Royal United Services Institute
for Defence and Security Studies

Occasional Paper

Creating Restraint in Cyberspace

Forward Cyber Operations and Theories of Restraint

Sean Atkins

Creating Restraint in Cyberspace

Forward Cyber Operations and Theories of Restraint

Sean Atkins

RUSI Occasional Paper, April 2021

Royal United Services Institute
for Defence and Security Studies

190 years of independent thinking on defence and security

The Royal United Services Institute (RUSI) is the world's oldest and the UK's leading defence and security think tank. Its mission is to inform, influence and enhance public debate on a safer and more stable world. RUSI is a research-led institute, producing independent, practical and innovative analysis to address today's complex challenges.

Since its foundation in 1831, RUSI has relied on its members to support its activities. Together with revenue from research, publications and conferences, RUSI has sustained its political independence for 190 years.

Published in 2021 by the Royal United Services Institute for Defence and Security Studies.

RUSI Occasional Paper, April 2021. ISSN 2397-0286 (Online).

Royal United Services Institute
for Defence and Security Studies
Whitehall
London SW1A 2ET
United Kingdom
+44 (0)20 7747 2600
www.rusi.org
RUSI is a registered charity (No. 210639)

Contents

Executive Summary

AS OFFENSIVE CYBER activities continue to trend in a dangerous direction, some states are turning to defensive strategies that involve cyber operations beyond the boundaries of their own systems. A commonly cited aim of this approach is to induce restraint in adversaries, but how to accomplish this goal remains unclear. To better understand the utility and risks in applying forward cyber operations to generate restraint in competitors, this paper examines how they might be applied under foundational theories of restraint: deterrence by punishment and compellence; deterrence by denial; entanglement; normative constraints and tacit cooperation. A structured analysis reveals three key implications and suggests associated recommendations for future policy development:

- Using forward cyber operations to induce restraint is not a straightforward matter of imposing costs. Instead, both costs and gains can be affected at multiple points of an adversary's calculus. The utility, requirements, and risks involved in doing so vary significantly across different theoretical pathways to restraint. Current cyber strategy development should expand scope to consider multiple paths to restraint, accounting for their distinct utility and risks, to offer policymakers greater flexibility in addressing a broader range of cyber threats (above and below the threshold of armed conflict). In general, the utility–risk analysis here suggests that forward cyber operations should prioritise: intelligence collection for deterrence by denial and as an enabler of other restraint pathways; and targeting adversary cyber-operations infrastructure.
- A state's ability to leverage partnerships (both internationally and domestically between government and industry) influences its ability to affect cyber adversaries' costs-gains calculus with forward cyber operations. Accordingly, governments must develop trusted, more operational cyber partnerships with key allies and select private firms that have the capability to exchange and leverage forward-derived information in implementing the various paths to restraint.
- Effective forward defence involves more than intelligence agencies and the military. Other government organisations possess distinct expertise, relationships and capabilities that can produce powerful forward effects (such as law enforcement agencies or economic organisations that can request or compel owners of forward infrastructure to reduce malicious actors in their networks). Meaningfully integrating a broader range of capable actors into strategies that use forward cyber operations for defence is essential. This, however, can require greater centralised control than some states currently exhibit to overcome the organisational and political impediments to improving collaboration and unity of effort. Possessing capability alone is not the full measure of a country's power. The ability to organise and employ it to effect is just as important.

Introduction

THE GROWING RISKS posed by cyber threats are driving several states to change their approach to cyber defence. Election interference, power grid disruption and malware designed to kill are the latest reminders that offensive cyber operations continue to trend in a dangerous direction.[1] As the recent US Cyber Solarium Commission noted, 'adversaries have moved beyond simple denial-of-service and website defacement campaigns to conducting intelligence collection, ransomware attacks, and destructive operations as well as disruptive attacks on critical infrastructure'.[2] In response, some governments are implementing policies of forward defence.[3] That is, surveilling and engaging cyber adversaries beyond the bounds of defended networks to respond to or undermine malicious operations. In the UK, for instance, there have been calls for the proportionate use of cyber capabilities in response to cyber attacks that breach international law.[4] In the US, the commander of Cyber Command argues that defending critical national interests requires cyber forces to 'operate against our enemies on *their* virtual territory'.[5] In its recent cybersecurity strategy, the Australian government described a policy of actively defending networks 'using both defensive and offensive tools'.[6] Some states in the Middle East are also integrating offensive cyber capabilities into defensive strategies.[7]

While officials have articulated a number of objectives for using forward cyber operations for defensive purposes, one commonly cited aim is to induce restraint in cyber adversaries. Precisely how these operations might influence restraint, what affects their utility and what risks are involved remain largely underexamined. Policy documents and speeches alternately refer to deterrence, norm enforcement or generally to imposing costs on cyber offenders to

1. James Ball, 'Russian Interference Threatens Elections Across the World – Including Ours', *Bureau of Investigative Journalism*, 13 November 2019; Andy Greenberg, 'How an Entire Nation Became Russia's Test Lab for Cyberwar', *Wired*, 20 June 2017; Martin Giles, 'Triton Is the World's Most Murderous Malware and It's Spreading', *MIT Technology Review*, 5 March 2019.
2. Cyberspace Solarium Commission, 'Cyberspace Solarium Commission Report', March 2020.
3. This paper uses the terms 'forward defence' and 'forward cyber operations' to describe the general use of external cyber operations for defensive purposes by a number of states. Importantly, this not synonomous with the US 'Defend Forward' policy, which is just one manifestation of forward defence.
4. Jeremy Hunt, speech at the National Cyber Security Centre, 23 May 2019, <https://www.gov.uk/government/speeches/foreign-secretary-speech-at-the-nato-cyber-pledge-conference>, accessed 15 March 2021.
5. Paul M Nakasone, 'A Cyber Force for Persistent Operations', *Joint Force Quarterly* (Vol. 92, 2019), pp. 10–14. Author's emphasis.
6. Australian Government, 'Australia's Cyber Security Strategy 2020', <https://www.homeaffairs.gov.au/cyber-security-subsite/files/cyber-security-strategy-2020.pdf>, accessed 25 October 2020.
7. Walid Tohme et al., 'Cyber Security in the Middle East: A Strategic Approach to Protecting National Digital Assets and Infrastructure', Strategy&, 2015.

shape their behaviour.[8] Policymakers appear to envision multiple pathways to restraint, but each holds distinct requirements and implications for how forward cyber operations might be conducted. Understanding how these operations fit within the logics of these paths to restraint is critical to achieving desired ends while reducing the risk of destabilisation and exacerbation of cyber conflict.[9]

This is not to say that governments view restraint as the sole or even primary objective of forward defence. For example, the US's 'Defend Forward' cyber strategy is closely coupled with the concept of 'persistent engagement', which primarily aims to continuously contest bad cyber behaviour to reduce the scope and influence of adversaries' cyber abilities.[10] Indeed, this concept was developed as a better-fitting alternative to traditional deterrence approaches because of their failure to achieve restraint in cyberspace below the threshold of armed conflict.[11] Despite this, policymakers still point toward the need for creating restraint in cyberspace and often see a role for forward cyber operations in achieving it. Furthermore, some of the strategic concepts that do not point primarily toward restraint have notions of restraint embedded within them. For instance, the creators of persistent engagement suggest forward cyber operations might act as tacit bargaining tools that develop restraint around commonly understood bounds of behaviour.[12] In short, while other aims exist, there is continuing policy interest in creating restraint and in the potential for forward cyber operations to influence it.

