TECHNOLOGY FOR GLOBAL SECURITY

CYBER OPERATIONS AND NUCLEAR WEAPONS

TECHNOLOGY FOR GLOBAL SECURITY SPECIAL REPORT

DR. JON LINDSAY
PROFESSOR AT THE MUNK SCHOOL OF GLOBAL AFFAIRS & PUBLIC POLICY;
DEPARTMENT OF POLICITAL SCIENCE AT THE UNIVERSITY OF TORONTO
June 20, 2019

CYBER OPERATIONS AND NUCLEAR WEAPONS

JON R. LINDSAY

JUNE 20, 2019

I. INTRODUCTION

In this essay, Jon Lindsay argues that: "As NC3 increasingly uses digital technologies to enhance efficiency and reliability, the cybersecurity of NC3 becomes a pressing concern. Adversaries have incentives to penetrate NC3 for intelligence in peacetime and for counterforce in wartime. Given the broad diffusion of cyber capabilities, furthermore, most nuclear weapon states also have some ability to do so, although the operational difficulties of gaining remote access to and covert control over NC3 cannot be overstated. Offensive cyber operations targeting NC3 introduce a number of underappreciated risks of organizational breakdown, decision making confusion, and rational miscalculation in a nuclear crisis."

Jon R. Lindsay is assistant professor at the Munk School of Global Affairs & Public Policy and the Department of Political Science at the University of Toronto.

Acknowledgments: The workshop was funded by the John D. and Catherine T. MacArthur Foundation.

A podcast with Jon Lindsay, Peter Hayes, and Philip Reiner on cyber operations and nuclear weapons is found here.

The views expressed in this report do not necessarily reflect the official policy or position of Technology for Global Security. Readers should note that Technology for Global Security seeks a diversity of views and opinions on significant topics in order to identify common ground.

Banner image is by Lauren Hostetter of Heyhoss Design.

CITATION

Lindsay R. Jon, "Cyber Operations and Nuclear Weapons," Tech4GS Special Reports, June 20, 2019, https://www.tech4gs.org/nc3-systems-and-strategic-stability-a-global-overview.html

II. TECH4GS SPECIAL REPORT BY JON R. LINDSAY
CYBER OPERATIONS AND NUCLEAR WEAPONS
JUNE 20, 2019

Summary

A nuclear weapon alone does not a deterrent make. A system of supporting technologies and organizations is required to make credible nuclear threats. Operational warheads, delivery platforms, and early warning satellites must be linked to political leaders and military commanders via a nuclear command, control, and communication (NC3) network. As NC3 increasingly uses digital technologies to enhance efficiency and reliability, the cybersecurity of NC3 becomes a pressing concern. Adversaries have incentives to penetrate NC3 for intelligence in peacetime and for counterforce in wartime. Given the broad diffusion of cyber capabilities, furthermore, most nuclear weapon states also have some ability to do so, although the operational difficulties of gaining remote access to and covert control over NC3 cannot be overstated. Offensive cyber operations targeting NC3 introduce a number of underappreciated risks of organizational breakdown, decision making confusion, and rational incentives to misrepresent the balance of power in a nuclear crisis. Risks and tradeoffs in NC3 have been inherent since the Cold War, but modern information technology heightens system complexity significantly, compounding additional problems created by nuclear multipolarity and the interdependence of nuclear and conventional command and control.

Digital Strangelove

The basic paradox of nuclear deterrence is that weapons too dangerous to use in war are threatened in order to prevent war.[1] The credibility of nuclear threats is a function of many things including the capability to inflict harm on the target, the political willingness to run risks of nuclear war, and the clear communication of ability and resolve. If a country cannot persuade its enemies that it can deliver punishment when redlines are crossed, and only when they are crossed, then a deterrent threat loses credibility. Yet the basic paradox creates difficult communication problems for any strategist.

To support nuclear deterrence strategies, nuclear command, control, and communications (NC3) systems must conform to strict 'always-never' criteria. An arsenal must always be ready when orders are given by legitimate authority but never usable by unauthorized persons or on accident.

---

[1] Bernard Brodie et al., *The Absolute Weapon: Atomic Power and World Order* (New York: Harcourt, Brace and Co., 1946); Thomas C. Schelling, *Arms and Influence: With a New Preface and Afterword* (New Haven, CT: Yale University Press, 2008).

Weapons that may not be usable when needed cannot be credibly threatened, and weapons that might be used even if no redline has been crossed give the target little reason to comply with threats. Unfortunately, there is an inherent tension in the 'always-never' criteria. Weapons that are always ready to launch at a moment's notice or triggered automatically are more likely to be used inadvertently. Weapons that are decoupled from NC3 and covered with safeguards against unauthorized launch are more likely to be destroyed or deactivated during a nuclear crisis. NC3 systems attempt to balance the dilemma with redundant communications, multiple contingency plans, and standard operating procedures drilled into military organizations or implemented in the routines of digital software.[2]

In the classic satire *Dr. Strangelove*, the Soviets over-optimize on the 'always' criterion by building a doomsday device that will be triggered automatically upon American attack, and the Americans over-optimize by delegating too much authority to a mentally unbalanced wing commander (General Ripper) and an indefatigable bomber pilot (Major Kong). Through a series of unfortunate events, they collectively end up violating the 'never' criterion through an unauthorized launch ordered by General Ripper and the unintentional detonation of the doomsday machine. None of the characters in the film fully understand how their tightly-coupled complex NC3 system will behave in the fog of war. All of the deterministic routines, communication protocols, and failsafe procedures designed to make NC3 more reliable, sadly end up making it less reliable.

Modern digital systems are not immune from these problems, and in some ways their complexity heightens the danger. The mad General Ripper was essentially a hacker who fooled loyal airmen into becoming agents of the apocalypse. Offensive cyber operations can, likewise, target NC3 to turn perfectly deterministic computers into spies or saboteurs. As if to drive home the point, the eponymous Dr. Strangelove explains that the doomsday weapon is "connected to a gigantic complex of computers [in which] a specific and clearly defined set of circumstances, under which the bombs are to be exploded, is programmed into a tape memory bank." Such a deterrent is terrifying, "because of the automated and irrevocable decision-making process which rules out human meddling." Strangelove assumes that Soviet NC3 systems are highly reliable, but his assumption would not be warranted if hackers could exploit a remote connection into the doomsday supercomputer. They might disable a remote deterrent by disconnecting sensors that detect incoming attacks or disabling the commands that trigger counterattacks. They might trigger the apocalypse on purpose if they are nihilists, or accidentally if they make mistakes during their intrusion, or, if they really have great intelligence and technical skill, they might trigger just enough bombs to doom the enemy but spare themselves. Another Cold War classic, *WarGames*, speculated that a hoodie-wearing teenager might inadvertently initiate World War III. A screening of *WarGames* allegedly prompted Ronald Reagan to formulate the first executive order for government cybersecurity.[3]

Two distinct types of risk emerge from the combination of cyber operations and nuclear weapons. The first is a function of the confusion and uncertainty created by sociotechnical complexity. All people make mistakes, and because hackers are people too, even the best hackers

---

[2] Paul J. Bracken, *The Command and Control of Nuclear Forces* (New Haven, CT: Yale University Press, 1983); Christopher A. Ford, "Playing for Time on the Edge of the Apocalypse: Maximizing Decision Time for Nuclear Leaders," in *Deterrence: Its Past and Future—Papers Presented at Hoover Institution, November 2010*, ed. George P. Shultz, Sidney D. Drell, and James E. Goodby (Stanford, CA: Hoover Institution, 2011).

