated attacks are at greater risk of leaving behind clues for forensic investigations in technical artifacts or other behaviors exposed to intelligence

---

[9] Kim Zetter, "How the Secretive Market for Zero-Day Exploits Works," *Slate*, July 24, 2015, http://www.slate.com/blogs/future_tense/2015/07/24/new_insights_into_zero_day_exploit_sales.html.
[10] Keir Lieber, "The Offense-Defense Balance and Cyber Warfare," in *Cyber Analogies*, ed. Emily O. Goldman and John Arquilla (Monterey, CA: Naval Postgraduate School, 2014), 96–107; Drew Herrick and Trey Herr, "Combating Complexity: Offensive Cyber Capabilities and Integrated Warfghting" (International Studies Association, Atlanta, 2016).
[11] Martin C Libicki, *Conquest in Cyberspace: National Security and Information Warfare* (New York, NY: Cambridge University Press, 2007), 31.
[12] Erik Gartzke and Jon R. Lindsay, "Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace," *Security Studies* 24, no. 2 (2015): 316–48.

collection, helping to attribute the identity of the intruder and, often more importantly, the nation-state sponsor.[13] More consequential attacks are also more likely to spur major investigations and create pressure to respond.[14] Attackers must exercise particular caution against resourced and resolved defenders.

Revolutionary Myths: Low Costs, High Rewards

In contemporary defense policy discourse there are three influential narratives of mounting cyber peril, corresponding roughly to the three operational modes of attack, exploitation, and defense. The most dangerous scenarios envision the *paralysis* of industrial control systems or military command and control through surprise attack by anonymous hackers. The imagined aggressor may be a revisionist state like China or Iran or a non-state anarchist or terrorist empowered by the information revolution. A second narrative offers an alternative to the shock of sudden catastrophe, warning instead of the long term *erosion* of economic and military competitiveness drained away through persistent computer espionage. The relentless theft of vital secrets stored on corporate and government networks produces a "death by a thousand cuts" as states give their firms an unfair commercial advantage and equip their military forces with potent countermeasures to U.S. strengths. In both of these scenarios, commercial hacking tools and ubiquitous connectivity give weaker states and terrorists provide a potent means to exploit and attack the expanding attack surface of digitally-dependent advanced industrial states. A third threat narrative concerns the transformation of internet architecture to decisively benefit one political group at the expense of the other. At one extreme, the growth of flexible social media

[13] Thomas Rid and Ben Buchanan, "Attributing Cyber Attacks," *Journal of Strategic Studies* 38, no. 1–2 (2015): 4–37.
[14] Jon R. Lindsay, "Tipping the Scales: The Attribution Problem and the Feasibility of Deterrence against Cyber Attack," *Journal of Cybersecurity* 1, no. 1 (2015): 53–67.

enables connected protesters to overwhelm and overthrow authoritarian regimes.[15] At the other extreme, authoritarian governments censor internet content and reconfigure internet governance to undermine the internet's potential for innovation and freedom. State paranoia about the threats of paralysis and erosion thus leads to digital *lockout* or "the end of the internet" as we know it.[16]

Threats of catastrophic attack, omniscient exploitation, and unassailable defense are myths because they imagine major rewards for little cost. The actual rewards of any given cyber campaign are rarely so great and the costs are rarely so trivial. Potential benefits of attack are discounted by uncertainty about the true value of the target to the adversary and the ability for the attacker to take advantage of it. Operative costs include the bureaucratic resources, development and testing requirements, human capital, and intelligence experience required to plan and run an effective covert cyber campaign. Setting aside the myths of low costs and high rewards (no free lunch), there are a variety of more realistic cyber operations with significant variation in their operative costs and benefits. A set of higher cost, and, potentially, higher reward complements enhance the capabilities of stronger actors who can master them. A much larger set of low cost, low reward irritants are available to weaker actors or even solitary individuals, but they provide only a small marginal return on a small investment.

Table 1: Types of Cyber Operation

Cyber Operation	**Revolutionary Myths**	**Operational Complements**	**Marginal Irritants**
Attack	Paralysis	Disruption	Hacktivism
Exploit	Erosion	Espionage	Fraud
Defend	Lockout	Control	Mobilization

[15] Larry Diamond, "Liberation Technology," *Journal of Democracy* 21, no. 3 (2010): 69–83.
[16] Jonathan L. Zittrain, "The Generative Internet," *Harvard Law Review* 119, no. 7 (May 1, 2006): 1974–2040.

Table 1 summarizes a typology of cyber harms across the three modes of attack, exploitation, and defense, from the mythic free-lunch varietals to more realistic complements and irritants. We use these terms to describe ideal type operations to focus attention on the relationship between costs and benefits, recognizing that many real-world activities going by the same names may blur these boundaries. We first describe the six types of operations we should expect to see from particular types of actors (i.e., complements and irritants) and then how they can be both used and limited through coercive strategies.

Complements: High costs, (Potentially) High payoffs

Operational complements are force-multipliers. They amplify the power of actors who have enough resources and expertise to figure out how to manage the complexity and uncertainty of ambitious intrusions. They are less useful as stand-alone substitutes for material power, but they have the potential to augment the effectiveness of military or intelligence activity in other domains. Cyber *disruption* includes cyber attacks against industrial control systems (ICS, to include Supervisory Control and Data Acquisition or SCADA subsystems) or military command and control systems (C4ISR). The most famous historical example is the Stuxnet attack on Iran's nuclear program, used by the U.S. and Israel in conjunction with diplomatic, sanctions, and intelligence pressure as a covert action alternative to airstrikes. The sophisticated worm required considerable intelligence, preparation, and technical expertise, yet it caused only a temporary loss of Iranian enrichment efficiency—it emphatically was *not* designed to paralyze the Iranian program.[17] Russia's BlackEnergy attacks on Ukraine's power grid likewise occurred only after years of intelligence probes and caused only a six hour disruption of electrical service.[18] Cyber attack can potentially substitute for electronic warfare in the suppression or destruction of enemy

[17] Lindsay, "Stuxnet and the Limits of Cyber Warfare."
[18] Kim Zetter, "Inside the Cunning, Unprecedented Hack of Ukraine's Power Grid," *WIRED*, March 3, 2016.

air defenses (SEAD/DEAD), as reportedly used by Israel to facilitate a 2007 raid on a Syrian nuclear complex.[19] Similarly, Russian hackers coordinated cyber attacks against Georgian government websites and communications in conjunction with its land invasion of South Ossetia.[20] However, planning complications, intelligence gaps, and uncertainties about unintended consequences (e.g., encouraging the propagation of cyber weapons by establishing a precedent for their use) have lead U.S. cyber planners to exercise restraint in considering cyber attacks against Libya and Syria.[21] Disruption can play a useful role in a combined arms military operation, but it takes significant organizational expertise and effort to integrate.