Given a continuing flow of alarming headlines, some might scoff at the idea of restraint in cyberspace.[13] Although, as the US Cyber Strategy highlights, states appear to be 'recklessly … engaging in … malicious cyber activities', restraint also seems to exist in cyberspace: the

8. See, for example, US Department of Defense, 'Summary: Department of Defense Cyber Strategy', 2018; Cyberspace Solarium Commission, 'Cyberspace Solarium Commission Report', p. 24; Paul M Nakasone and Michael Sulmeyer, 'How to Compete in Cyberspace: Cyber Command's New Approach', *Foreign Affairs*, 25 August 2020; Jeremy Hunt, speech given at Glasgow University, 7 March 2019, <http://www.ukpol.co.uk/jeremy-hunt-2019-speech-on-cybersecurity/>, accessed 15 March 2019; Australian Government, 'Australia's Cyber Security Strategy 2020', p. 20.
9. Jason Healey, 'The Implications of Persistent (and Permanent) Engagement in Cyberspace', *Journal of Cybersecurity* (Vol. 5, No. 1, 2019).
10. Michael P Fischerkeller and Richard J Harknett, 'Persistent Engagement and Cost Imposition: Distinguishing Between Cause and Effect', *Lawfare*, 6 February 2020.
11. It is important to note here that traditional deterrence by high-cost punishment is only one theoretical pathway to restraint and its failure is found in one part of the cyber competition spectrum (below the threshold of armed conflict).
12. Michael P Fischerkeller and Richard J Harknett, 'Persistent Engagement and Tacit Bargaining: A Path Toward Constructing Norms in Cyberspace', *Lawfare*, 9 November 2018; Michael Fischerkeller, 'Persistent Engagement and Tacit Bargaining: A Strategic Framework for Norms Development in Cyberspace's Agreed Competition', Institute for Defense Analysis, November 2018.
13. For recent examples, see *Reuters*, 'Solarwinds Hack Was "Largest and Most Sophisticated Attack" Ever: Microsoft President', 15 February 2021; Frances Robles and Nicole Perlroth, '"Dangerous Stuff": Hackers Tried to Poison Water Supply of Florida Town', *New York Times*, 8 February 2021.

lights are still on in New York City, Russian gas pipelines still operate and aircrafts still cycle unimpeded through Beijing Capital International Airport.[14] While restraint appears to exist, its scope and why its strength varies are still understudied, forcing those who research international competition in cyberspace back to foundational theories to understand state behaviour in a relatively new context.

In order to better understand the utility and risks in using forward cyber operations to generate restraint in competitors, this paper examines them within the context of foundational theories of restraint: deterrence by punishment and compellence; deterrence by denial; entanglement; normative constraints and tacit cooperation. While the current cyber literature includes an increasingly rich debate on deterrence in cyberspace, application to forward cyber operations for defence is still thin and other theories of restraint remain relatively underexamined.[15] Often, they are simply considered as different flavours of deterrence. However, although each of the theoretical pathways share the logic of affecting competitors' costs and benefits from cyber operations, each is distinct in the types of costs and benefits implicated and how they are affected. Furthermore, concepts beyond the ubiquitously considered deterrence by punishment might offer ways to achieve restraint on a broader range of cyber threats, including those below the level of an armed attack ('grey zone' threats).

For this paper, forward cyber operations are considered to involve activities conducted in cyberspace beyond the boundaries of the networks owned and operated within a state. When used for defensive purposes, they can be leveraged for 'the proactive observing, pursuing, and countering of adversary operations'.[16] This excludes activities conducted within home networks (often described as 'blue' cyberspace), including both traditional passive defensive operations and more active approaches (for example, internal threat hunting). Rather, it involves actions in adversary-owned or -controlled infrastructure (described as 'red' cyberspace) as well as in systems that are owned and operated in third-party states ('grey' cyberspace).[17] These operations can be for gathering information on adversary activities and for producing cyber effects, whether virtual (as with data manipulation or changing system operating characteristics) or potentially physical (as with disruption of electricity distribution through manipulation of industrial control systems). They can also be used to engage adversary cyber forces, to include targeting their operational infrastructure used to command and control, manoeuvre and stage offensive cyber operations.

The analysis that follows frames restraint as the result of a cost–benefit calculus by competitors considering malicious activity. Inducing restraint is therefore a matter of manipulating either or both sides of this calculus to make the costs of cyber offence outweigh its potential gains. Each

14. White House, 'The Cyber Strategy of the United States of America', 2018.
15. See, for example, Joseph S Nye Jr, 'Deterrence and Dissuasion in Cyberspace', *International Security* (Vol. 41, No. 3, Winter 2016/17), pp. 44–71; Erica D Borghard and Shawn W Lonergan, 'The Logic of Coercion in Cyberspace', *Security Studies* (Vol. 26, No. 3, 2017), pp. 452–81.
16. Cyberspace Solarium Commission, 'Cyberspace Solarium Commission Report'.
17. US Joint Chiefs of Staff, 'Cyberspace Operations', Joint Publication 3-12, 8 June 2018.

theory considered here affects this calculus in different ways. Deterrence by punishment and compellence exact pain costs that prevent potential or stop ongoing transgression. Deterrence by denial raises the costs of offending operations and reduces the gains they might offer by increasing defence and resilience capabilities. Entanglement involves costs to offenders by virtue of being a part of interdependent systems while norms involve costs from breaking collectively established bounds of behaviour. Restraint developed through tacit cooperation can involve pain and denial costs that build up over iterated interactions between competitors to outweigh gains.

This paper applies a structured analysis approach, first by individually considering each theory of restraint and examining how forward cyber operations might fit within their logic. Next, it identifies the implications for their use, particularly those related to its utility, requirements and the potential risks involved. The risk analysis focuses on the potential for escalation (in and out of cyberspace) and of producing unwanted collateral effects. The paper concludes with a summary of findings and key recommendations for future policy development.

I. Deterrence by Punishment and Compellence

DETERRENCE BY PUNISHMENT and compellence are built on the logic that the probability of enemy attacks can be reduced by convincing an adversary that attacking will result in a net loss.[18] As Thomas Schelling famously argued, 'the power to hurt is most successful when held in reserve … it is the threat of damage, or of more damage to come, that can make someone yield'.[19] For deterrence by punishment, this damage is threatened in advance of a possible attack to dissuade an adversary. Under compellence, it is threatened to convince an adversary that an ongoing attack is not worth continuing. Importantly, in its cost–benefit calculus, an adversary not only considers the direct cost of the threatened response. Crossing or flirting with the punishment trigger line also involves embarking on a path of competitive interaction filled with uncertainty regarding exactly where it leads or what further damage may be suffered along the way.[20] Because of the mutual inability to predict or fully control the consequences of these interactions, this adds risk of further costs to an adversary's calculus.[21] The nuclear weapons context, where much of this theory was developed, offers an illustrative example of these concepts. Inducing nuclear restraint takes the familiar form of threatening assured destruction through a secure second-strike capability. The influence of future risk on the cost–benefit calculus can be seen throughout the Cold War, where the uncertainty around the chance of escalation to a nuclear conflict affected restraint at lower levels of conflict.[22]