[3] Fred Kaplan, "'WarGames' and Cybersecurity's Debt to a Hollywood Hack," *The New York Times*, February 19, 2016.

make mistakes. Failures to fully understand the consequences of complexity can lead to unintended consequences in the behavior of tightly coupled NC3 systems and the software code that attacks or defends it.[4] The second and more insidious type of risk stems from rational incentives to escalate created by cyber operations. Secret cyber operations undermine the clear communication on which credible deterrence depends.[5] As Dr. Strangelove chides the Soviet ambassador, "the whole point of the doomsday machine is lost if you keep it a secret!" Information asymmetries created by cyber operations can narrow or close the crisis bargaining range, which in effect trades the ability to limit damage in a nuclear war for greater risks of starting a nuclear war. The interdependence of conventional and nuclear command and control systems may also create rational incentives to escalate to the nuclear level once a conventional war begins.[6]

Advanced technologies make familiar tradeoffs more complex. Digital systems have the potential to improve the 'always-never' criteria that NC3 must meet. Unfortunately, the price of more sophisticated NC3 is more complex failure modes and a larger attack surface. Hackers that cripple enemy NC3 undermine the 'always' criterion for the enemy, but their accomplishment cannot revealed. Glitches and confusion in NC3 or the intrusions into it undermine the 'never' criterion, but the risks are hard to understand. *Dr. Strangelove* remains disconcertingly relevant in the 21st century because the risks of accident and escalation are ultimately rooted in human organizations and strategic incentives rather than the vintage of information technology.

This paper provides an overview of the problems inherent in studying and managing the interactions between cyber and nuclear weapons. It first sketches the cyber vulnerabilities to NC3 posed by various political actors and then contrasts the operational, strategic, and proliferation characteristics of the cyber and nuclear domains. Finally it catalogues the potential escalation pathways activated by their combination, distinguishing pathways that arise through confusion from those that arise through strategic interaction.

Cyber Vulnerabilities of Nuclear Weapons

The modern nuclear weapons enterprise connects some of the most complex computational systems ever built to the most dangerous weapons in history. NC3 is the nervous system of nuclear deterrence. It includes, or interacts through, several different components:

- Early warning satellites and radars detect events like enemy missile launches that indicate whether or not deterrence has failed.
- Command and control systems aggregate intelligence data for political leaders and military commanders and enable them to send instructions to operational units in the field. Secure, authenticated, redundant communication networks tie everything together.

[4] Stephen J. Cimbala, *Nuclear Weapons in the Information Age* (Continuum International Publishing, 2012); Andrew Futter, *Hacking the Bomb: Cyber Threats and Nuclear Weapons* (Washington, DC: Georgetown University Press, 2018).

[5] Erik Gartzke and Jon R. Lindsay, "Thermonuclear Cyberwar," *Journal of Cybersecurity* 3, no. 1 (February 2017): 37–48.

[6] Avery Goldstein, "First Things First: The Pressing Danger of Crisis Instability in U.S.-China Relations," *International Security* 37, no. 4 (2013): 49–89; James M. Acton, "Escalation through Entanglement: How the Vulnerability of Command-and-Control Systems Raises the Risks of an Inadvertent Nuclear War," *International Security* 43, no. 1 (August 1, 2018): 56–99, https://doi.org/10.1162/isec_a_00320.

- Operational units deployed on land, in the air, or under the waves must be able to receive and authenticate instructions and send back status reports. Platform diversity and dispersal ensures that at least some nuclear forces can survive an enemy attack and be available for retaliation, i.e., a secure second strike.
- National intelligence reporting and political assessment functions, while not strictly part of the NC3 enterprise, enable leadership to determine the context, nature, extent, and stakes of the crisis.
- Missile defense systems, also not strictly part of NC3 but important in the broader context of nuclear deterrence and warfighting, provide an additional, if slight, layer of protection for friendly weapons and cities.
- All of the sensors, data processing systems, communication links, and weapons platforms used in NC3 are the product of an upstream research, development, test, evaluation, and procurement system that must also be monitored for reliability and security.

Advanced information technologies can improve NC3. An expanded range and number of sensors can improve early warning while the ability to aggregate and analyze data from sensors and other information sources can improve the ability to distinguish true warnings from false indications. Robust networks can connect multiple command centers and dispersed weapon systems to improve survivability and enhance collective situational awareness. Better cryptographic authentication protocols can improve trust among people in the system. High fidelity targeting data and precision guided weapons can increase confidence in destroying enemy targets or intercepting enemy missiles. Monitoring, reporting, auditing, and authentication schemes enable network operators to detect and correct data processing problems.

Unfortunately, NC3 systems have vulnerabilities. Nuclear safety experts have compiled a disconcerting list of computer glitches, loose components, early warning radar faults, and human mistakes that resulted in close calls but also reveal the potential for malicious interference.[7] In 2013 the commander of U.S. Strategic Command stated he was "very concerned with the potential of a cyber-related attack on our nuclear command and control and on the weapons systems themselves."[8]

Concerns about offensive cyber operations targeting NC3 are historically plausible. The U.S. military has long been committed to counterforce, reflected by decisions about force structure, posture, and operational doctrine.[9] A Cold War program known as Canopy Wing combined electronic warfare and information operations to degrade Soviet nuclear and conventional control; according to Warsaw Pact officials who became aware of the program through espionage, it "sent ice-cold shivers down our spines."[10] The revelation of Stuxnet demonstrated American and Israeli willingness to conduct intrusions into and disrupt nuclear infrastructure

---

[7] Shaun Gregory, *The Hidden Cost of Deterrence: Nuclear Weapons Accidents* (London: Brassey's, 1990); Scott D. Sagan, *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons* (Princeton, NJ: Princeton University Press, 1995); Eric Schlosser, *Command and Control: Nuclear Weapons, the Demascus Accident, and the Illusion of Safety* (New York: Penguin, 2014).

[8] "Hearing to Receive Testimony on U.S. Strategic Command and U.S. Cyber Command in Review of the Defense Authorization Request for Fiscal Year 2014 and the Future Years Defense Program," § U.S. Senate Committee on Armed Services (2013), p. 10.

[9] Austin Long and Brendan Rittenhouse Green, "Stalking the Secure Second Strike: Intelligence, Counterforce, and Nuclear Strategy," *Journal of Strategic Studies* 38, no. 1–2 (2014): 38–73.

[10] Benjamin B. Fischer, "CANOPY WING: The U.S. War Plan That Gave the East Germans Goose Bumps," *International Journal of Intelligence and CounterIntelligence* 27, no. 3 (2014): 439.

(and also highlighted the operational challenges of crafting and conducting remote covert sabotage).[11] The Israelis reportedly used cyber operations to disable Syrian air defenses during their 2007 raid on a Syrian nuclear facility.[12] The United States may have conducted cyber and electronic warfare operations to "remotely manipulate data inside [the Democratic People's Republic of] Korea's missile systems."[13] While this was reportedly part of a U.S. counterproliferation initiative, this episode highlights the potential utility of employing cyber and electronic warfare measures "left-of-launch" in an actual warfighting scenario.

[11] Jon R. Lindsay, "Stuxnet and the Limits of Cyber Warfare," *Security Studies* 22, no. 3 (2013): 365–404.
[12] David A. Fulghum, "Why Syria's Air Defenses Failed to Detect Israelis," *Aviation Week, Ares Blog* (blog), October 3, 2007.
[13] David E. Sanger and William J. Broad, "Trump Inherits a Secret Cyberwar Against North Korean Missiles," *The New York Times*, March 4, 2017.