Less cost-intensive but hardly inexpensive, cyber *espionage* is the use of computer network exploitation to access the secrets of an economic competitor or political adversary. Espionage has the potential to alter the balance of power over time, but realizing this advantage in the form of competitive products or effective countermeasures requires an actor to leverage complementary strengths. Chinese APT campaigns have received great notoriety penetrating Western firms, governments, and non-governmental organizations, stealing far more data than Chinese human intelligence (HUMINT) ever managed alone. Yet the theft of secret data is only the first step in converting espionage into competitive advantage to produce, in the words of former National Security Agency (NSA) Director General Keith Alexander, "the greatest transfer of wealth in history." Cyber spies must extract valuable "needles" from a petabyte-scale "haystack" of junk data and then successfully disseminate the take to customers who can make sense of and absorb the stolen data into production or decision processes. While China has

---

[19] David A. Fulghum, "Why Syria's Air Defenses Failed to Detect Israelis," *Aviation Week, Ares Blog*, October 3, 2007. The cyber explanation has been disputed in this particular case, but the general concept is certainly feasible.
[20] Deibert R.J, Rohozinski R, and Crete-Nishihata M, "Cyclones in Cyberspace: Information Shaping and Denial in the 2008 Russia-Georgia War," *Security Dialogue* 43, no. 1 (2012): 3–24.
[21] Ellen Nakashima, "U.S. Accelerating Cyberweapon Research," *Washington Post*, March 18, 2012; David E. Sanger, "Syria War Stirs New U.S. Debate on Cyberattacks," *The New York Times*, February 24, 2014.

invested large sums in improving its capacity to absorb foreign technology, it is still playing catchup to Western innovation.[22]

To capture the positive network effects and efficiency gains that information technology makes possible,[23] networks must be defended against intrusion and misuse. Cyber *control* is neither cheap nor absolute. Indeed, to the degree that disruption and espionage are possible (or hacktivism, fraud, or mobilization, for that matter), perfect control (lockout) is not. This claim is generally uncontroversial, as cyber defense is widely held to involve difficult coordination problems in an offense-dominant medium. Some aspects of cyber defense have private goods characteristics such as firewall and intrusion detection systems protecting the owner's network perimeter, but cybersecurity is also beset by public goods problems.[24] Examples include users who opt not to patch their systems and end up hosting botnets that attack other users and software venders who neglect the development security features in the rush to get their products to market. Meanwhile, offensive cyber threats may change their signatures faster than defenders can keep up.[25] Even authoritarian governments cannot achieve absolute advantage in the arms race with technologically-savvy dissidents, at least as long as those states also desire digital access to international economic transactions. Yet weak dissidents face even greater challenges assuring control of their social networks, since regime actors can use not only espionage but also more draconian enforcement measures against them. The relationship between internet control,

---

[22] Jon R. Lindsay and Tai Ming Cheung, "From Exploitation to Innovation: Acquisition, Absorption, and Application," in *China and Cybersecurity: Espionage, Strategy, and Politics in the Digital Domain*, ed. Jon R. Lindsay, Tai Ming Cheung, and Derek S Reveron (New York: Oxford University Press, 2015).

[23] Erik Brynjolfsson and Adam Saunders, *Wired for Innovation: How Information Technology Is Reshaping the Economy* (Cambridge, MA: MIT Press, 2010).

[24] Ross Anderson and Tyler Moore, "The Economics of Information Security," *Science* 314, no. 5799 (October 27, 2006): 610–13; Paul Rosenzweig, *Cyber Warfare: How Conflicts in Cyberspace Are Challenging America and Changing the World* (Santa Barbara, Calif.: Praeger, 2013).

[25] Dale Peterson, "Offensive Cyber Weapons: Construction, Development, and Employment," *Journal of Strategic Studies* 36, no. 1 (2013): 120–24.

innovation, and freedom is too complicated to analyze further in this chapter, but that very complexity makes either version of lockout an unlikely possibility.[26]

Irritants: Low costs, Low payoffs

Operational complements can make the strong even stronger, if coordinated with other sources of strength. Contrary to conventional wisdom about the asymmetric nature of cyber warfare, operational complements tend to advantage nation state actors, especially great powers with institutional resources who can overcome planning and intelligence challenges and price in the risk of failure. Marginal irritants, by contrast, are widely affordable for all types of actors, weak or strong, and they can be employed with lower risk of adverse consequences. The risks are low because those who have the power to intervene to stop or punish irritant behavior often do not have the motivation to do so. While irritants are often illegal under domestic statutes, law enforcement authorities often do not launch sufficiently aggressive investigations, for want of resources or authorization, to discover and sanction the perpetrators. Irritant attackers can thus hide safely behind their digital anonymity, whereas the use of offensive complements described above would provoke a more concerted investigation and response. By the same token, the expected rewards are low.

The vast majority of empirically-observable cyber attacks can be grouped into the category of *hacktivism*, which includes distributed denial of service (DDoS) attacks that knock servers offline, website defacements, defamation, and other forms of online protest. An early example of hacktivism as defined here is the barrage of DDoS attacks and defacement from

---

[26] Ronald Deibert et al., eds., *Access Contested: Security, Identity, and Resistance in Asian Cyberspace Information Revolution and Global Politics* (Cambridge, MA: MIT Press, 2012); Sarah McKune, "'Foreign Hostile Forces': The Human Rights Dimension of China's Cyber Campaigns," in *China and Cybersecurity: Espionage, Strategy, and Politics in the Digital Domain*, ed. Jon R. Lindsay, Tai Ming Cheung, and Derek S Reveron (New York: Oxford University Press, 2015).

Russian nationalists that wracked Estonia in 2007 following the removal of a Soviet memorial in Tallinn. Similar activity from Chinese nationalists is also common in China's periodic tensions with Taiwan and Japan. The anarchist collective Anonymous has embarrassed several firms and government agencies by illicitly acquiring and then publicly posting (doxing) confidential data. Colloquially "hacktivism" is sometimes used to cover what we call mobilization in our typology (i.e., when hacktivists seek to encourage one another and create support for a cause). The same technical methods used for criminal exploitation or what we categorize as fraud can also be considered hacktivism if conducted for the purposes of defamation or political influence rather than financial gain (e.g., when hackers steal and expose confidential information after the manner of Wikileaks). Hacktivist techniques might also be combined with OCO disruption as in North Korea's attack on Sony which destroyed hard drives and released embarrassing internal memos, an episode which attempted but ultimately failed to prevent the release of a satirical comedy, exemplifying the limitations of anonymous cyber coercion.[27] Hacktivist attacks can grab headlines and be embarrassing or financially costly to those they target, but they usually subside within a few news cycles. At the technical level, mitigation techniques for things like DDoS are readily available. DDoS and defacements have become prominent during almost any period of political tension as a form of nationalist protest, and there is often much ambiguity behind whether government actors or nationalist citizens are responsible, a plausible deniability that can aid either hacktivists seeking to avoid punishment or targets that are not motivated enough to punish them.