In analysing how use of forward cyber operations might figure into this logic, strategists must examine how it relates to four fundamental conditions of deterrence by punishment and compellence: communication of a threat; credibility of the threat; meaningful potential punishment cost; and reassurance that no costs will be imposed while restraint holds.[23]

---

18. Thomas C Schelling, *Arms and Influence* (New Haven, CT: Yale University Press, 1966); Glenn H Snyder, *Deterrence and Defense: Toward a Theory of National Security* (Princeton, NJ: Princeton University Press, 1961).
19. Schelling, *Arms and Influence*, p. 2.
20. *Ibid.*, p. 97.
21. *Ibid.*, pp. 92–96, 109.
22. See Robert Jervis, *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Nuclear Armageddon* (Ithaca, NY: Cornell University Press, 1990); Sarah Kreps and Jacquelyn Schneider, 'Escalation Firebreaks in the Cyber, Conventional, and Nuclear Domains: Moving Beyond Effects-Based Logics', *Journal of Cybersecurity* (Vol. 5, No. 1, 2019).
23. For a broader analysis of coercion in cyberspace, see Borghard and Lonergan, 'The Logic of Coercion in Cyberspace'.

First, a clearly communicated threat is required. This is a particularly challenging task in cyberspace where secrecy is a necessity for operations that rely on exploitation of vulnerabilities that can often be easily overcome with simple system updates or software patches. Despite this, in some cases deterrers may attempt to communicate threats through intentionally observable cyber actions. For example, in 2016 Russian cyber operators compromised US power grid control rooms but held back from disrupting electricity distribution.[24] This might be interpreted as a signal that Russia intends to hold critical infrastructure at risk for deterrent purposes. Reception becomes problematic here, however, as the source, intent and meaning of the activity are left for the recipient to discern, opening significant space for misperception.[25]

One area where forward cyber operations may possess an advantage in communication involves what Schelling described as 'connectedness'.[26] He argued that coercive threats usually benefit from connection between the 'proscribed action and the threatened response'. Non-cyber threats lack this kind of inherent connection to cyber provocations, relying again on the coercer's perception and ability to put the disparate pieces together themselves. For example, in 2018 the US threatened economic consequences to induce Chinese cyber restraint in its longstanding commercial espionage campaign.[27] To be effective, these threats required explicit and unambiguous messaging that connects the potential punishment to the offending activity. This, however, is often a more complex endeavour than might be expected. In this US–China case, the same economic punishment tool was also applied to over 100 other demands associated with their ongoing trade war, resulting in muddled messaging.[28] Cyber-based deterrent threats, on the other hand, could offer clearer communication through their 'connectedness' to the activity to be deterred.

---

24. Steven Musil, 'Russian Hackers Accessed US Electric Utilities' Control Rooms', *CNET*, 24 July 2018, <https://www.cnet.com/news/russian-hackers-reportedly-gained-access-to-us-utility-control-rooms/>, accessed 15 March 2021. For a deeper look at the Russian cyber operations discussed here, see Andy Greenberg, *Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers* (New York, NY: Anchor, 2020).
25. For a more detailed discussion of signalling and its limits in cyberspace, see Ben Buchanan, *The Hacker and the State: Cyber Attacks and the New Normal of Geopolitics* (Cambridge, MA: Harvard University Press, 2020). For rich discussions on signalling in cyberspace specifically in relation to deterrence, see Jon R Lindsay, 'Tipping the Scales: The Attribution Problem and the Feasibility of Deterrence Against Cyberattack', *Journal of Cybersecurity* (Vol. 1, No. 1, September 2015), pp. 53–67; Jacquelyn G Schneider, 'Deterrence In and Through Cyberspace', in Jon R Lindsay and Erik Gartzke (eds), *Cross-Domain Deterrence: Strategy in an Era of Complexity* (Oxford: Oxford University Press, 2019), pp. 95–120.
26. Schelling, *Arms and Influence*, p. 88.
27. For additional details on China's industrial cyber campaign, see Lorand Laskai and Adam Segal, 'A New Old Threat: Countering the Return of Chinese Industrial Cyber Espionage', Council on Foreign Relations, 6 December 2018; Cissy Zhou, 'China to Face "Pain" in Meeting US Trade War Demand on Cybertheft, But Will Beijing Yield?', *South China Morning Post*, 26 February 2019.
28. Zhou, 'China to Face "Pain" in Meeting US Trade War Demand on Cybertheft, But Will Beijing Yield?'.

Second, deterrence by punishment also requires that deterrers' threats are credible. Achieving credibility involves not only possessing the *capability* to carry out the threat, but also demonstrated or perceived *resolve* to do so. Credible capability relies on two constituent factors: instrument recognition and effect recognition.[29] To continue the Russian power grid operations example, in gaining access to US power grid controls, Russia revealed the *instrument*. In its previous cyber attacks that disrupted electricity transmission in Ukraine, Russia demonstrated the *effect*.[30] Here again, however, demonstrating capability also risks losing it. Exposing details of the instrument and effect to achieve credibility can increase the likelihood of losing the access or ability to affect targeted systems.[31]

The credibility of a threat is also a matter of the deterrer's resolve, which is linked to its interests at stake. If a state's interests are large enough to justify the risks associated with acting on the threat, then the resolve is credible.[32] Because of their generally lower associated risks, forward cyber operations may have an advantage over traditional military means when it comes to credible resolve. A target of deterrence may believe a cyber threat before a threat of an airstrike because the risks involved to the deterrer are more in line with the interests at stake. Of course, this increase in credibility of resolve comes at the expense of degree of deterrent cost imposed, which influences the power of the threat. This is addressed below.

Third, threatened operations must meaningfully influence competitors' cost–benefit calculus. Unfortunately, creating cyber effects that impose outright deterrence-level costs requires technical sophistication and access to challenging targets, both of which can be difficult to attain and maintain. As a result, governments may be unable to threaten deterrence-level costs in cyberspace and are even less likely to make that deterrent enduring. For instance, what level of cyber-imposed cost would be required to deter future Russian cyber-based election interference operations? The demonstrated value in amplifying division in its targets and reduced trust in their institutions likely outweighs the potential costs of similar operations that might be conducted against Russia in response. Furthermore, this cost calculation may not be as straightforward as deterring states may hope. For instance, how does holding Russia's power grid at risk to deter similar attacks on US infrastructure figure into the Russian cost–benefit calculus?[33] While mirror cyber operations may send clear signals, does Russian leadership value uninterrupted electricity in the same way as US leadership? Similar targets of deterrence operations may not hold the same value across different states. As David Edelman argues, 'restraining cyberattacks requires

29. R David Edelman, 'Cyberattacks in International Relations', PhD dissertation, University of Oxford, 2013.
30. Joe Slowik, 'CRASHOVERRIDE: Reassessing the 2016 Ukraine Electric Power Event as a Protection-Focused Attack', Dragos, 15 August 2019, <https://www.dragos.com/wp-content/uploads/CRASHOVERRIDE.pdf>, accessed 12 April 2021.
31. For further analysis of how the need for deception in cyber operations affects deterrence in cyberspace, see Lindsay, 'Tipping the Scales'.
32. Daryl G Press, *Calculating Credibility: How Leaders Assess Military Threats* (Ithaca, NY: Cornell University Press, 2007), p. 20.
33. David E Sanger and Nicole Perlroth, 'U.S. Escalates Online Attacks on Russia's Power Grid', *New York Times*, 15 June 2019.

more than simply amassing greater, similar capabilities in the manner of most conventional and nuclear deterrence'.[34]

Beyond these challenges, forward cyber operations may be able to threaten potential costs both as a broader manipulator of risk and as a facilitator of other deterrent action. Even if threatened cyber-imposed costs are not capable of outright deterrence, these operations are a step down Schelling's uncertain 'path of adversarial interaction' that may bring higher costs later.[35] As a hypothetical example, a state may threaten a twofold cyber response to any cyber attack on its critical infrastructure. While the cyber-imposed costs themselves may not outweigh the benefit provided to the transgressor, the doubling of costs increases risk by putting the states' interactions on an escalatory trajectory of uncertain future costs. In this way, forward cyber operations may be able to serve as a tool of risk manipulation that raises the uncertainty and potential costs for cyber offenders.