Table 1: Cyber vulnerabilities of nuclear weapon systems and countermeasures

Nuclear Segment	**Attack Vectors**	**Consequences**	**Countermeasures**
Research, development, testing, simulation, maintenance	Supply chains, manufacturing, supercomputers, technical personnel	Unreliable systems, readiness delays, financial expenditure	Onshore logistics, redundant supplies, redundant verification protocols, cyber and physical security best practices
Early warning satellites & radars	Directed energy, communication links, ground station processing	False positives (spoofing), false negatives (blinding)	Redundant sensors, multiple phenomenology, all-source intelligence fusion
Crisis intelligence & assessment	Deception, social media flooding & manipulation, false flag, jamming	Confusion, misattribution, threat inflation, centralization error	Counterintelligence, active defense, all-source fusion, multiple advisors, public affairs strategy, crisis hotline
Command, control, communications, computing architecture	Access, authentication, confidentiality, integrity of network operations	Unauthorized launch, accidental launch, launch failure, targeting error	Redundant communication & authentication, limited connections, heterogeneous systems, network monitoring, cybersecurity best practices
Operational units, delivery platforms & warheads	Supply chain, C3 network, telemetry	Launch failure, guidance failure, self-destruction	Redundant systems (overkill), testing, authentication
Missile defense	Supply chain, sensors, C3 network	Detection, tracking, interception failure	Redundant sensors, networks, interceptors; inspections, cybersecurity best practices

It is beyond the scope of this paper to examine all of the technical vulnerabilities of NC3 and potential consequences of compromise.[14] Table 1 provides a brief summary of potential threats and consequences, together with defensive measures that might be taken to mitigate the exposure to vulnerabilities, of different segments of the nuclear enterprise. These are listed in roughly temporal order, from the advanced preparation of forces, to early warning, command and control, operational forces, and missile defense.

[14] For more detailed assessment see Page O. Stoutland and Samantha Pitts-Kiefer, "Nuclear Weapons in the New Cyber Age: Report off the Cyber-Nuclear Weapons Study Group" (Washington, DC: Nuclear Threat Initiative, September 2018); Beyza Unal and Patricia Lewis, "Cybersecurity of Nuclear Weapons Systems: Threats, Vulnerabilities and Consequences" (London: Chatham House, the Royal Institute of International Affairs, January 2018).

Redundancy is a recurrent defensive principle across categories: parallel or backup sensors, communications links, authentication protocols, and weapons can help to identify and overcome errors caused by malicious hacking. Insofar as cyber-attacks target single points of failure or corrupt key data, reducing the number and exposure of critical nodes and increasing error detection and correction should be a priority. System heterogeneity is another principle that can be used to avoid technical monocultures that are easier to target, although the system engineering integration challenges mount considerably with heterogeneity. Redundancy is expensive, of course, and often imposes a processing burden on human organizations. The 'always-never' tradeoff rears its head again.

Comparing the Nuclear and Cyber Domains

The cyber and nuclear domains are very different. In many ways they are complete opposites. This section compares operational, strategic, and proliferation characteristics. Operational factors (Table 2) describe the nature of the capability, to include the weapon itself and the organizational capacity needed to wield it. Strategic factors (Table 3) describe the nature of interaction between political competitors. Proliferation factors (Table 4) describe the supply and demand side factors that affect the ability and willingness of different types of actors to acquire the capability. Understanding all three is important for understanding the increasingly multipolar cyber and nuclear arenas, namely what types of actors are likely to acquire capabilities and employ them in different situations.

While cyber warfare is sometimes likened to a new weapon of mass destruction capable of great societal disruption, this comparison is misleading.[15] Comparing operational capabilities (Table 2), nuclear weapons create severe and irreversible damage, and they do it very quickly. Cyber operations generally create little to no damage, and once discovered their effects can often be remediated. Moreover, most cyber operations are conducted over a long period, maintaining persistence where they can collect information or exert influence indefinitely (quite in contrast to the popular trope of cyber at the speed of electrons). Defense against nuclear warheads speeding in from outer space is extremely difficult, and the costs of missing even one are extremely high (a city lost). Missile defenses are improving but still unreliable and "left of launch" cyber and electronic warfare methods are fraught with difficulty, as discussed presently. Cyber defense is also widely thought to be difficult (and relatively harder than offense), since attackers craft their intrusions to evade defenses and complicate attribution; however, network monitoring and counterintelligence techniques can be effective and the offense-defense balance starts to look more favorable to defense against complex, high-value targets.[16]

One striking difference in the management of the two domains is the centralization of most states' nuclear enterprise to include an elaborate regime of tests and exercises and to realize 'always-never' criteria. Cyber command and control tends to be more distributed, leveraging networks of compromised computers in the target network and possibly on the public internet.

---

[15] Joseph S Nye Jr., "Nuclear Lessons for Cyber Security?," *Strategic Studies Quarterly* 5, no. 4 (2011), http://www.au.af.mil/au/ssq/2011/winter/nye.pdf.

[16] Rebecca Slayton, "What Is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment," *International Security* 41, no. 3 (January 1, 2017): 72–109; 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.

Cyber operations rely on the commercial technologies and shared connections of a global domain rather than the specialized and

Table 2: Comparison of operational factors in the nuclear and cyber domains

Operational Factors	**Nuclear Weapons**	**Offensive Cyber Operations**
Damage mechanism	Direct blast, fire, radiation	Indirect influence & espionage
Damage severity	Extreme—cities, populations, planet	Generally low, higher levels possible
Reversibility	Low—devastation, lingering effects	High—repair, redirect, reinstall
Delivery	Stand-off weapons, missiles	Remote access, social engineering
Payload	Thermonuclear detonation	Control, persistence, exfil, disrupt
Feasibility of defense	Limited—low probability intercept	Hard but possible—counterintel
Timeline	Rapid delivery, instant devastation	Tactically fast, operationally slow
Command & control	Centralized	Decentralized
System complexity	Complicated but well-characterized	Extreme diversity & connectivity
System jurisdiction	National command authority	Globalized civil society
Targeting	Deliberate planning process	Detailed intel and extensive planning
Reliability	Production, testing, models, exercise	Custom engineering each use

highly controlled technologies of the nuclear realm. Nuclear weapons can be tested on an instrumented range, or computationally simulated with precise scientific models, to create confidence in each standardized weapon design, but almost every cyber operation relies on custom engineering for each new target and considerable intelligence preparation (potentially including human intelligence) to tailor the intrusion and payload for novel circumstances. Operational aspects inform but do not necessarily determine the strategic utility of particular types of weapons. Utility ultimately depends on incentives for action or restraint given expectations of the other player(s)'s choices. How should we understand incentives? Strategic concepts for nuclear weapons are more mature than for cyber, which is not surprising since the strategic studies community in many ways was a product of the nuclear revolution. Furthermore, the problems of intelligence and subterfuge have received comparatively little attention in the field of international relations (but this is changing).

Table 3: Comparison of strategic factors in the nuclear and cyber domains

Strategic Factors	**Nuclear Weapons**	**Offensive Cyber Operations**
Intellectual concepts	Mature	Immature
Informational quality	Can and must be revealed	Deceptive methods require secrecy
Commitment problem	High costs of retaliation	Revelation undermines capability
Political usefulness	Enhance status quo stability	Marginal revision of power balance
Warfighting utility	Limited (counterforce school)	Force multiplier (C4ISR)
Deterrence utility	Credible protection of vital interests	Limited (reputation for cyber ability)
Compellence utility	Less credible for revising status quo	Limited (ransomware)
Interdependence	Mutual exposure	Shared protocols, networks, trade
Political logic	Mutually assured destruction	Intelligence & covert action
Conflict spectrum	Extreme high-end war	Low end-gray zone
Incentives for restraint	Nuclear retaliation and escalation	Loss of access & cross-domain action
Cross-domain effects	Cyber weakens nuclear deterrence	Nuclear bounds on cyber severity

The nuclear and cyber domains have markedly different informational characteristics. The basic strategic facts about nuclear weapons are easily appreciated. The high costs of nuclear use and potential retaliation make nuclear threats most credible when used to defend vital interests like regime survival but less credible for compellent threats that seek to revise the status quo.[17] Importantly, nuclear weapons can and must be revealed to establish a credible deterrent threat. By contrast, cyber intrusions cannot be revealed if the attacker wants them to remain viable. The 'cyber commitment problem' refers to the unsuitability of cyber means for coercion due to the fact that revelation of the threat, which is needed to separate credible threats from bluffs, would enable the target to patch or reconfigure its networks to mitigate the threat.[18] The imperative to maintain access to target systems and preserve the viability of shared networks is a reinforcing

[17] Todd S. Sechser and Matthew Fuhrmann, *Nuclear Weapons and Coercive Diplomacy* (New York: Cambridge University Press, 2017). Cf., Matthew Kroenig, *The Logic of American Nuclear Strategy: Why Strategic Superiority Matters* (New York: Oxford University Press, 2018).