[27] Stephan Haggard and Jon R. Lindsay, "North Korea and the Sony Hack: Exporting Instability Through Cyberspace," AsiaPacific Issues (Honolulu, HI: EastWest Center, May 2015).

Cyber *mobilization* is the use of social media to coordinate protest activity online and on the ground. We categorize this as a mode of defense because dissidents seek to use social media as designed to realize its positive network benefits. Recent large scale political protest movements associated with the Arab Spring or Ukrainian EuroMaidan leveraged social media to mobilize and organize protesters.[28] However, whatever success they enjoyed owed more to the tenacity of the protestors, their physical presence in number, and the restraint of government security forces. The limits of mobilization were highlighted in the abortive Iranian "Green Revolution" of 2009 and the Chinese "Jasmine Revolution" of 2010, where government security forces exploited the use of social media to identify and punish protesters.[29] The defenses of mobilization are weak because software products must be distributed wholesale to interconnected users with uneven technical skills. State-sponsored attackers can exploit this vulnerability; for example, the Chinese state sought to provide spyware to protesters in Hong Kong in 2014.[30] Mobilization can also take on virulent or socially dysfunctional forms such as cyber bullying and its industrial scale Chinese analogue known as "human flesh search," a virtual lynch mob defaming corrupt officials, disgraced celebrities, and unfortunate citizens alike.[31]

Financially-motivated computer crime or *fraud* accounts for the overwhelming empirical manifestation of cyber insecurity. There is a complex, segmented, global market for cybercrime divided into interdependent advertising, theft and fraud, and technical support rackets. This

---

[28] Diamond, "Liberation Technology."
[29] Evgeny Morozov, *The Net Delusion: The Dark Side of Internet Freedom*, Reprint edition (New York: PublicAffairs, 2011).
[30] Shannon Tiezzi, "China's Cyber War Against Hong Kong Protestors," *The Diplomat*, October 1, 2014, http://thediplomat.com/2014/10/chinas-cyber-war-against-hong-kong-protestors/.
[31] Saul Levmore and Martha C. Nussbaum, eds., *The Offensive Internet: Speech, Privacy, and Reputation* (Cambridge, MA: Harvard University Press, 2011).

underground ecosystem also supports more sophisticated OCO complements by providing malware and compromised hosts to facilitate intrusion. However, APTs that conduct targeted espionage require much more skill and effort per target and assume greater risk. Unlike APTs, retail cybercrime is untargeted and scales more easily to exploit millions of potential victims, since it only has to be successful a fraction of a percent of the time to be profitable. Underground revenues total hundreds of billions of dollars annually, although the vast majority of cyber criminals actually make very little money because of rampant dishonesty in the underground economy and nontrivial law enforcement risks. Even spectacular compromises of millions of credit card accounts do not readily translate into handy profits because translating that data into a useable monetary instrument is difficult.[32]

To sum up, cyber attack, exploitation, and defensive operations can be employed as costly complements to enhance other advantages an experienced and resourced actor might have, or as inexpensive irritants for minor gain. Myths of grave harms are based on an unrealistic assessment of the operational barriers and strategic risks involved. For the most part our discussion so far has focused on the ways in which cyber means can provide some direct benefit to adjust the balance of power, however marginally, via OCO or DCO support to military operations, intelligence collection, information system control, symbolic protest, political mobilization, financial gain, etc. Some mention of deterrent threats (or their absence) has been unavoidable even in the discussion of cyber harms, such as the unwillingness of law enforcement

[32] Kirill Levchenko et al., "Click Trajectories: End-to-End Analysis of the Spam Value Chain," in *Proceedings of the 2011 IEEE Symposium on Security and Privacy*, SP '11 (Washington, DC, USA: IEEE Computer Society, 2011), 431–46; Ross Anderson et al., "Measuring the Cost of Cybercrime," in *The Economics of Information Security and Privacy*, ed. Rainer Böhme (Berlin: Springer-Verlag, 2013), 265–300; Cormac Herley, "When Does Targeting Make Sense for an Attacker?," *IEEE Security & Privacy* 11, no. 2 (2013): 89–92; Jianwe Zhuge et al., "Investigating the Chinese Online Underground Economy," in *China and Cybersecurity: Espionage, Strategy, and Politics in the Digital Domain*, ed. Jon R. Lindsay, Tai Ming Cheung, and Derek S Reveron (New York: Oxford University Press, 2015).

to pursue all irritants, or the willingness of strong actors to use disruptive complements when backed up by other capabilities. We now turn from harmful means to coercive ends, and from the power to hurt to the power to persuade.

Cyber Coercion

As Robert Art and Kelly Greenhill point out in chapter 1, coercion is different from pure harm, even though coercing may require causing some harm in the process of creating credible threats of more harm. Harm pure and simple (brute force), aims to change the balance of capabilities between adversaries in the present, while coercion uses harm or threats of harm to influence an adversary's decision-making in the future. Coercion is a signaling process which attempts to link particular behaviors to unpleasant consequences in the mind of the opponent. It targets the willingness of the opponent to endure suffering or comply with demands and can include *deterrence*, to prevent something from happening, or *compellence*, to cause something to happen, as well as more complicated forms of signaling which we will not address in this chapter.

The future-directedness of all forms of coercion seems to create a problem in cyberspace. How can an adversary be made to understand a credible threat of future harm via network connection and yet voluntarily maintain the connections on which that future harm depends? As mentioned above, cyber intrusions depend on deception because logical, massless code cannot kinetically force its way through anything. If software vulnerabilities are highlighted by an explicit threat to exploit them unless the target complies with a demand, then the target can patch or otherwise neutralize the threat. Therefore, if attackers rely on the difficulty of attribution to protect themselves, then they cannot easily make credible coercive demands which would reveal their identity and methods. Moreover, if bad intelligence or buggy malware in the threatened cyber attack causes unexpectedly high or low damage when exercised, then the punishment may

have little resemblance to the threat. Vague threats from anonymous sources lack credibility, fail to precisely specify the action proscribed or demanded, and offer little reassurance that the coercer will withhold punishment if the target complies. Furthermore, a necessary reliance on deception and ambiguity creates considerable potential for the misperception of coercive signals.[33]

Nevertheless, cyber operations are not completely devoid of coercive potential, even if they are more limited compared to traditional means of aggression. Cyber operations can be employed in some circumstances for deterrence and compellence, especially if expectations for success are limited. The most promising coercive cyber tools are complements (rather than irritants) because they have the potential to impose higher costs in conjunction with other (non-cyber) tools. Table 2 summarizes strategies for coercion using cyber ("within domain") and "cross-domain" means.