Separately, forward operations might facilitate or combine with cross-domain deterrent capabilities that offer greater flexibility and range of costs imposed. Eric Gartzke and Jon Lindsay recognised that classical deterrence did not fit many of today's security challenges, arguing that it may instead rely on the use of 'capabilities of one type to counter threats or combinations of threats of another type'.[36] For instance, threatening a traditional kinetic military response for catastrophic cyber attacks that result in loss of life and large-scale material loss offers a deterrent-level cost that cyber-based options are unlikely to achieve. Further down the conflict spectrum where the vast majority of cyber competition takes place, threatening economic or diplomatic costs may provide flexibility when cyber-imposed costs are still of insufficient weight or less responsive. In both cases, forward cyber operations can contribute to cross-domain deterrence by collecting intelligence on competitor cyber operations. Information that attributes attacks and exposes their details can support timely and well-targeted sanctions in response. They can also help reduce the uncertainty about the source of attacks that may make deterrers think twice about retaliating with higher-cost punishment, thus increasing the credibility of deterrent threats.

Finally, effective deterrence by punishment requires that deterrers reassure their targets that costs will not be imposed while they refrain from the proscribed activity. If an adversary believes that the threat might be acted on whether or not it crosses the deterrent threshold, then it has little incentive not to pursue whatever benefits it can gain by doing so anyway.[37] However,

34. Edelman, 'Cyberattacks in International Relations'.
35. Schelling, *Arms and Influence*, pp. 97–98, 105.
36. Erik Gartzke and Jon Lindsay, 'Cross-Domain Deterrence: Strategy in an Era of Complexity', Office of Naval Research, July 2014, <https://quote.ucsd.edu/deterrence/files/2014/12/EGLindsay_CDDOverview_20140715.pdf>, accessed 15 March 2021.
37. Jeffrey W Knopf, 'Varieties of Assurance', *Journal of Strategic Studies* (Vol. 35, No. 3, 2012), pp. 375–99. For instance, during the Second World War both the UK and Germany responded with bolstered aggression, not acquiescence, to the strategic bombing of their homeland. See Robert A Pape, *Bombing to Win: Air Power and Coercion in War* (Ithaca, NY: Cornell University Press, 1996).

providing this reassurance in cyberspace can be a complicated endeavour. Many cost-imposing cyber operations first require gaining access to and manoeuvring within adversary systems. Determining whether these activities are intended to establish a deterrent capability and will stop short of implementing disruptive effects is difficult to discern for target states analysing the ones and zeros.[38] This might be addressed by communicating intent via other channels but, again, this risks exposing operations that rely on secrecy to be effective.

Use of cyber operations for deterrence or compellence can also bring a risk of escalation and collateral effects. Risk of escalation (either within or outside cyberspace) from forward cyber operations is largely tied to the type of operation and system being targeted. The requirement for significant pain imposition via cyber means under deterrence by punishment or compellence narrows target options down to those that can create higher-level costs for an adversary, and these can involve greater potential risk for escalation. For example, compromising a state's critical infrastructure, such as water or power systems, where costs imposed might be high enough to deter or compel, will likely incur a very different response than disrupting the systems it uses to conduct malicious cyber activity. In conducting operations against costly targets, escalation might occur along two distinct pathways. First, an adversary might respond outright in an escalatory manner. Coercion has proven difficult throughout history and can sometimes result in increased offence rather than restraint.[39]

A second pathway, particularly under deterrence by punishment, involves inadvertent escalation due to misperception of the deterrer's intent. By necessity, cyber operations are conducted in secret, leaving little more than ones and zeros to intuit its intent. For instance, if an adversary is detected inside a nuclear power company's network, it might be for the purpose of: stealing power plant details to advance indigenous nuclear technology on the cheap; conducting reconnaissance to enable future destructive operations and hold it at risk; or conducting destructive operations at that moment. Operations under all three intents require the intrusion and the same activities up until moments before an actual attack takes place. Misinterpreting intent can be costly. As Ben Buchanan highlights, this risk of misperception can cause 'hostilities in what would otherwise have been a peaceful situation'.[40]

The risk of unintended collateral effects produced by forward operations is also a significant concern here. Cyber operations, like their kinetic counterparts, involve a degree of uncertainty about what will actually take place in execution. For instance, coding mistakes or lack of understanding of the targeted systems can lead to malware spreading or creating effects beyond these systems. Targeting the critical systems required for deterrence or compellence opens the door to high-cost mistakes that generate unwanted backlash as well as send incorrect signals to competitors. These systems can often overlap with or support the functioning of other systems, reducing predictability and raising the potential for unintended cascading effects. For example,

38. Ben Buchanan, *The Cybersecurity Dilemma: Hacking, Trust and Fear Between Nations* (Oxford: Oxford University Press, 2016), p. 20.
39. Pape, *Bombing to Win*.
40. Buchanan, *The Cybersecurity Dilemma*.

financial and communications systems are deeply dependent on electricity distribution. Because many of these systems are shared across borders, they might also affect neighbouring states' critical systems. The result is higher uncertainty as to the full extent of the effects deterrent or compellent cyber operations might create.

Uncertainty in attribution adds another dynamic to the risk of unintended effects. Consistently clear and timely attribution of attacks in cyberspace can be difficult.[41] As Bruce Schneier highlights, 'packets don't come with return addresses'.[42] Even with recent advances in attribution capabilities, achieving it with enough certainty to justify a high-cost attack in response and at a speed that facilitates a timely deterrent effect is still challenging. The reality is that, unless there is a claim of responsibility by the offender, there is likely to exist some level of uncertainty about the true culprit and the degree of this uncertainty determines the associated risk of imposing the threatened cost on the wrong state.

Where using cyber operations as a deterrent or compellent tool might be risky, leveraging them to support cross-domain activities can be less so. As described above, cyber operations can contribute to cross-domain cost imposition (whether kinetic, diplomatic or economic) by collecting intelligence on competitor cyber operations. The targets of these types of activities are likely to be the systems that adversary cyber operators use to conduct their operations. The potential for outright escalation, escalation due to misinterpretation, or unintended collateral effects is largely reduced with these targets. Because an adversary's cyber operations infrastructure is less useful for imposing high costs, if the intrusion is identified it is more likely to be correctly interpreted as spying (or perhaps an attempt at operational disruption) than a costly attack requiring an escalatory response. The target itself can imply intent, reducing the uncertainty that increases risk of escalation. Though reduced, this is not to say these operations are without risk. Still present, as with all operations, is the possibility of mistakes and misinterpretation that can have increased escalatory or collateral risk if the targeted cyber operations infrastructure happens to overlap with other important systems.