[18] Erik Gartzke and Jon R. Lindsay, "The Cyber Commitment Problem and the Destabilization of Nuclear Deterrence," in *Bytes, Bombs, and Spies: The Strategic Dimensions of Offensive Cyber Operations*, ed. Herbert S. Lin and Amy B. Zegart (Washington, DC: Brookings Institution Press, 2019).

source of restraint in the cyber domain distinct from the fear of retaliation (usually in different domains, e.g., economic sanctions or conventional military response).[19]

These extreme informational differences tend to push nuclear and cyber operations to the opposite ends of the conflict spectrum. Nuclear weapons threaten total war to protect vital interests. Cyber operations pursue marginal revision in the distribution of power and benefits by conducting intelligence and covert action in peacetime as well as war (where they can be a force-multiplier for action in other domains like land, air, and sea). The interaction of the cyber and nuclear domains has different consequences depending on which is the main domain of action. Nuclear threats can serve implicitly or explicitly to bound the severity of cyber attacks (as suggested by the 2018 Nuclear Posture Review from the Trump administration). Yet NC3 targeting cyber operations, which rely on deception, tend to undermine the stability of nuclear deterrence, which relies on transparency.[20] Anything that contributes to information asymmetry regarding the balance of power and resolve is a potential source for bargaining failure and war.[21]

The operational and strategic factors result in very different proliferation dynamics (Table 4). The operational factors (Table 2) imply quite different financial, technical, organizational, and market barriers to entry. Cyber operations do require a little more than just technical expertise and an internet connection, namely the organizational capacity to collect intelligence and conduct covert activity, but this is slight compared to the scientific and military infrastructure required for nuclear weapons. The strategic factors (Table 3) result in additional disincentives to acquire nuclear weapons, namely the risk of preventative war and sanctions from powers who would prefer not to be coerced by them.[22] Deterrence and counterproliferation might be backed up by a normative system of arms control treaties and inspection regimes. By contrast, there are few strategic disincentives to the acquisition of cyber capabilities and many incentives for states looking for a way to work around the strengths of their adversaries. Given the reliance on deception for cyber operations, cyber arms control proposals are inherently incredible (i.e., they require participants to promise not to lie, which is a good way to cover lying behavior).

[19] Jon R. Lindsay, "Restrained by Design: The Political Economy of Cybersecurity," *Digital Policy, Regulation and Governance* 19, no. 6 (2017): 493–514.

[20] Gartzke and Lindsay, "Thermonuclear Cyberwar."

[21] 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.

[22] Nuno P. Monteiro and Alexandre Debs, "The Strategic Logic of Nuclear Proliferation," *International Security* 39, no. 2 (October 1, 2014): 7–51.

Table 4: Comparison of proliferation factors in the nuclear and cyber domains

Proliferation Factors	**Nuclear Weapons**	**Offensive Cyber Operations**
Financial barriers	Higher	Lower
Technical barriers	Expertise, fuel cycle, labs, testing	Computer science expertise
Operational barriers	Delivery systems, NC3	Intelligence & operations
Market barriers	Tightly controlled	Thriving white/gray hat market
Strategic barriers	Risk of preventative war & sanctions	None
Dual-use	Nuclear energy	Ubiquitous digital networks
Norms	"Nuclear taboo" & historic non-use	Weak norms & endemic usage
Arms control treaties	NPT, PTBT, Seabed, Space, START…	Cybercrime, economic espionage
Enforcement	Intelligence, inspections, sanctions	Infeasible: secrecy & deception
Capable actors	Nine nuclear weapon states	Many states & non-state actors
Terrorist threat	High barriers, but "loose nuke" risk	Attractive & available at low end

Cyber Capabilities of Nuclear Weapons States

The net result is that supply-side and demand-side barriers have limited nuclear proliferation to date to only nine states, and there are strong disincentives for any of them to proliferate to terrorists.[23] Cyber proliferation is comparatively unrestrained, however, especially at the low end of the conflict spectrum where cyber criminals and spies alike can and do flourish. Low-end activity, moreover, to include planting propaganda and agitating on social media, could potentially be destabilizing in a nuclear crisis if it muddied political and intelligence assessments of the situation and enemy intentions. The attribution problem, and thus the risk of misattribution, is also greater at the lower end, but then the consequences are lower too (unless admixed with crises and operations in other domains).[24]

At the high end of the conflict spectrum, which certainly includes the use of offensive cyber operations against NC3, fewer actors can be expected to have the ability and willingness to

[23] Scott D. Sagan, "The Causes of Nuclear Weapons Proliferation," *Annual Review of Political Science* 14, no. 1 (2011): 225–44, https://doi.org/10.1146/annurev-polisci-052209-131042; Keir A. Lieber and Daryl G. Press, "Why States Won't Give Nuclear Weapons to Terrorists," *International Security* 38, no. 1 (July 1, 2013): 80–104.

[24] Lindsay, "Tipping the Scales."

conduct successful operations. Targeting NC3 inevitably requires lengthy reconnaissance to gather detailed intelligence, specific technical expertise on the NC3 and weapon systems of the enemy state, testing and rehearsal to assure commanders that the operation will work as intended, and some assessment of the enemy organization's routines and human behavior to understand how the target will behave with or compensate for degraded NC3. The Stuxnet operation targeting Iranian nuclear enrichment took many years of careful preparation and lots of intelligence resources, and that was simply targeting the fuel cycle.[25] Targeting NC3 of an operational deterrent should be expected to be even more difficult and sensitive. Any sensitive targeted operation must be carefully planned and monitored via supporting command and control networks to receive feedback on the progress of the intrusion and to push updates and instructions to the attack code, and external sources of intelligence are needed to search for indications of compromise. These are difficult tasks for even mature signals and human intelligence agencies like the NSA and CIA.

The operational hurdles screen out many would be NC3 cyber operations. Yet, an actor that risks penetrating enemy NC3 will almost certainly want some sort of insurance policy (other counterforce capabilities or a robust deterrence posture) in case the operation is compromised or fails. Therefore NC3 intrusions are most likely to be conducted by other nuclear powers with the ability and willingness to backstop cyber operations with other forms of power. As a general matter, cyber operations are not as important on their own, except as a form of intelligence gathering and covert influence, which only matter on the margins. The risks of cyber-NC3 operations are only offset by their potential benefit in a nuclear warfighting scenario, or for gathering intelligence to support that eventuality. Cyber operations thus become most relevant as a force multiplier combined with other domains. Cyber-NC3 is most relevant when the attacker has nuclear weapons and cyber-NC3 destabilization is most likely to be felt in a brinksmanship crisis between nuclear rivals.

Table 5 lists the nine nuclear weapons states plus Iran, which is the state most likely to make a dash for the bomb should the Joint Comprehensive Plan of Action (JCPOA) continue to unravel following the Trump administration's default on the deal. The size and posture of the nuclear arsenals of these states vary considerably, with the former Cold War superpowers in a class unto themselves. Minimal deterrence postures with declaratory no first use pledges (e.g., India and China) are generally believed to be good for strategic stability vis-à-vis deterring nuclear attack on the home state. Most of these states, however, also hope to use their arsenals for something more, such as deterring significant non-nuclear attacks, extending deterrence to allies, or coercive diplomacy. More ambiguous and aggressive postures, which are designed to make nuclear weapons more usable in more situations, carry increased risks of strategic instability.[26]

The states also vary in their proficiency in cyber offense and cyber defense. The United States stands head and shoulders above the rest in terms of intelligence capacity and political economic advantages in cyberspace. Nuclear rivalries can be assessed for the potential for strategic instability resulting from cyber-nuclear interactions. The most dangerous situation is an asymmetric dyad where a small nuclear arsenal with weaker cyber defenses faces a stronger nuclear power with potent cyber offense capability, such as North Korea vs. Unites States or, potentially, Iran vs. Israel. Defenders with larger arsenals

[25] Slayton, "What Is the Cyber Offense-Defense Balance?"