Table 2: Cyber Coercion

	Deterrence	**Compellence**
Cyberspace	Detection Denial Deception	Latency Extortion Seduction
Cross-Domain	Retaliation Disconnection	Escalation Protection

Cyber Deterrence

The great ambiguity of attribution in cyberspace is generally thought to undermine the credibility of retaliatory threats and thus cyber deterrence in general.[34] However, these assumptions are not categorically true across the entire range of cyber operations. Pervasive surveillance or the threat

[33] Gartzke, "The Myth of Cyberwar."

[34] For extensive discussion see Martin C. Libicki, *Cyberdeterrence and Cyberwar* (Santa Monica, CA: RAND, 2009); National Research Council, ed., *Proceedings of a Workshop on Deterring Cyberattacks: Informing Strategies and Developing Options for U.S. Policy* (Washington, D.C.: National Academies Press, 2010).

of detection can augment general deterrence. Public investment in cyber defense can improve deterrence by denial. Defensive deception can improve both punishment and denial strategies.

Insofar as cyber operations are particularly well suited for intelligence, they should also be expected to enhance the intelligence dimensions of coercion, discussed in chapter 2 of this volume. A reputation for skill at surveillance by any means can dissuade targets from planning and executing harmful operations. Intelligence collection can provide the direct benefit of targeting data and it can help, potentially, to shift the balance of power over time by transferring knowledge and evening the playing field (with all the caveats regarding absorption mentioned above). Moreover, the fact or fear of surveillance can also encourage paranoia and force a target to adopt onerous security measures. Extensive signals intelligence (SIGINT) monitoring of underground groups forces many of them to rely on couriers who are slower, more expensive, and less efficient than mobile phones. Battlefield targets that adopt debilitating "emissions control" postures effectively commit "EMCON suicide." Paranoia that detection will be followed by swift precision strikes or law enforcement action may cause targets to forgo misbehavior altogether. Likewise, citizens subject to continuous monitoring via public cameras and internet surveillance tend to internalize obedience to the state or at least curtail observably deviant behavior. Pervasive cyber espionage may be undesirable from a civil liberties perspective but may be effective for suppressive coercion.

Deterrence by *detection* works when a potential cyber attacker fears that the probability of detection is high and is concerned about the consequences of getting caught—detection provides the option for targeted retaliation. While attribution is not strictly needed for punishment (think of a minefield or retributive razing of a village), prompt detection and

convincing attribution does lower the costs and thus the credibility of punishment.[35] Credibility here is enhanced by cultivating a reputation for skilled cyber exploitation, which in turn improves the credibility of threats of punishment by whatever means, cyber or cross-domain. While Edward Snowden's leakage of top secret NSA documents certainly compromised technical intelligence sources, it also (inadvertently) helped the U.S. to advertise the technical skill of the NSA. This leak was credible because it was also costly in terms of lost sources and, potentially, lost market share for US firms wittingly or unwittingly collaborating with the NSA. Michel Foucault introduced the metaphor of the "panopticon" to describe how pervasive state surveillance deters social deviance.[36]

A reputation for skilled cyber defense, above and beyond the ability simply to detect threats, enhances deterrence by *denial*. Cyber defense can block, parry, or redirect intrusions. Public knowledge of the robustness of defense makes attackers worry that their efforts might be futile, even dangerous. Defensive aptitude can be signaled through costly investment in cyber intelligence, law enforcement, and regulatory agencies and the advertisement of success in detecting and thwarting attacks. Creation of U.S. Cyber Command to defend military networks, major exercises like the Cyber Storm series, and budgetary investment in training and capabilities all provide signals of American commitment to the defense of its warfighting networks. It is not necessary for all intrusions to be prevented, moreover. In the case of Chinese APTs, the public disclosure of Chinese tradecraft and PRC government responsibility, heightens the suspicion future Chinese intruders must face when trying to hide or deny their involvement in espionage. Indeed, Chinese APT activity paused and reconfigured following the February 2013

[35] Lindsay, "Tipping the Scales."
[36] Michel Foucault, *Discipline & Punish: The Birth of the Prison*, trans. Alan Sheridan (New York: Random House, 1977).

exposé by cybersecurity firm Mandiant, which identified a specific Chinese military unit in Shanghai involved in the exploitation of 141 English-speaking targets worldwide via a variety of lapses in Chinese tradecraft (e.g., operator reuse of names and emails, command and control servers located in China, operators checking personal Facebook accounts from their attack infrastructure, etc.).[37]

Deception is an underappreciated strategy that is particularly promising for network protection. Ruses, honeypots, digital bait, data obfuscation, and more aggressive counterintelligence techniques have already been employed by security engineers. Deception can confuse, delay, misdirect, or even harm the attacker, for instance by enabling the exfiltration of harmful malware to infect the attacker's home networks. Whereas pure deterrent strategies punish intrusion and pure denial strategies impede it, deception actually encourages intrusion, but then turns it against the intruder. For this reason, deception is rightly considered a distinct protective strategy, even though in practice it operates to reinforce defense or deterrence by punishment or denial. While deception has always been available and has been practiced in the past, the growth in complexity of information technology from the telegraph to the internet has made deception more possible and useful than ever before. The supposed offense dominance of cyberspace is actually a result of the increased potential for deception, especially for less sophisticated gambits, but defenders can use deception too. There are clearly some operational and legal challenges associated with cyber deception, and all forms of active defense. Credible

[37] Mandiant, "APT1: Exposing One of China's Cyber Espionage Units," White Paper, (February 2013), http://intelreport.mandiant.com/Mandiant_APT1_Report.pdf; "M-Trends: Beyond the Breach, 2014 Threat Report" (Mandiant, April 2014).

deception is difficult, so as with control complements generally, the advantage in deception usually goes to the stronger and/or better organized actor.[38]