41. This is not to say that attribution is always difficult, but the difficulty varies based on a number of political, technical and capability factors. Additionally, recognition that an attack has taken place can take months to years and achieving attribution with decision-level confidence is often a time-consuming process. For discussions of the attribution challenge, see Thomas Rid and Ben Buchanan, 'Attributing Cyber Attacks', *Journal of Strategic Studies* (Vol. 38, No. 1–2, 2015), pp. 4–37; Lindsay, 'Tipping the Scales'.
42. Bruce Schneier, 'Attack Attribution in Cyberspace', Schneier on Security blog, 8 January 2015, <https://www.schneier.com/blog/archives/2015/01/attack_attribut.html>, accessed 15 March 2021.

II. Deterrence by Denial

RESTRAINT CAN BE induced not just by threatening harm, but also by raising the costs of conducting transgressive operations and reducing the gains adversaries might achieve through them.[43] Manipulating the operational costs for adversaries requires building defensive capabilities that increase the investment needed to achieve desired effects. To alter the gains side of adversaries' cost–benefit calculus means developing resilience to deny the value of attacks even if they overcome those defensive capabilities. As Joseph Nye describes, good cyber defences can 'reduce the incentive for some attacks by making them look futile'.[44] Forward cyber operations have significant potential to support a deterrence-by-denial strategy through provision of intelligence to bolster defence and resilience efforts as well as by disrupting threats closer to their source.

Cyberspace itself is composed of a vast web of interconnected networks, not all of which maintain high defensive standards. Without knowing where and how attackers will concentrate attention, cyber defenders are left scrambling to secure all possible vulnerabilities. Attackers need only identify the weak points within a system and focus resource investment there. As a result, the costs of effective defence in cyberspace often outweigh the costs of attack. Perhaps more importantly, however, it means that the gains made through attacks more easily exceed the investment made in conducting them. While many policymakers and scholars have noted the apparent offence–defence imbalance in cyberspace, it is also clear that the balance is not a static feature of cyberspace.[45] In fact, some researchers have recognised that it varies and can be manipulated to shift advantage to the defence, opening the feasibility of deterrence-by-denial strategies.[46]

---

43. Glenn H Snyder, 'Deterrence and Power', *Journal of Conflict Resolution* (Vol. 4, No. 2, 1960), pp. 163–78.
44. Nye, 'Deterrence and Dissuasion in Cyberspace'.
45. William J Lynn III, 'Defending a New Domain: The Pentagon's Cyberstrategy', *Foreign Affairs* (Vol. 89, No. 5, September/October 2010), p. 98; Lucas Kello, 'The Meaning of the Cyber Revolution: Perils to Theory and Statecraft', *International Security* (Vol. 38, No. 2, 2013), pp. 7–40; John Arquilla, 'Cyberwar Is Already Upon Us,' *Foreign Policy*, 27 February 2012; Martin C Libicki, *Cyberdeterrence and Cyberwar* (Santa Monica, CA: RAND, 2009); Joseph S Nye Jr, 'Cyber Power', Harvard Kennedy School, May 2010; Keir Lieber, 'The Offense–Defense Balance and Cyber Warfare', in Emily Goldman and John Arquilla (eds), *Cyber Analogies* (Monterey, CA: Naval Postgraduate School, 2015), pp. 96–107.
46. Rebecca Slayton, 'What Is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment', *International Security* (Vol. 41, No. 3, Winter 2016/17), pp. 72–109; Erik Gartzke and Jon R Lindsay, 'Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace', *Security Studies* (Vol. 24, No. 2, 2015), pp. 316–48.

Forward cyber operations can contribute to redressing the defensive burden in two ways. First, intelligence collection where threat actors operate externally can identify adversary tools, patterns of behaviour and infrastructure components used to manoeuvre or stage attacks. This information can be used to focus limited defensive resources to where the threat is greatest, raising the costs for attackers who must invest more to achieve desired effects. In 2019, the director of GCHQ, for example, described developing the ability to share 'time-critical, secret information in a matter of seconds' with increasing access for targeted businesses that are particularly critical.[47] In the US, Cyber Command officials have highlighted the potential to feed forward-derived information back to the Department of Homeland Security, the FBI, or through public–private information sharing and analysis centres to defend critical infrastructure.[48] Ahead of the 2018 election cycle, for example, US Cyber Command deployed personnel to countries with recent exposure to the election cyber threats it was then facing.[49] Working closely with the host countries, these 'hunt forward missions' provided details on malicious cyber actors and their operations that fed security preparations for US elections. The UK's National Cyber Security Centre has expanded this approach, sharing information on Russian cyber activity with 16 NATO Allies and a number of countries outside the Alliance to help counter threats.[50]

Second, forward cyber operations can also contribute to deterrence by denial by disrupting malicious cyber activity. Identifying and gaining access to adversary cyber operations infrastructure can provide the deterrer with the ability to contest threat actors on their terrain. Theoretically, the deterrer could then engage threat actors before they ever reach their targets. Varying degrees of disruption can be achieved across multiple points of an attacker's kill chain.[51] For example, forward cyber operations can be employed to affect an attacker's ability to conduct preparatory reconnaissance or disrupt command and control of its malware. Importantly, this is not just the realm of military action. Although outside the general conversation of forward cyber operations used for defence, law enforcement and economic government entities can and have performed similar functions. In the past, for example, the FBI and Microsoft have partnered to take down massive botnets using active means that resemble forward cyber operations.[52]

In addition to the direct effect described here, forward cyber operations for intelligence collection and malicious activity disruption might also influence an adversary's cost–benefit calculus by undermining trust in their offensive capability. The very presence of the deterrer's cyber

47. Jeremy Fleming, speech at CYBERUK conference, 24 April 2019, <https://www.gchq.gov.uk/speech/director-s-speech-at-cyberuk-2019>, accessed 15 March 2021.
48. Mark Pomerleau, 'In Era Of "Defend Forward", What Does Success Look Like?', *Fifth Domain*, 24 April 2019.
49. Nakasone and Sulmeyer, 'How to Compete in Cyberspace'.
50. Jeremy Hunt, speech at the National Cybersecurity Centre.
51. The 'cyber kill chain' outlines the steps involved in the conduct of cyber operations. See Lockheed Martin, 'The Cyber Kill Chain', <https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html>, accessed 15 March 2021.
52. *BBC News*, 'FBI and Microsoft Take Down $500m-Theft Botnet Citadel', 6 June 2013; Violet Blue, 'Microsoft, Law Enforcement Disrupt Sprawling Dorkbot Botnet', *ZDNet*, 4 December 2015, <https://www.zdnet.com/article/microsoft-law-enforcement-disrupt-sprawling-dorkbot-botnet/>, accessed 25 October 2020.

operators on adversary infrastructure, if detected, may lead to reduced confidence in whether they will be effective when needed. Recognition of compromises to cyber attack infrastructure raises questions of how much was revealed about their operations or whether that intrusion included placement of unseen countermeasure capabilities, both of which increase uncertainty about the likelihood of future operational success. Because potential attackers must consider the likelihood of success in their cost–benefit calculus, degrading confidence in whether their cyber capabilities will work increases the value of deterrence by denial.