[26] Vipin Narang, *Nuclear Strategy in the Modern Era: Regional Powers and International Conflict* (Princeton University Press, 2014).

Table 5: State nuclear and cyber capabilities

State	Warheads (deployed) [27]	Nuclear posture	Cyber Offense [28]	Cyber Defense[29]	Nuclear rivals	Cyber-Nuclear risk from rivals
USA	6450 (1750)	Extended counterforce	Top	Top 0.91 (2nd)	Russia, China, NK	Moderate
Russia	6850 (1600)	Ambiguous, triad + NSNW	Leading	Leading 0.79 (10th)	USA, China, UK, France	Moderate
China	280	Minimal, NFU	Leading	Maturing 0.62 (32nd)	USA, India, Russia	Moderate
UK	215 (120)	General	Leading	Leading 0.78 (12th)	Russia	Low
France	300 (280)	General	Maturing	Leading 0.81 (8th)	Russia	Low
India	130-140	Minimal, NFU	Maturing	Maturing 0.68 (23rd)	Pakistan, China	Low
Pakistan	140-150	Asymmetric	Maturing	Maturing 0.45 (66th)	India, Israel	Moderate
Israel	80	Ambiguous, Opaque	Leading	Maturing 0.69 (20th)	Pakistan, (Iran)	Low
N Korea	10-20	Ambiguous, Asymmetric	Maturing	Maturing 0.53 (52nd)	USA	High
Iran	0	Latency	Maturing	Maturing 0.49 (59th)	USA, Israel	High

[27] SIPRI Yearbook 2018, "World Nuclear Forces," https://www.sipri.org/yearbook/2018/06
[28] Author assessment
[29] ITU Global Cybersecurity Index 2017, https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-GCI.01-2017-R1-PDF-E.pdf. The GCI is a composite measure (with rank across countries included) of legal, technical, organizational,

or more competent cyber defense would tend to reduce the incentives for rivals to engage in NC3 cyber attacks. States like the United States, Russia, and China all face moderate risk because, while they have retaliatory forces and/or competent cyber defenses, they also face an adversary with the capability and possibly the motivation to intrude. Table 5 provides a rough assessment of the risks of nuclear escalation due to relative cyber offensive and defensive capabilities in the context of rival nuclear balances.

Cyber-to-Nuclear Escalation Scenarios

The combination of offensive cyber operations and nuclear weapons creates many different pathways for escalation. The notion of strategic stability is contested, but I use it here to refer to the marginal risk of the outbreak of nuclear war or escalation of a war to a higher level of intensity as a result of cyber-nuclear interactions.[30] The risks of breakdown in complex systems and human confusion in the fog of war receive a lot of attention in the literature on organizational reliability, and NC3 architects are generally familiar with these sorts of risks (even if they have not, or cannot, eliminate them).[31] Because NC3 systems interface directly with human decision makers, degradation of NC3 under cyber attack certainly carries the potential to degrade the quality of decision making. Most analysis of cyber-nuclear risk falls into this general category.[32]

Less appreciated but in some ways more worrisome is the potential for rational incentives to misrepresent the balance of power resulting from information asymmetry in strategic interaction created by cyber operations.[33] Cooler heads and more rational thinkers might avoid disaster in the fog of war scenarios but not in situations where rational incentives to misrepresent the truth create divergent assessments about the possible outcomes of conflict. Effective deterrence requires an actor reveal a willingness and capacity to punish the target under some particular circumstances. Deterrence, as well as negotiated settlements to crises that restore deterrence, thus depend on common knowledge about the balance of power, mutual interests, and the expected costs and outcomes of war. Computer networks increase the complexity of the system, which increases intrinsic uncertainty, and hackers rely on deception to exploit NC3, which increases strategic information asymmetry. Uncertainty and deception both undermine common knowledge. Any daylight between the ways in which actors in a crisis understand each other and their actual capabilities and interests creates possibilities for accident and miscalculation.

---

capacity building, and cooperation to enhance national cybersecurity for both the government and private sector. Singapore is the only state to receive a higher GCI than the United States.

[30] Generally see Elbridge A. Colby and Michael S. Gerson, eds., *Strategic Stability: Contending Interpretations* (Carlisle Barracks, PA: Strategic Studies Institute, U. S. Army War College, 2013).

[31] Sagan, *The Limits of Safety*; Diane Vaughan, *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA* (Chicago: University of Chicago Press, 1996); Charles Perrow, *Normal Accidents: Living with High Risk Technologies*, 2nd ed. (Princeton, NJ: Princeton University Press, 1999); Scott A. Snook, *Friendly Fire: The Accidental Shootdown of U.S. Black Hawks over Northern Iraq* (Princeton, NJ: Princeton University Press, 2000); Karl E. Weick and Kathleen M. Sutcliffe, *Managing the Unexpected: Sustained Performance in a Complex World*, 3rd ed. (Hoboken, New Jersey: John Wiley and Sons, 2015).

[32] Cimbala, *Nuclear Weapons in the Information Age*; Andrew Futter, "Hacking the Bomb: Nuclear Weapons in the Cyber Age" (ISA Annual Conference, New Orleans, 2015).

[33] Gartzke and Lindsay, "The Cyber Commitment Problem and the Destabilization of Nuclear Deterrence."

Table 6: Cyber-nuclear escalation mechanisms by escalating actor and timing

<table>
<tr><th>Actor</th><th>Peace</th><th>Crisis</th><th>War</th></tr>
<tr><td>Attacker</td><td>Bolt from the blue</td><td>Cyber commitment problem
Counterforce overconfidence</td><td>Cyber effectiveness window</td></tr>
<tr><td rowspan="2">Either</td><td colspan="3">Accidental launch</td></tr>
<tr><td>n/a</td><td>Fog of cyberwar</td><td>Damage limitation window
Cross-domain retaliation
Targeting error</td></tr>
<tr><td rowspan="2">Target</td><td colspan="2">Cybersecurity dilemma</td><td rowspan="2">Use it or lose it window
Conventional entanglement
Gambling for resurrection</td></tr>
<tr><td>Foiled prevention
Prior capitulation</td><td>Misinterpreted warning
Mistaken attribution
Foiled pre-emption</td></tr>
<tr><td rowspan="3">3rd Party</td><td colspan="3">Unauthorized launch</td></tr>
<tr><td colspan="2">False flag operation</td><td rowspan="2">n/a</td></tr>
<tr><td></td><td>Confusion from cyberspace</td></tr>
</table>

Table 6 and Table 7 list twenty-one different ways in which cyber operations can exacerbate the risks of nuclear escalation (with further discussion of individual mechanisms in the appendix below). An attempt has been made to identify distinct escalation pathways in terms of their strategic logic to illustrate that there is no one type of risk associated with cyber, although some are more dangerous than others. Specific pathways could be multiplied indefinitely by enumerating operationally different segments of the nuclear enterprise (Table 1) and particular vulnerabilities that these different mechanisms might exploit. No effort is made here to do so, nor would that be wise in this forum. No claim is implied that all of these mechanisms are distinct; indeed, some of them can and must interact (e.g., the fog of cyberwar features in many possible scenarios).