The traditional distinction between punishment and denial is muddied for forms of cyber defense that involve counterattack and deception against the intruder. U.S. officials have begun to announce that the cyber attribution problem is not as daunting as once believed and that attackers will be met with a decisive response in cyberspace or elsewhere.[39] A reputation for skill at cyber operations is useful not only for the deterrence of cyber attacks (because of the risk of detection and punishment or of denial) but also for more general coercion, if they augment the effectiveness of other threats or undermine the target's confidence in defending against them. So far the U.S. has demonstrated the greatest capacity for and willingness to use cyber operations through its Olympic Games program which allegedly produced the Stuxnet attack on Iran. Unfortunately, this same case highlights the deterrent limitations of cyber punishment, as Iran continued to enrich uranium and even accelerated the modernization and relocation of its program after 2010 while also pursuing its own offensive cyber program.[40]

There is no reason for threats of punishment or *retaliation* to be limited to cyber actions. In fact, the most important bounds on the severity of cyber aggression probably have nothing to do with the technology of cyberspace. Victims of aggression can and likely will look to responses not only in kind but also through whatever other means they possess, ranging from conventional military retaliation, irregular warfare and covert subversion, trade and financial sanctions, and so forth. The problem of attribution is often thought to preclude the ability to

---

[38] Gartzke and Lindsay, "Weaving Tangled Webs."
[39] Elisabeth Bumiller and Thom Shanker, "Panetta Warns of Dire Threat of Cyberattack on U.S.," *The New York Times*, October 11, 2012.
[40] Christopher Bronk and Eneken Tikk-Ringas, "The Cyber Attack on Saudi Aramco," *Survival* 55, no. 2 (April 3, 2013): 81–96; Lindsay, "Stuxnet and the Limits of Cyber Warfare."

retaliate, but as mentioned above, this problem is overstated for any serious attack. The question for an aggressor contemplating truly serious cyber harm is not whether but how the victim will retaliate. The only exception would be for the attacker to launch a cyber attack so devastating that it effectively paralyzes all ability for the target to retaliate, but as we argued above, such paralysis is a myth for operational reasons alone—a splendid first cyber strike is simply too complex and full of uncertainties to reliably deliver its results—let alone the risks of cross-domain retaliation. If the victim retains significant options in other domains where the attacker's ability to resist is slight, then the attacker has strong incentives to avoid provoking a response in those domains. Even if the target is asymmetrically more dependent on the internet, making disruptive cyber attack seem like an attractive possibility, advanced industrial countries are for the same reason more likely to have other advanced military and economic options available where the asymmetries do not favor the attacker. Serious attacks invite serious responses, which need not be in kind.

Perhaps the simplest form of cross-domain response to cyber threats is to forgo the use of the cyber domain altogether. While it is hard if not impossible to limit exposure to nuclear weapons and even a determined conventional assault, the risk of cyber attack can be completely eliminated by *disconnection* from digital networks. The internet is an artificial environment and connection to it is voluntary. Individuals, organizations, and states retain the ability to unplug completely, limit their online transactions, or erect various barriers to connection. Obviously disconnection is not very feasible commercially, socially, and militarily today, but this is more of an indicator of how positive the benefits of interconnection are compared to the perceived risks. If the risks were perceived as extreme, then firms and states could go back to making a living as they did before 1991 (when the World Wide Web came to be). This is a cross-domain threat

because it entails exiting the cyber domain to some degree to engage in more traditional economic and military transactions. The threat of disconnection follows from the more general logic of international institutions, where contracts must be self-enforcing.[41] On the internet as in institutions, ties among egoistic actors under anarchy must be mutually beneficial. If the internet is a bad deal for actors, they can throw up boundaries or exit cyberspace altogether. The threat of voluntary disconnection is especially relevant for repeated interactions, or repeated exploitation, rather than a one-shot "bolt from the blue" cyber attack, which is better countered with cross-domain retaliation. The threat of disconnection is implicit in the voluntary nature of connection to the internet, and the potential loss of the ability to make future attacks exercises a deterrent effect on attacks in the present. An aggressor who does not want to lose access to cyber complements for espionage and disruption it has invested so much in developing will show restraint in their employment. This does not mean that coercion cannot take place online, but dependence on mutual interconnection bounds coercion by excess value. One implication is that the countries that can be most coerced on the internet will be those that have the most to lose by leaving it.

Detection—leveraging cyberspace as a panopticon—removes the cloak of anonymity cyber attackers depend on and facilitates retaliation. Denial—a reputation for effective defense—counters attackers who believe they can, nonetheless, maintain their tactical covertness. Deception—using attackers' strengths in stealth against them—reinforces both punishments and denial. Importantly, all three of these strategies are useful for the entire range of cyber operations, complements as well as irritants, although for many irritants the effort for protection may not be worthwhile. By contrast, the other two cross-domain strategies—retaliating by any

[41] Kenneth A. Oye, ed., *Cooperation Under Anarchy* (Princeton University Press, 1986).

means necessary or disconnecting from the threat altogether—are aimed more at deterring high impact (complement) aggression or keeping aggression confined more toward the lower (irritant) end of the spectrum. As we discuss in the conclusion, this gives rise to a cyber analogue of the stability-instability paradox.

Cyber Compellence

Schelling and others have argued that compellence is harder than deterrence. Deterrence dissuades while compellence persuades. Deterrence stops something in motion while compellence starts something at rest. Deterrence need only signal, "Don't cross this line," while compellence must signal, "Move across this line, and only so far." If the U.S. threatens or conducts a cruise missile attack against a state believed to be harboring a suspected leader of al-Qaeda (i.e., using the base state coercion strategy described by Keren Fraiman in this volume), the goal is to get the target to extradite the al-Qaeda leader. Would a nonlethal cyber attack or threat of attack be able accomplish the same thing? When would coercion work because of cyber means but not because of something else? Is there any target that cyber tools have particular power to compel? If cyber deterrence is thought to be difficult, surely cyber compellence should be more difficult still.

The biggest obstacle to cyber coercion is the difficulty of credible signaling about potential harm when it depends on secrecy to be harmful. Advertised cyber threats that are specific enough to be credible can be neutralized through patching, reconfiguration, or other countermeasures. Sacrifice of the anonymity on which offensive deception depends exposes the cyber attacker to retaliation. Coercive cyber threats thus tend to be more generalized, which undercuts their effectiveness in targeted or crisis situations. The effectiveness of deterrence by deception (the cyber panopticon) depends more on generalized concerns that an intelligence

service might be reading email or snooping around networks than any immediate threat of detection. This form of deterrence is especially effective when employed by a strong state with a sophisticated digital surveillance system. A compellent analogue is the generalized threat of escalation from cyber exploitation to disruptive attack. Detection carries the latent possibility of punishment, and exploitation the latent possibility of attack.