As may be evident in the preceding discussion, using forward cyber operations for deterrence by denial also comes with implications for relations between government and private industry as well as between allied states. More specifically, maximising the effect of these operations requires robust collaboration across both of these lines. A two-way partnership between government and industry is essential. To orient their activities in a useful direction, government cyber operators benefit from information regarding the networks they aim to defend and competitor activity on these networks. In the opposite direction, cyber infrastructure owner-operators benefit from forward-derived threat intelligence to shift cybersecurity resources to where they are needed most. Since the vast majority of critical infrastructure is owned and operated by private entities, this approach necessitates a deep and dynamic relationship between government and key industry entities.

International partnerships can provide access to much-needed threat intelligence as well as a pathway for further raising adversaries' operational costs. The US Cyber Command's 'hunt forward' deployments provided insight into threat actor behaviour but required the invitation of host governments and the ability to work closely with existing and new partners. Furthermore, the wider intelligence on cyber threat activity is shared and then leveraged to increase the defence of targeted systems, the greater the operational costs to adversaries. The Five Eyes partners have bolstered sharing of timely cyber threat intelligence, for example, which is likely to raise the challenge to adversary operations across participating countries.[53]

As with deterrence by punishment and compellence, the risk of escalation and collateral effects is tied to the types of cyber operations and targets involved. Forward cyber operations for denial, whether in the form of intelligence collection or disruption activities, are focused on the systems that adversaries use to prepare and conduct their offensive cyber operations. This reduces the potential for both outright escalation and escalation due to misinterpretation, as well as unintended collateral effects. In contrast to the operations against high-value critical systems of deterrence by punishment and compellence, there is less cause for escalation. Focus on cyber operations infrastructure also reduces room to misinterpret denial actions as escalatory moves. Despite the generally lower risk in denial operations, there are still pathways to escalation and production of collateral effects. First, if denial targets overlap or connect with higher-value critical systems, room for misinterpretation of denial-intentioned activities

53. Five Country Ministerial, 'Official Communique 2018', <https://www.homeaffairs.gov.au/about-us/our-portfolios/national-security/security-coordination/five-country-ministerial-2018>, accessed 25 October 2020.

increases. Furthermore, operational mistakes can also undermine this lower-risk advantage. For instance, if an operation to disrupt malicious activity is designed in such a way that allows it to spread or take effect beyond the targeted system, the chance for collateral consequences increases. Even these risks are reduced, however, if the method of forward operations is through partner-provided access to information, as with the 'hunt forward' missions described above.

While still developing, some recent scholarship also suggests that reducing adversary cyber capability in this manner may incentivise escalatory behaviour.[54] It indicates that in some cases, states might opt for more escalatory options if denied the use of their lower-level, less escalatory cyber tools. For example, reducing Iran's ability to conduct cyber attacks as a way to retaliate for similar attacks against it might drive consideration of kinetic alternatives. In short, cyber operations offer states a way to engage in competition without significantly increasing tensions. They might operate as a sort of pressure release, offering stabilising, non-lethal options that are less threatening than traditional weapons.[55] Focusing forward operations on increasing defences rather than direct operational disruption may be a more desirable option in these instances.

54. Brandon Valeriano and Benjamin Jensen, 'How Cyber Operations Can Help Manage Crisis Escalation with Iran', *Washington Post*, 25 June 2019.
55. Jason Healey and Robert Jervis, 'The Escalation Inversion and Other Oddities of Situational Cyber Stability', *Texas National Security* Review (Vol. 3, No. 4, 2020), pp. 30–53.

III. Entanglement and Norms

COSTS FOR CYBER offence that might dissuade future malicious actions can also come from non-deterrent sources. For example, dissuasion by entanglement involves costs to offenders by virtue of being a part of interdependent systems. Alternatively, normative costs are associated with breaking collectively established bounds of behaviour.

Entanglement

The cost–benefit calculus of potential state cyber attackers may be influenced by the degree to which they are connected to interdependent systems (cyber and non-cyber) that their operations might affect.[56] Described as 'entanglement', this works when cyber attacks result in costs to members across the system including the attacker, not just the target.[57] Entanglement costs are realised when cyber attacks disrupt the economic, diplomatic and strategic relationships and the information systems that might underpin them. As an example, when making network partnering decisions, the organisations that manage the operation of the internet rely on reputation indicators and records of abuse or mismanagement in their decisions.[58] Theoretically, if states leveraged their private sector elements to conduct activities counter to the common interest (such as internet routing hijacking, which Chinese companies have been accused of), they risk potential exclusion from the system and the loss of gains it provides.[59] Potential attackers are dissuaded by both the current and future value they receive from their connection to these systems.

There are even indications that entanglement-based restraint may be exhibited in cyberspace. State-based cyber attacks on systems critical to the functioning of the global financial system appear to be largely confined to those states least likely to suffer from entanglement costs. Attacking critical financial services infrastructure is an attractive cyber target due to its high value and number of potentially vulnerable systems. However, the most prominent state cyber attackers of the global financial system are confined to those least connected to it: North Korea and Iran.[60]

---

56. Nye, 'Deterrence and Dissuasion in Cyberspace'.
57. Joseph S Nye Jr and Robert O Keohane, 'Power and Interdependence Revisited', *International Organization* (Vol. 41, No. 4, 1987), pp. 725–53.
58. Jesse H Sowell, 'Finding Order in a Contentious Internet', PhD dissertation, Massachusetts Institute of Technology, 2015.
59. Chris C Demchak and Yuval Shavitt, 'China's Maxim – Leave No Access Point Unexploited: The Hidden Story of China Telecom's BGP Hijacking', *Military Cyber Affairs* (Vol. 3, No. 1, 2018), pp. 1–9.
60. See, for example, US District Court, Southern District of New York, 'United States of America v. Ahmad Fathi, Hamid Firoozi, Amin Shokohi, Sadegh Ahmadzadegan, Omid Ghaffarinia, Sina Keissar, and Nader Saedi', 24 March 2016, <https://www.justice.gov/opa/file/834996/download>,

Forward cyber operations may be able to contribute to entanglement by influencing its cost mechanisms through transparency imposition. Forward-derived intelligence can be used to expose the details of offending activities to generate entanglement costs. For instance, China's long-running cyber-based industrial espionage campaign, while advancing immediate interests, may also impose longer-term diplomatic and economic costs on itself through reputation destruction in an international trade system on which it depends.[61] Huawei's struggle in selling 5G infrastructure in some states reflects growing concerns regarding future malicious behaviour by China, based in part on exposure of past behaviour.[62]

Furthermore, facilitating entanglement through transparency might also create accountability for cyber actors operating beyond the state's direction. Official state cyber actors or state-supported 'cyber mercenaries' may on occasion use their skills for personal gain by conducting criminal activity.[63] A recent example is found in a group of Chinese operators that hack into healthcare and telecommunications firms for the government and later conducts ransomware and digital currency attacks for personal profit.[64] Exposing this activity and its potential cost implications for the entangled state can generate recognition by national leadership that results in housecleaning and greater controls on cyber proxies.