The mechanisms are arranged across columns temporally, because different risks can manifest during peacetime, during a nuclear brinksmanship crisis, and during war (where escalation constitutes a widening of conflict from the conventional to the nuclear level or to a higher rung on the nuclear escalation ladder). Some manifest across a range of temporal periods (e.g., accidental launch is always possible). The mechanisms are arranged across rows by the actor that makes the decision to escalate, which may or may not be the same actor that makes the decision to conduct cyber operations.

Table 7: Cyber-nuclear escalation mechanisms

Mechanism	Escalator?	Revealed?	Rational?	Plausible?
1. Accidental launch	Either	Yes	Error	Low
2. Bolt from the Blue	Attacker	No	Bounded	No
3. Cross-domain retaliation	Either	Yes	Bounded	Possible
4. Counterforce overconfidence	Attacker	No	Bounded	Possible
5. Confusion from cyberspace	3rd Party	Yes	Bounded	Possible
6. Conventional entanglement	Target	Yes	Bounded	Possible
7. Cyber commitment problem	Attacker	No	Yes	Yes
8. Cyber effectiveness window	Attacker	No	Maybe	Possible
9. Cybersecurity dilemma	Target	Yes	Bounded	Possible
10. Damage limitation window	Either	Maybe	Maybe	Yes
11. Gambling for resurrection	Target	Yes	Bounded	Possible
12. False flag operation	3rd Party	No	Bounded	Low
13. Fog of cyberwar	Either	Yes	Bounded	Yes
14. Foiled preemption	Target	Yes	Bounded	Possible
15. Foiled prevention	Target	Yes	Bounded	Yes
16. Misinterpreted warning	Target	Yes	Bounded	Possible
17. Mistaken attribution	Target	Yes	Bounded	No
18. Prior capitulation	Target	Yes	Bounded	Possible
19. Targeting error	Either	Yes	Error	Low
20. Unauthorized launch	3rd Party	No	Error	Possible
21. Use it or lose it window	Target	Yes	Bounded	Possible

The mechanisms may further differ by whether or not they are revealed to the target. Notably, most of them involve some awareness on the part of the target that an enemy cyber operation is occurring. A lack of target knowledge of a strike may be important because incentives to decide to escalate or negotiate, or to panic in the case of less than rational thinking, are based on information available about the potential outcomes of different courses of action and their associated costs. The target of an attack can become aware or strongly suspect that its NC3 is being degraded by cyber attack. Alternately, especially if the attacker is highly skilled, the effects of the attack may be invisible to decision makers. How decision makers react to the information revealed depends on assumptions about rationality.

Notably, only one of the mechanisms involve rational cost-benefit calculations: the cyber commitment problem which causes bargaining failure through rational incentives to misrepresent

strength. Two of these mechanisms might be rational under some additional assumptions: the damage limitation window which creates incentives to initiate counterforce strikes while command and control capabilities are still available, and the cyber effectiveness window which causes an attacker to lose the benefits of its cyber operations as wartime conditions alter target systems. The majority of the mechanisms, however, involve some sort of sub-rational thinking or error, which includes misperception, confusion, and bounded rationality. These scenarios involve some deviation from strictly rational utility maximization. Actors may be risk averse for gains and risk accepting for losses. They may think "hot" or emotionally under pressure. They may use heuristics and "thin slice" rather than gathering or using all the information available. Bounded rational thinking has been demonstrated in experiments and found to be relevant in historical case studies.[34] Finally, three of the mechanisms involve no rationality at all (accidental and unauthorized launch and targeting error); they involve some sort of malfunction that is unexpected from the perspective of the primary crisis actors.

Table 7 includes a qualitative judgment as to the plausibility of each scenario depending on the complexity of the cyber operations involved and assumptions needed to get to nuclear escalation. Four are assessed as plausible because there are easily conceivable scenarios in which cyber use raises the risks of escalation. Many are considered merely possible risks because they require numerous supporting assumptions to significantly raise escalation risk. Other scenarios seem to require more heroic assumptions. Moreover, not all of them carry the same risks of escalation, plausibility aside. Thus 'foiled prevention'—the discovery of an adversary's attempt to use cyber operations to degrade a latent nuclear capability—is very plausible but mainly serves to heighten mistrust and tension in future nuclear crises, should they transpire, and is thus only a very indirect escalatory risk factor.

The most dangerous scenario is the 'cyber commitment' problem. One can argue the plausibility of this scenario given the significant difficulty and sensitivity of mounting a successful disarming counterforce cyber operation against what should be considered to be hardened NC3 targets. However, the confluence of hidden operations (which produce information asymmetry) and strategic rationality (which makes the expected value of war greater than the expected value of settlement) make it a dangerous eventuality indeed, since even perfectly clear-thinking actors will have incentives to initiate a nuclear war for the sake of damage limitation if the bargaining range closes. In most other scenarios, sober appraisal of costs and benefits is more likely to reveal incentives to avoid escalation.

Conclusion

The interaction between cyber operations and nuclear weapons is a complex problem. The complexity and danger of interaction is only partly a function of technology but also a function of significant political, economic, and strategic differences associated with the use of cyber and nuclear capabilities. There are many potential vulnerabilities in the expanding attack surface of modern NC3, but the challenges associated with exploiting those vulnerabilities under actual operational and political conditions are nontrivial, to put it mildly. As a result, not every state

[34] J. M. Goldgeier and P. E. Tetlock, "Psychology and International Relations Theory," *Annual Review of Political Science* 4, no. 1 (2001): 67–92; Janice Gross Stein, "The Micro-Foundations of International Relations Theory: Psychology and Behavioral Economics," *International Organization* 71, no. S1 (April 2017): S249–63.

will have the same capability to conduct cyber operations against enemy NC3, and non-state actors are unlikely to have much capability at all.

Furthermore, not all nuclear rivalries will be destabilized to the same degree by cyber operations given asymmetric NC3 vulnerabilities, cyber capabilities, and expected consequences. In rivalries where cyber-NC3 interactions do manifest, furthermore, not all pathways to escalation are equally plausible or dangerous. The sheer variety of pathways in combination with the complexity of the systems involved and the political complexity of any crisis, however, should give one pause. The combination of cyber and nuclear capabilities in the twenty-first century may in effect constitute a doomsday device that might be triggered despite the wishes of the principal actors to avoid it.

It must be emphasized that all of these scenarios, like nuclear war generally, are all very unlikely. Nevertheless, given the high costs of nuclear war, every risk factor matters. There is some concern that these cyber-NC3 mechanisms do raise the marginal risk of nuclear war, thereby making a highly unlikely event slightly more likely.

APPENDIX

This appendix provides further description of each of the escalation mechanisms listed in Table 6 and Table 7. They are listed alphabetically.

1. Accidental launch

Technical glitches in NC3 system or cyber intrusion payload automatically trigger a launch. Human decision making is incidental to the error (i.e., downstream commanders may carry through on the order to launch). This results in a non-strategic escalation in the sense that neither the attacker nor the target deliberately decides to launch. This category encompasses a wide diversity of inadvertent breakdowns of the 'never' criterion.