Escalation *latency* is not a specific threat of harm, but more of a generalized paranoia that "our networks are already so penetrated that resistance is useless!" If intelligence intrusions into vital systems are detected, the target cannot be sure that the same intrusion might be used to activate or deliver a more dangerous payload. The potential for escalation from exploitation to attack is latent most of the time in reality, yet the inherent ambiguity about purpose and scale encourages some paranoia: what if the networks on which our prosperity and strength depends were turned against us? What if defense against societal paralysis is as fruitless as defense against cybercrime? What if all it takes is a change of mind by the adversary to convert probes into destruction? Technological, intelligence, and operational constraints in fact render the fungibility of technique from exploitation to attack less of a realistic concern than oft feared (because engineering destruction on command requires significant additional expertise and testing), or restrict it only to large actors with resources and experience to make disruptive complements work. Nonetheless, the prospect of ongoing technological innovation and falling barriers of entry to cyber attack tempt many observers to take this dormant potential seriously. Since this fear rests on a misperception—escalation is operationally nontrivial, victims of network disruption often find alternatives, and both sides will learn these points as the crisis drags on—cyber latency is not very reliable for compellence on its own. But anything that

creates fear can be at least a little bit useful in a broader gambit; again cyber appears useful as a complement to, not a substitute for, other means of coercion.

Specific and credible coercive signaling is more feasible in the realm of irritant cybercrime beneath the threshold of state-sanctioned retaliation. There exist *extortion* scams that use malware (ransomware) to disable a computer or lock out access to data unless the victim pays into an account which the anonymous attacker can access.[42] This strategy uses the threat of disconnection in a compellent rather than a deterrent role (extortion is also analogous to denial as both threaten failure of the target's cyber operations, the failure of normal operations or the failure of an attack). Cyber blackmail is only useful at small scales where the ransom is less than the reconstitution cost of the embargoed system or the cost of contacting law enforcement. It thus can be most effective against victims who themselves might be on the wrong side of the law, e.g., a threat to embargo a bookie's books or computers that support an illegal racetrack. Authoritarian states can also use threats of disconnection from the internet to cow media outlets and muzzle dissidents. Attempts to cut off a population from the internet altogether, for example in Egypt during the Arab Spring or in China during the Xinjiang unrest of 2009, bleed into the category of control rather than coercion, although expectations that a "kill switch" might be thrown could have a slight persuasive effect on dissidents (who would surely worry more about riot police if it came to that). One can conjure up scenarios where a cyber attack causes costly malfunctions but for some reason audience costs prevent the victim from publicly admitting to the damage or the mechanism. Joel Brenner imagines a scenario where China blackmails a US President by knocking out large sections of the US power grid run that only China can repair through its monopoly on a particular type of generator, but only if the US recalls a carrier strike

[42] Ian Urbina, "Hackers Find New Ways to Breach Computer Security," *The New York Times*, June 21, 2014.

group en route to Taiwan.[43] Importantly, all such examples of cyber extortion can only work where the victims are weak or lack alternatives or recourse for some reason. Brenner's scenario would almost certainly backfire on the Chinese.

We discussed above how deception can enhance deterrence by punishing intruders who steal baited data or attack honeypot systems. Deception can also enhance compellence by making the desired action appear more attractive to the target than it actually is. The strategic essence of deception is persuading an adversary to voluntarily take action that is not, ultimately, in its interest. Compellent deception can be described as *seduction*. It is the bread and butter of commonplace internet fraud and one of the charms of the Nigerian princess who needs your help to recover lottery winnings held by an alien hovering over London. Cyber seduction might also be useful for sending bogus orders to enemy troops in the field instructing them to retreat or stow their weapons, or to lure enemy commanders into an ambush. Importantly, as with other cyber complements, the window of vulnerability created through a seductive ruse is only useful if the seducer plans to exploit it in the terrestrial domain.

OCO can be used conjunction with other forms of military force—e.g., disabling early warning or command and control networks in support of an air campaign—to make a threatened intervention seem more potent, if an actor is inclined to threaten military action. That is, a hostile adversary armed with cyber weapons in addition to tanks, fighters, and missiles might appear more able to overcome defenses and thus be better able to issue compellent threats. Likewise, a preparatory cyber campaign to destabilize communications or carry out deception operations could signal a willingness to employ more meaningful forms of aggression in a process of cross-domain *escalation*. Certainly the presence of aggression online helps to cultivate an air of crisis

[43] Brenner, *America the Vulnerable*, chap. 7.

which could be useful for strategies of risk or brinksmanship. Something like this seems to have happened in Ukraine in 2014 where cyber attacks accompanied the incursion of Russian commandos to take Crimea without resistance, while mechanized Russian forces mobilized and exercised on the border.[44] Another route to coercion involves outbursts of DDoS attacks and website defacement by civilians to cultivate an air of crisis and provide elites with a diplomatic argument that their hands are tied by popular nationalism. The hands-tying argument loses credibility if the state has demonstrated an ability in tamp down online outbursts at will, as China has. Importantly, all of these mechanisms depend on the coercer having the ability to escalate beyond the cyber domain, leveraging the fear of substantive military assault. Cyber aggression is used to signal the risk or potency of cross-domain escalation.

Cyber defense can likewise enhance the credibility of coercive military threats. If a coercer has and is known to have robust cyber *protection*, then the target will have less faith in preempting or defending against the threatened punishment. Conversely, a coercer who lacks adequate cyber defenses is like the proverbial stone-thrower in a glass house. Threats will lack credibility if exercising them means assured retaliation, in this case a cyber counterattack on military C4ISR that delays or degrades the ability to carry out the promised punishment. But if cyber defense can be assured, then it is easier—or at least less hard—to project military power in the service of coercive diplomacy.[45] The protection of computer networks is thus an important complement to the issuance of cross-domain threats that depend on them. Protection is the

[44] David E. Sanger and Steven Erlanger, "Suspicion Falls on Russia as 'Snake' Cyberattacks Target Ukraine's Government," *The New York Times*, March 8, 2014.
[45] Michael Fortmann and Stefanie von Hlatky, "The Revolution in Military Affairs: Impact of Emerging Technologies on Deterrence," in *Complex Deterrence: Strategy in the Global Age*, ed. T. V. Paul, Patrick M. Morgan, and James J. Wirtz (Chicago: University of Chicago Press, 2009), 304–20.

inverse of disconnection, seeking to persuade the target that cyber systems will not be unplugged.