Norms

Normative considerations can also dissuade aggressive cyber behaviour by imposing costs that 'damage an actor's soft power beyond the value gained from a given attack'.[65] The logic of normative-based restraint is founded on the idea that collective expectations of proper behaviour, from informal patterns to applied law, can lead to self-restraint.[66] Distinct from other theories of restraint, the influence of norms involves a sense of 'oughtness', a standard of appropriate behaviour for actors with a given identity.[67] They evolve in a 'life cycle' with a period of contested emergence before reaching a critical mass of acceptance (a tipping point), and

accessed 20 October 2020; Ben Buchanan, 'How North Korean Hackers Rob Banks Around the World', *Wired*, 28 February 2020.

61. Diane Bartz and Jack Stubbs, 'U.S., Allies Slam China for Economic Espionage, Spies Indicted', *Reuters*, 20 December 2018.
62. Lindsay Maizland and Andrew Chatzky, 'Huawei: China's Controversial Tech Giant', Council on Foreign Relations, 6 August 2020.
63. For additional work on cyber mercenaries, see Tim Maurer, *Cyber Mercenaries: The State, Hackers, and Power* (Cambridge: Cambridge University Press, 2018).
64. FireEye, 'Double Dragon: APT 41, a Dual Espionage and Cyber Crime Operation', August 2019, <https://content.fireeye.com/apt-41/rpt-apt41/>, accessed 25 October 2020.
65. Nye, 'Deterrence and Dissuasion in Cyberspace'.
66. Martha Finnemore, 'Cybersecurity and the Concept of Norms', Carnegie Endowment for International Peace, 30 November 2017; Edelman, 'Cyberattacks in International Relations'.
67. Martha Finnemore and Kathryn Sikkink, 'International Norm Dynamics and Political Change', *International Organization* (Vol. 52, No. 4, 1998), pp. 887–917.

then broad adoption and internalisation.[68] States and non-governmental entities have for years pursued efforts to build normative constraints on offensive cyber activity, but progress has been slow at best. Since 2004, for example, the UN has commissioned Groups of Government Experts to develop common ground on acceptable cyber behaviour.[69] The Estonian-led Tallinn Manual process aims to do the same from a legal perspective, discerning how existing international law applies to cyberspace.[70] Direct bilateral agreements have also been used to attempt to codify normative boundaries.[71] Even private industry players, impatient with struggling government efforts, are developing international norm-building processes.[72]

Critically, the influence of norms is based on more than written agreements. It relies on interactions between states that establish and maintain normative boundaries, and this is where forward cyber operations might contribute to norm-based restraint. First, forward operations can help develop and then enforce established normative boundaries by imposing costs that communicate precisely where bounds exist, recognition of their transgression and signalling potential escalation if ignored. Importantly, the costs for offenders do not need to reach outright deterrence level. However, this approach does suffer from the same communication and signalling challenges discussed above. Second and similar to entanglement, forward intelligence operations can be used to expose 'out of bounds' behaviour and engage normative mechanisms such as reputational costs. This imposed transparency may be particularly influential for states with demonstrated concern for reputation and international standing.

As with deterrence by denial, partnerships play a central role in pursuing restraint by entanglement and norms. The influence of both depends on recognition and action by a large enough group of international and private actors to generate sufficient costs to outweigh an attacker's potential gain. This itself requires trusted partnerships with capable private and allied state actors to facilitate information sharing, confidence in the veracity of shared information and consensus building.

As with deterrence by punishment, using forward cyber operations to impose direct costs opens the door to greater risk of escalation and collateral effects, though perhaps to a lesser degree. Creating disruptions in adversary systems increases the chance of generating an escalatory

68. *Ibid.*, p. 888.
69. UN Office for Disarmament Affairs, 'Developments in the Field of Information and Telecommunications in the Context of International Security', <https://www.un.org/disarmament/ict-security/>, accessed 15 March 2021.
70. NATO Cooperative Cyber Defence Centre of Excellence, *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (Cambridge: Cambridge University Press, 2017).
71. The White House, President Barack Obama, 'Remarks by President Obama and President Xi of the People's Republic of China in Joint Press Conference', Washington DC, 25 September 2015, <https://obamawhitehouse.archives.gov/the-press-office/2015/09/25/remarks-president-obama-and-president-xi-peoples-republic-china-joint>, accessed 15 March 2021.
72. Microsoft, 'A Digital Geneva Convention to Protect Cyberspace', Microsoft Policy Papers, <https://query.prod.cms.rt.microsoft.com/cms/api/am/binaryspo/RW67QH>, accessed 15 March 2021.

response (either outright or through misinterpretation). The lower-value targets that this approach involves, however, likely decreases this risk as compared to deterrence by punishment or compellence. The risk of producing unwanted collateral effects still remains. Imposing costs through transparency requires forward-derived information on adversary cyber activity, which, as with deterrence by denial, involves methods and targets that have a lower risk of escalation and collateral effects.

A final consideration in using forward operations to promote normative bounds is found in the irony that it presents. In a space where norms are largely nascent and ill-defined, forward operations in support of norms may themselves appear to be operating outside normative bounds. For instance, a cost-imposing cyber operation intended as a normative boundary signal to an offending state may be interpreted as more evidence of the unconstrained nature of competition in cyberspace. The fundamental differences in how leading cyber powers perceive what constitutes acceptable behaviour complicates their application for normative advances. The result in this case could be increased potential for escalation within cyberspace.

IV. Tacit Cooperation

SITUATED BETWEEN DETERRENCE and norms, restraint can also be developed between competitors through interactions that act as a process of tacit bargaining. The central logic here is that mutually agreed bounds of behaviour can evolve between adversaries through back-and-forth competitive interactions that build costs which, over time, outweigh the potential gains from transgression. Similar to deterrence, it relies on imposing costs on offenders. Unlike deterrence, however, it does not rely on increasing defences or passively threatening a response that outweighs the benefits of offence. Instead, it relies on actively imposing smaller costs that add up gradually. Similar to norms, the ultimate goal is to develop mutually agreed bounds of expected behaviour. Distinct from norms, however, these bounds are not established through a sense of 'oughtness'.

Much of the foundation for the concept of tacit cooperation was built during the Cold War by scholars seeking to discern the characteristics of competition under the shadow of nuclear weapons and high uncertainty of other states' intentions. Herman Kahn described how, through instrumental competitive interactions over time, adversaries develop a mutual understanding of the bounds and shape of their 'agreed battle': the legitimate and illegitimate moves, what is 'within the rules', and what is escalatory behaviour.[73] Schelling discussed how adversaries might have aligned interest in restraint in certain areas and how this common interest is often realised through tacit bargaining, or communication through action, rather than explicit negotiation under contexts of limited trust.[74] The result is shared expectations of conduct within a defined competitive space and general restraint beyond those bounds.[75]

The idea of tacit cooperation was further developed in the early 1980s using game theory to model how these interactions between competitors develop into mutual restraint.[76] The resulting theory proposed that tacit cooperation can develop through an exploratory process of trial and error, restraint and offence. Over the course of these iterated interactions, 'clear patterns of