2. Bolt from the blue

An attacker is so supremely confident in its ability to disable the target's command and control, and any retaliatory moves for that matter, that it launches a surprise disarming attack in peacetime as a sort of preventative war. This scenario strains credibility. Given the costs and risks of war, bolts from the blue don't happen.[35]

3. Cross-domain retaliation

Many imagined escalation scenarios involve the use of nuclear weapons for pre-emption or retaliation. Yet it is possible for an actor to opt for an attack in some other domain, such as a large-scale cyber attack on critical infrastructure with significant loss of life (electrical power, nuclear power plants, air-traffic control, etc.). An actor in a nuclear crisis absent of any cyber attacks on NC3 might attempt such a cyber attack in an attempt to de-escalate the crisis. Miscalculation could result in retaliation instead. This scenario makes strong assumptions about whether the target of a large-scale cyber attack would perceive it as so provocative that it would be willing to retaliate with nuclear weapons, but the escalatory potential of cyber operations is a matter of debate.[36] This scenario is more remote from the core concern of this paper with cyber-NC3 interactions and requires numerous supporting assumptions to produce nuclear escalation. Nonetheless, the expanded portfolio of options available to crisis actors can be expected to complicate the common knowledge requirements for strategic stability.[37]

4. Counterforce overconfidence

False optimism is a common cause of war.[38] This is the cyber-nuclear version. The attacker who has conducted a cyber attack on enemy NC3 is overconfident about its effectiveness. This has the equivalent effect of creating an information asymmetry that leads to bargaining failure per the cyber commitment problem, but this is based, perhaps in addition to it, on a misperception of the attacker's own performance. Thus the attacker, in the false belief that the target will be able to fully retaliate, feels that it can launch a disarming first strike with reduced consequences.

---

[35] Betts, Richard, *Surprise Attack: Lessons for Defense Planning* (Washington, D.C.: Brookings Institution Press, 1982).

[36] Jacquelyn Schneider, "The Information Revolution and International Stability: A Multi-Article Exploration of Computing, Cyber, and Incentives for Conflict" (Ph.D. Dissertation, George Washington University, 2017).

[37] James D. Morrow, "International Law and the Common Knowledge Requirements of Cross-Domain Deterrence," in *Cross-Domain Deterrence: Strategy in an Era of Complexity*, ed. Jon R. Lindsay and Erik Gartzke (New York: Oxford University Press, 2019).

[38] Geoffrey Blainey, *Causes of War, 3rd Ed.* (New York: Simon and Schuster, 1988).

5. Confusion from cyberspace

This is a subset of the fog of cyberwar problem but is broken out separately to highlight the importance of increased complexity in the external media and intelligence environment as a result of cyberspace. Cyber events external to the crisis dyad, ranging from technical operations that affect NC3 to worrisome events in the social media environment or intelligence reporting streams, create additional stress and confusion that heighten the risk of miscalculations.

6. Conventional-nuclear entanglement

Cyber attacks on networks that are used for both conventional C2 and NC3 are attacked in a conventional war. The target becomes worried that this is actually preparatory to nuclear war (misinterpreted warning), or degrading the usability of nuclear forces (use it or lose it), or degrading the opportunity to attack enemy nuclear forces if needed (damage limitation window).[39]

7. Cyber commitment problem

A cyber attack that disables NC3 cannot be revealed to the target for coercion before or during a crisis. This degrades the common knowledge of the outcome of the war and its costs, closing the bargaining range. Each actor is resolved, one with false optimism, the other unable to dissuade it of the illusion, and one or both decides to take its chances in war instead.[40]

8. Cyber effectiveness window

This is a close cousin to the damage limitation window but is a problem peculiar [do you mean particular?] to the attacker and potentially operable beyond damage limitation scenarios. This problem refers to the fact that the systems penetrated by cyber operations start to change during war as they are damaged and as the enemy introduces wartime reserve modes and makes other modifications. The result is that the attacker has a closing window of cyber effects during war, which could create pressures to act fast to attack targets before losing (or immediately after losing) whatever targeting or suppressive benefits are provided by the cyber operation.

9. Cybersecurity dilemma

This scenario works by either the attempted prevention or pre-emption mechanisms. The difference in this case is that the intrusion discovered by the target is not in fact intended by the intruder to do any damage for the sake of prevention or preemption. The discovered intrusion is actually just intended for reconnaissance. Unfortunately it uses the same close access intrusion and exploitation methods that a disruptive attack protocol would use. The intruder might actually have completely defensive intentions in probing NC3 to verify that the actor does indeed have a capable and credible deterrent, which would in most situations promote strategic stability. However, in this situation the intruder probes the NC3 system for intelligence but the defender cannot tell whether the intrusive malware also includes, or might download, some dangerous attack payload. After all, Stuxnet started off as a network reconnaissance operation before it started messing with valves and rotors at Natanz. The target knows that a successful cyber attack on its NC3 would disable its nuclear deterrent at comparatively low cost for the attacker, so it believes it lives in an offense dominant world. Yet the target cannot easily distinguish whether

---

[39] Goldstein, "First Things First"; Caitlin Talmadge, "Would China Go Nuclear? Assessing the Risk of Chinese Nuclear Escalation in a Conventional War with the United States," *International Security* 41, no. 4 (April 1, 2017): 50–92; Acton, "Escalation through Entanglement."

[40] Gartzke and Lindsay, "The Cyber Commitment Problem and the Destabilization of Nuclear Deterrence."

the intrusion is a defensive reconnaissance or a preparation for attack. This is the "doubly dangerous" world in Robert Jervis' classic analysis of the security dilemma and extended by Ben Buchanan for cybersecurity.[41] The target discovers an NC3 intrusion and assumes the worst. If the ambiguous discovery is made in a crisis it escalates per the attempted pre-emption scenario. If discovered in peacetime, future crises become more dangerous per the attempted prevention scenario.

10. Damage limitation window

Many of the scenarios here include a comment to the effect that additional assumptions are needed to move from heightened escalation risk to an actual decision to escalate. Others offer a few possibilities by considering escalatory pressures that result from strategic interaction. This is akin to the classic scenario described by Schelling where both sides rush to execute a disarming first strike. The window need not be mutual, however. In particular, the cyber attacker may believe that its operation will only be effective in impairing NC3 for a limited amount of time before wartime operations start invalidating operational assumptions. The target, on the other side, may assume that cyber attacks will begin to limit its ability to launch damage limitation strikes, even if the forces themselves are not at risk.[42]

11. Gambling for resurrection

Cyber attacks alert the target that it is probably going to lose a war. It decides to attack in an attempt to 'escalate to deescalate.' That is, the actor hopes that by using nuclear weapons in some limited way (an atmospheric burst for signaling, a demonstration in a peripheral area, or limited attacks on enemy forces), it will demonstrate its resolve to stay in the fight and convince its enemy that the war is not worth continuing. This requires some boundedly rational thinking or great risk acceptance given the premise of knowledge that the war is all but lost.

12. False flag operation

In this variant of the unauthorized launch scenario, a third party doesn't directly initiate a launch but rather takes actions through cyberspace or through NC3 that persuade one party of a nuclear dyad that the other is about to attack. Additional assumptions are needed to explain why the actor expects to benefit from starting a nuclear war, which may not be rational.

13. Fog of cyberwar

Some version of subrational or boundedly rational thinking is often invoked in crisis escalation scenarios. Time pressure, panic, and exhaustion can degrade the quality of decision making in crisis or war. In this scenario the confusion is heightened by cyber attacks that degrade the quality of incoming warning data and reports from operational units and outgoing requests for information and instructions to units. A variant of this problem might emerge not from cyber attacks on NC3 per se but on the intelligence and assessment functions, perhaps even including social media manipulation, that complicate assessment of the nature and stakes of the crisis. It can be an exacerbating factor in several of the other scenarios. It is treated as a separate scenario from misinterpreted warning because warnings of incoming missiles may indeed be accurate while other reports from friendly units or command decision making may be impaired.

---

[41] Ben Buchanan, *The Cybersecurity Dilemma: Hacking, Trust and Fear Between Nations* (New York: Oxford University Press, 2017).

[42] Acton, "Escalation through Entanglement."

The fog of cyberwar concept can also be distinguished from several other mechanisms by the defender's suspicion or awareness that cyber attacks are occurring. As systems begin malfunctioning, it becomes obvious that the enemy has launched a cyber attack. In other scenarios below the enemy attack is never revealed. In this case, however, the revelation of cyber operations in the midst of a crisis create panic and confusion that muddle decision making. Further assumptions are needed to explain why a greater sense of uncertainty and a lack of confidence in NC3 does not lead to de-escalation. From a purely rationalist perspective, the target's dawning awareness of cyber attacks could also plausibly reveal information about an unfavorable balance of power leading an actor to terminate a crisis rather than risk paying greater costs by continuing or escalating a war it is likely to lose.