To sum up, latency exploits the ambiguity between cyber exploitation and attack to create fears of harmful punishment if the target doesn't comply with demands. Extortion denies the target's use of cyber resources for blackmail purposes. Seduction uses deceptive techniques to lure the target into a position where compellent punishment and denial will be more effective. Escalation uses cyber aggression to signal a risk of more punishing consequences in other domains. Protection of cyber assets improves the credibility of military threats in other domains. This whole discussion shows how closely related deterrence and compellence are. Deterrence can facilitate aggression by blocking counterattack. This may be especially the case for cyber operations, particularly for strong nations like the United States that can use the deterrent threat of military retaliation to cover the employment of offensive cyber tools. As with all forms of coercion, cyber compellence appears to be more complicated than deterrence, even as they share many constraints and signals. For both deterrence and compellence, cyber coercion is most effective when employed as a complementary adjunct with capabilities to hurt in other domains. The capacity of cyber means to operate as a substitute is highly constrained and effective mostly for aggression of the irritant class.

Misperception

No discussion of coercion would be complete without some attention to the psychological dimension. After all, coercion seeks to influence the mind of the opponent by generating credible signals of future costs and benefits associated with taking (avoiding) certain behavior. These signals can be missed, garbled, or misinterpreted. If, as we argue, cyberspace is a domain that is

especially conducive to deception, then misperception is liable to be a major problem.[46] Cyber operations characterized by secrecy and ambiguity often lack the clarity associated with, say, a military invasion of territory. This secrecy and duplicity inherent in cyber operations makes it easy to misinterpret signals, if indeed it is possible to signal at all. There is an emerging consensus in political science that uncertainty is a major—if not the major—cause of war.[47] Uncertainty in the cyber domain makes it unstable, especially at the low end, even as improved knowledge of the likely costs of war in other domains places a bound on conflict at the higher end. Cyber operations are most useful in an intelligence role, and they can potentially convey information about interests (including the willingness to escalate) without actual fighting. Deliberate collection and pervasive leaks all enhance transparency, which should make conflict less likely. Furthermore, cyberspace is a manmade construct of commonly embraced and mutually beneficial protocols—interoperability is the condition for the possibility of cyber operations—so states collaborate in making internet infrastructure more stable and reliable.[48]At the same time, the secrecy of cyber operations, their complexity, and the asymmetry of access to high quality intelligence may only exacerbate problems of uncertainty.

Misperception has the potential to cause inadvertent escalation. In a crisis combining secret OCO with threats to act in other military domains, adversaries may end up inflicting and accepting greater costs than they would have been willing to pay at the start of the crisis. For

---

[46] Rose McDermott, "Decision Making Under Uncertainty," in *Proceedings of a Workshop on Deterring Cyberattacks: Informing Strategies and Developing Options for U.S. Policy*, ed. National Research Council (Washington, D.C.: National Academies Press, 2010), 227–42.

[47] James D. Fearon, "Rationalist Explanations for War," *International Organization* 49, no. 3 (1995): 379–414; Erik Gartzke, "War Is in the Error Term," *International Organization* 53, no. 03 (1999): 567–87; Jeffrey M. Kaplow and Erik Gartzke, "Knowing Unknowns: The Effect of Uncertainty in Interstate Conflict" (International Studies Association, 56th Annual Convention, New Orleans, 2015).

[48] Lindsay, "The Impact of China on Cybersecurity"; Daniel W. Drezner, "The Global Governance of the Internet: Bringing the State Back In," *Political Science Quarterly* 119, no. 3 (2004): 477–98.

example, the use of cyber attacks to degrade an opponent's strategic command and control systems, either as part of a coercive risk strategy or to facilitate limited conventional strikes, could put the opponent into a "use it or lose it" situation.[49] The dangers of inadvertent escalation through the cyber domain are thought to be particularly salient in a US-China conflict scenario due to dangerous combination of US preferences for aggressive C4ISR counterforce operations, the consolidation of nuclear and strategic missile forces in the PLA Rocket Force, Chinese and American convictions that cyber warfare must be used preemptively, and nationalist pressures in both states to retaliate for costly losses.[50]

Escalatory danger is theoretically most stark when the instability of the cyber domain (a result of ambiguity and deception) is combined with the relative stability of the nuclear domain (a result of the transparently horrific costs of nuclear war). OCO penetration of nuclear command and control could lead one side to run greater risks during a brinksmanship contest in the knowledge that the other side's capabilities are degraded, while the other side remains resolved to resist in false confidence that its nuclear deterrent remains intact. The counterforce advantage of OCO depends on its secrecy and thus cannot be revealed, but this erodes the signaling advantages of nuclear weapons which must be revealed. This tension between "winning" and "warning" recurs in many instances of cross-domain deterrence but is dangerously extreme in the cyber-nuclear combination.[51] Fortunately such a crisis is also extremely unlikely (although, unfortunately, slightly less unlikely than during the Cold War).

---

[49] Barry R. Posen, *Inadvertent Escalation: Conventional War and Nuclear Risks* (Ithaca, N.Y.: Cornell University Press, 1991).
[50] David C. Gompert and Martin Libicki, "Cyber Warfare and Sino-American Crisis Instability," *Survival* 56, no. 4 (2014): 7–22; Avery Goldstein, "First Things First: The Pressing Danger of Crisis Instability in U.S.-China Relations," *International Security* 37, no. 4 (2013): 49–89.
[51] Erik Gartzke and Jon R. Lindsay, "Thermonuclear Cyberwar," *SSRN Working Paper*, July 5, 2016.

The Stability-Instability Paradox in Cyberspace

In a half-century of internet history we have seen far more cyber espionage and crime than disruption and warfare, and more marginal irritants than operational complements. The attacks that do occur are usually minor and reversible, like website defacement and service denial, rather than serious and destructive, like attacks on industrial control systems. The distribution of actual harms inflicted through cyber operations reflects a lot of minor aggressions and very little of anything major.

This trend is partly explained by barriers to entry. We distinguished irritant from complement cyber harms based on the low costs (and low rewards) of the former and the higher resource and effort requirements (and potentially higher rewards) of the latter. Anyone can get into cybercrime. It is a highly differentiated market to facilitate easy entry and exchange. By the same token, it's hard for most cybercriminals to make much money. High-end espionage is more complicated because it focuses on particular targets and heterogeneous networks rather than indiscriminate predation on homogenous assets. Even successful intrusions do not translate into successful absorption and application, thus fears of erosion of competitive advantage are largely mythical. Cyberwarfare (disruption) is more complicated still and, for now, is mainly only a nation state competency. It is possible to imagine sophisticated non-state groups overcoming the technical, organizational, and intelligence hurdles to conduct seriously disruptive attacks, even as systemic paralysis remains out of reach for all. No one cannot take anonymity for granted if the offense is serious. Moreover, all digitally-savvy actors, including terrorists, are more likely to find the internet useful in an adjunct role supporting other types of operations rather than as a vector for staging a major attack.[52] Apart from the inherent complexity of attack planning,

[52] David C. Benson, "Why the Internet Is Not Increasing Terrorism," *Security Studies* 23, no. 2 (2014): 293–328.

moreover, defenders also tend to invest more effort protecting high-reward targets than they do against low-reward targets, which further exacerbates the differential barriers to entry.