14. Foiled preemption

In this scenario, a cyber attack attempting to disrupt or disable NC3 is discovered in the midst of a crisis. Unlike the fog of war scenario in which the target panics and starts a war, in this case the target is able to respond to the incident and defeat the attempted pre-emption. This episode reveals to the target that the cyber attacker had hostile intent. The target is more likely to believe that the attacker was willing to initiate or in the process of initiating a pre-emptive nuclear strike. The target decides that crisis bargaining has failed and war is inevitable, so it decides to launch with its NC3 fully intact. Further assumptions are needed to explain why the target decides that bargaining is no longer possible and deterrence is no longer possible now that it has restored confidence in its NC3 and its ability to inflict terrible costs on the enemy.

15. Foiled prevention

This scenario works by a similar logic, i.e., making subsequent crises more dangerous. The difference is that the discovery of the enemy cyber operation does not take place in the midst of a crisis. Instead the discovery occurs in peacetime and involves the use of cyber operations to prevent the acquisition of an operational nuclear deterrent. At the very least the discovery of sabotage provides confirmation for the target that the attacker fears its possession of a nuclear weapon, which further justifies its reasons for wanting one. Furthermore, the attacker's use of covert cyber means rather than something more definitive like an airstrike suggests that the attacker faces some constraints in what it can or is willing to do for counterproliferation. The compromised cyber operation becomes a signal of restraint that increases the confidence of the proliferator that it can make a dash for a bomb with reduced fear of preventative war.
This scenario is more dangerous if the counterproliferation operation involves cyber means that can also feasibly be used for counterforce preemption (i.e., the U.S. "left of launch" scenario vs. North Korean missile launches as distinguished from Stuxnet, which only targeted the Iranian fuel cycle). In this case, as with the previous scenario, the discovery of the cyber attack prompts the target to audit and improve its cyber defenses. Yet this scenario does not require the heroic assumption that the winner of a nuclear crisis allows the loser to retain an operational arsenal. The target improves its NC3 security and gains greater confidence in subsequent crises. A more resolved target becomes harder to coerce and heightens the possibility of nuclear risk taking and escalation by other pathways here. Additional assumptions about the subsequent crisis are needed to explain escalation.

16. Misinterpreted warning

In the classic false positive scenario, a false reading from a sensor or on a display screen leads an actor to assess that an enemy has decided to attack. Historical near-misses like Canadian geese or lunar reflections mistaken for missile indications by radar operators or a training tape of Soviet

missile attack loaded into live NORAD systems fall into this category. This is similar to the accidental launch scenario in that no one fully intends the resulting war but with the difference that misguided human decision making plays a key role in the decision to launch. This scenario is incomplete without a further account of why actors would feel pressured to pre-empt rather than wait for further confirming or disconfirming evidence. The use-or-lose scenario below is one possible explanation, should the false warning activate fears of losing the ability to retaliate or limit damage. Most false positive scenarios will usually assume fog of war below as well. It is notable that there have been hundreds if not thousands of false alarms in the history of NC3 without any decision to launch, even in the midst of crisis. Is restraint to date just a matter of good luck or a feature of high reliability organizations?

17. Mistaken attribution

In this scenario, cyber attacks have become visible as above (misinterpreted warning), but the actor responsible for the cyber attacks is misidentified. This raises the risk of retaliation against the wrong target, leading to further retaliation in a vicious spiral of miscalculation. This scenario is often evoked to speculate about how nonstate or weak state actors might run a "false flag" operation to trick nuclear rivals into attacking each other. As in the misinterpreted warning scenario, additional assumptions are required to explain the gullibility and impulsiveness of the target state(s). It also requires strong assumptions about the difficulty of attribution in cyberspace which may not be warranted, especially given the high stakes involved and limited pool of actors capable of performance-degrading NC3 intrusion.[43]

18. Prior capitulation

In the fog of war scenario, the revelation of cyber attacks is assumed to make escalation more likely. However, as mentioned, the revelation of compromised NC3 could also lead a rational actor to de-escalate the crisis given new information about its deteriorating bargaining position. In this scenario the heightened risks of strategic instability emerge across rather than within crises. The target, which has just capitulated, is assumed to retain its arsenal. It then proceeds to repair and harden its NC3 and cyber defenses. Then, during the next crisis, that actor's mistrust and resolve will be further increased. Furthermore, targets that select themselves into a subsequent crisis will have already taken into account the possibility that their systems might be hacked by the enemy. Indeed, the rival power has demonstrated a willingness to use cyber operations against the target's NC3 in the previous crisis that it won. Thus the target has already discounted its expected effectiveness due to possible cyber attacks and still decided that it was willing to run nuclear risks. Such a target is highly resolved to take extreme risks, which undermines strategic stability and exacerbates the risks in the other scenarios. Additional assumptions about the subsequent crisis are needed to explain escalation.

19. Targeting error

This scenario assumes that nuclear war has already been initiated but is inadvertently widened. For whatever reason, rational or not, a state has decided to launch a nuclear strike. However, because NC3 or the platform itself has been hacked, the weapon does not go where it is supposed to. Think Chinese Embassy bombing with nuclear weapons. Either the nuclear strike hits a third party or some more sensitive target (i.e., a bomb intended as a demonstration shot over an unpopulated area falls short and hits a city). This carries high risks of triggering a retaliatory

[43] Thomas Rid and Ben Buchanan, "Attributing Cyber Attacks," *Journal of Strategic Studies* 38, no. 1–2 (2015): 4–37; Lindsay, "Tipping the Scales."

spiral when the injured party believes they have been attacked on purpose. The actor that made the mistake thus gets far more escalation than it bargained for. Like accidental launch, this category describes a nonstrategic escalation because a random error is to blame. The unauthorized launch scenario below would include deliberate interference in targeting systems to widen the target. Additional assumptions are required to explain why the actor that made the mistake cannot communicate its error and make conciliatory gestures to the injured party.

20. Unauthorized launch

In this scenario, insider threats (e.g., a General Ripper) or external hackers with falsified credentials engineer and authorize a launch that the national command authority does not desire. This scenario can be considered nonstrategic because the key states (the attacker and the target) are not witting to impending war. Strategic considerations may be part of the subversive actor's theory of victory (again, the Ripper scenario), but they do not start the train in motion. The insider threat scenario could be the result of a breakdown in civil-military relations or a wildly successful human intelligence penetration. This category is an egregious breakdown of the 'never' criterion and is distinguished from the accidental launch scenario by deliberate subversive action with the intent to launch weapons. Further strong assumptions are needed to explain why the subversive agent would want to deliberately start a nuclear war the political principal prefers to avoid, to say nothing about how the agent is able to subvert counterintelligence efforts and operational safeguards against this very scenario.

21. Use it or lose it window

Cyber attacks do not remain hidden, but rather their effects alert the target to the degradation of its nuclear forces. The target worries that ongoing or future cyber attacks are inevitable and it will lose the ability to command and control its forces, or the forces themselves, or else it begins to panic or become engaged. The target rushes to launch its weapons while it still can, motivated either to limit damage or inflict punishment. This is not strictly a rational move as the revelation of an increasingly unfavorable military situation together with the expectation of further retaliation following escalation should make negotiation more attractive. Combined with other mechanisms, however, such as gambling for resurrection, conventional entanglement, or damage limitation windows, bounded rationality might make nuclear use attractive.

III. ENDNOTES

IV. TECHNOLOGY FOR GLOBAL SECURITY INVITES YOUR RESPONSE

Technology for Global Security invites your responses to this report. Please send responses to: info@tech4gs.org. Responses will be considered for redistribution to the network only if they include the author's name, affiliation, and explicit consent