Operational costs are only half the story, however. Deterrence also helps to explain the distribution of cyber aggression. Restraint is built into strategic interaction in cyberspace, under most conditions. OCO as a direct harm (i.e., brute force to change the balance of power) is likely to stay beneath the threshold at which the harm inflicted relative to the benefits of connection is great enough to trigger disconnection. The exception is certain situations where pervasive surveillance is used to compel miscreants to avoid using the internet altogether, thus realizing some benefit from their disconnection. Usually, however, political economic actors want to continue to be able to use the internet productively even as they cheat at the margins of mutual agreement about the benefits of connectivity. Cyber attack as a direct harm will likewise be contained to situations where the disruption of computation is minor enough so as not to trigger cross-domain retaliation for serious loss of life or incapacitation of critical infrastructure, or in situations where cyber disruption provides a tactical window of opportunity for a broader combined arms military operation. Cyber options are attractive as substitutes for force only when they are calibrated to enable the attacker to realize some benefit without the exposure to risks that the use of force usually involves. Limited cyber attacks are most appealing to those who have the capacity to conduct them when aggressors are deterred from using more violent measures.

The combination of cross-domain deterrence and voluntary connection to the internet gives rise to a variant of the classic stability-instability paradox. In Glenn Snyder's original articulation, mutually assured destruction could deter nuclear war but was not credible for, and

might even encourage, limited conventional or proxy conflict.[53] According to Robert Jervis, "To the extent that the military balance is stable at the level of all-out nuclear war, it will become less stable at lower levels of violence."[54] Cyber capacity is a poor substitute for nuclear weapons, myths of paralysis notwithstanding, yet there is a similar logic constraining the distributions of harms which are possible via information technology. To extend this logic to the cyber domain, a variety of deterrent mechanisms contain the most disruptive types of cyber attacks yet fail to contain, and even enable, a wide variety of online espionage, subversion, symbolic protest, and criminal predation. In cyberspace we observe a rather stable damage contest (i.e., no paralysis and only limited disruption) but a very unstable intelligence-counterintelligence contest (lots of espionage and fraud vying with efforts at control and mobilization). Thus the actors that have the ability to carry out highly destructive cyber attacks (mainly state actors for now) lack the motivation to attack. Yet these same actors as well as many others have both the ability and motivation to inflict irritant aggression with little fear of suffering consequences By and large, cyber options fill out the lower end of the conflict spectrum where deterrence is not as credible or reliable. The very few cases of physically disruptive cyber attack we do observe—mainly powerful states conducting covert action, subversive propaganda, or battlefield support operations against militarily weaker opponents—have notably involved stronger actors who not only have the capacity to plan and conduct a sophisticated attack but also have the ability to deter retaliation against their use of OCO.

The cyber variant of the stability-instability paradox has a slightly different logic, however. In the nuclear realm, actors cannot disconnect from the threatened harm, and this is

---

[53] Glenn H Snyder, *Deterrence and Defense: Toward a Theory of National Security* (Princeton, N.J.: Princeton University Press, 1961).
[54] Robert Jervis, *The Illogic Of American Nuclear Strategy* (Ithaca, NY: Cornell University Press, 1984), 31.

what makes the threatened destruction both mutual and assured. When there are many missiles with many warheads, the chance of intercepting them on the ground through a disarming counterforce strike or in the air through ballistic missile defense with any confidence is depressingly small. Not so in cyberspace, where connection to the internet and acceptance of connections through it is voluntary. Attackers thus rely on deception to exploit vulnerabilities and ensure they can access their targets. However, offensive deception can fail in the fog of cyberwar, and defenders can be deceptive as well; both are more likely with high-reward targets, where cross domain deterrence is also more credible. The need to preserve internet connections to facilitate ongoing and future deception as well as the need to preserve stealth to avoid the consequences of getting caught imposes discipline on attackers.

Actors cannot enjoy the substantial benefits of interconnection without accepting some risk of exploitation and attack. Perfect defense (internet lockout) with advantage accruing exclusively to one political group or another, is not feasible. Moreover, because cyber harms share similar techniques, the observed abundance of exploitation represents a latent potential for attack. The latent escalatory potential of even minor irritants leads to rampant fears of unrestrained catastrophe, to be sure. Yet this latent potential is difficult to harness for targeted coercion because the threat is self-effacing. Declared cyber threats that highlight the target's vulnerability to exploitation are readily mitigated. Instead, the ineradicable threat of cyber catastrophe, so long as the internet continues to be useful, creates a general if diffuse deterrent effect among all parties who value their connection to the internet. No one who wants to make money on the internet really wants to have a cyberwar.

Large powers like the U.S. are highly dependent on the internet but also highly skilled at inflicting harm through a variety of means. Poor powers across the digital divide may have a

smaller digital attack surface, while medium powers may have vulnerability but lack a range of forces to deter attacks. This might imply a curvature to the utility of cyber coercion. Strong and capable countries are vulnerable to cyber harm but can deter through other military instruments. Poor states are not vulnerable. It may be the prosperous small or digitally developing who are in the most trouble, since they cannot credibly deter and are highly dependent on the internet. The cases of Estonia and Ukraine are suggestive. The information revolution is often thought to be a boon to non-state actors, and indeed it is, but mainly in the irritant class of cyber operations. Moreover, the increasing ubiquity and sophistication of information technologies can be expected to have something of a democratizing effect on intelligence and counterintelligence techniques whereby firms and citizens will have access to and be concerned about the types of things that were historically the purview of obscure state intelligence organs. However, it would be a mistake to use the increasing ferment of low-intensity information contests to infer the shape of higher intensity activity. On the contrary, the traditional logic of war will continue to dominate the expression of cyber aggression.

Because threatened internet harms depend on voluntary connections in the first place, and as many actors have alternative means to inflict (cross-domain) harm in retaliation, the coercive utility of cyberspace is actually somewhat limited. At the same time an ever increasing variety of irritants and more temperamental operational complements becomes available for political interaction. The “net” result is that opponents have strong incentives to impose costs via the internet but also to keep those costs low enough to preserve interconnection and avoid retaliation. Therefore, contests in damage will remain relatively stable while contests in intelligence will be increasingly unstable. The human-built world is becoming more complex, to

be sure, but it is not necessarily more dangerous. As long as it is desirable to connect to the internet tomorrow, there will be only limited harm via the internet today.