73. Herman Kahn, *On Escalation: Metaphors and Scenarios* (London: Routledge, 2017); Fischerkeller, 'Persistent Engagement and Tacit Bargaining: A Strategic Framework for Norms Development in Cyberspace's Agreed Competition', p. 2.
74. Thomas Schelling, *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1960).
75. Today, these Cold War theoretical foundations have been adapted by Michael Fischerkeller and Richard Harknett to shape the US's operational approach through the Defend Forward cyber strategy, called Persistent Engagement. See Fischerkeller and Harknett, 'Persistent Engagement and Tacit Bargaining: A Path Toward Constructing Norms in Cyberspace'; Fischerkeller, 'Persistent Engagement and Tacit Bargaining: A Strategic Framework for Norms Development in Cyberspace's Agreed Competition'.
76. See Robert Axelrod, *The Evolution of Cooperation* (New York, NY: Basic Books, 1981); Kenneth A Oye, *Cooperation Under Anarchy* (Princeton, NJ; Princeton University Press, 1986).

mutually understood behaviour' may develop between egoists free of central control.[77] More specifically, tacit cooperation can develop when: individual actors have a sufficiently large chance to meet again that gives them a stake in their future interactions (long shadow of the future); and individual actors are prepared to reciprocate either cooperation or defection from cooperation. In short, continual reciprocal cost imposition for offending actions can lead to the expectation of future costs that add up to outweigh the gains from offence, establishing restraint for a particular type of action.

A commonly cited example that illustrates this concept is that of an evolved tacit truce, or a 'live-and-let-live' system, between British and German soldiers in certain areas along the western front during the First World War.[78] Robert Axelrod describes how in trench warfare, opposing units faced each other in immobile sectors over long periods of time, making possible the iterated interactions that can develop into patterns of agreed bounds of behaviour. In the 'quiet sectors' of the front, tacit cooperation evolved through iterated reciprocation of attack and restraint, first during certain periods such as meal or ration times and then expanding from there.[79]

This theory of restraint seems particularly attractive for relatively immature domains of interaction, like cyberspace, where bounds are not yet clearly defined and competitors are still developing basic capabilities and exploring their utility and costs. Furthermore, for states, the global landscape of interconnected networks that modern societies increasingly rely on is a space of shared vulnerability and iterated interaction. A context of common vulnerability and certainty in future interaction indicates that cyberspace is likely fertile ground for exploratory actions to develop into tacit bargaining that results in patterns of mutually understood bounds of behaviour. As might be expected, there is evidence that this exploratory process of iterated trial and error, action and response is already underway between states in cyberspace. For example, back-and-forth interactions that characterise this sort of tacit bargaining within an emerging area of agreed battle are found between numerous competitive dyads, including:

- Russian and US compromises of each other's power grids.[80]
- Israeli and Iranian disruptive attacks on each other's civilian critical infrastructure.[81]
- The two-decades-long cyber tit-for-tat between India and Pakistan.[82]

77. Axelrod, *The Evolution of Cooperation*; Robert Axelrod and William D Hamilton, 'The Evolution of Cooperation' *Science* (Vol. 211, No. 4489, 1981), pp. 1390–96.
78. Axelrod, *The Evolution of Cooperation*.
79. *Ibid.*, p. 79.
80. Musil, 'Russian Hackers Accessed US Electric Utilities' Control Rooms'; Sanger and Perlroth, 'U.S. Escalates Online Attacks on Russia's Power Grid'.
81. Ronen Bergman and David M Halbfinger, 'Israel Hack of Iran Port Is Latest Salvo in Exchange of Cyberattacks', *New York Times*, 19 May 2020.
82. Marie Baezner, 'Hotspot Analysis: Regional Rivalry Between India-Pakistan: Tit-for-Tat in Cyberspace', Center for Security Studies, ETH Zürich, August 2018.

Furthermore, it is not difficult to foresee how some of these interactions might lead to mutual restraint. For instance, in the Israel–Iran case, the costs of continued back-and-forth disruptive attacks on civilian infrastructure along with the demonstrated likelihood of future cost imposition for offence may begin to outweigh the benefits each gains from these types of attacks. In the latest iteration of this cyber competition, Israel responded to an Iranian attack on a regional water system by disrupting operations at a major Iranian port. Given that both countries are sufficiently wired to offer a substantial number of similar future targets, it is possible that both will at some point decide this component of their cyber competition no longer has utility, shaping their 'agreed battle' to exclude certain civilian infrastructures.

The basic requisites for cooperation theory (reciprocity and the promise of future interaction) reveal where forward cyber operations might have utility in promoting tacit cooperation as well as where it is likely limited. The fact that states, great and small, are increasingly interconnected and competing through cyberspace makes the possibility of future interaction a near certainty. Within this context, forward cyber operations can theoretically operate as a tool of reciprocation with their iterated use in return for offence, or non-use in return for non-offence. Although not as clearly connected, cooperation-building responses can also take the form of cross-domain actions that are facilitated by forward cyber operations, similar to those described in Chapter I. For example, intelligence on competitor cyber activities derived through forward operations can support the application of economic or diplomatic consequences that act as tit-for-tat responses.

Similar to the previous analysis of cyber operations in support of deterrence by punishment, however, clearly linking the response to the offending action is difficult. In particular, slow recognition or attribution of transgression can delay responses that perform better when temporally close to the offending action. Effective reciprocation requires quick and reliable information about the other's actions.[83] Even if offences are identified and attributed to a state actor, difficulties in determining intent can complicate tacit bargaining for both the victim and the offender who must interpret the meaning of the other's actions. This is yet another area where forward intelligence collection operations can help reduce the uncertainty surrounding adversary activities and intentions. Schelling's 'connectedness' may also help overcome some of this challenge. For instance, responding to transgressions with attacks on an offender's cyber operations infrastructure or on mirror targets can communicate direct linkage between offending behaviour and the response. For example, in addition to taking place soon after Iran's cyber attack on regional water infrastructure, Israel's retaliation targeted Iranian civilian infrastructure which made it clearer on the receiving end what out-of-bounds behaviour this response was directed at.

The importance of the clear tacit communication that this implies also extends to clarity over the desired bounds themselves. For two competitors to converge on a commonly understood bound of behaviour through tacit bargaining by action, simplicity is paramount. Distinctions between

83. Oye, *Cooperation Under Anarchy*.

acceptable or unacceptable actions must be all-or-none in form, not matters of degree.[84] This reveals another limitation in using forward cyber operations as a tacit bargaining tool: while they may appear fast in execution, they often require long lead times for preparation. In practice, this might lead to use of operations that trade availability for clarity in desired bounds.

Forward cyber operations may also have particular advantages when applied under cooperation theory. Whereas the lower costs that cyber operations are typically able to impose is a disadvantage under deterrence by punishment, it can work to its benefit here. Unlike deterrence by punishment actions, cooperation theory cyber responses can be smaller reprisals with lower risk, and thus greater credibility, as threats to potential offenders calculating future aggregated costs of transgression. Additionally, as some scholars have noted in non-cyber contexts, high potential costs for making cooperative gestures for non-offence may make them too risky, preventing this element of tacit bargaining interactions.[85] In cyberspace, however, most potential costs are low enough to reduce the risk involved with making cooperative gestures that help tacitly define acceptable bounds of behaviour.

The types and targets of forward cyber operations employed under cooperation theory can vary greatly; thus, so does the potential risk of escalation and collateral effects. Using cyber operations for tacit bargaining or intelligence collection to facilitate it (in cyberspace or cross-domain) is likely less escalatory when targeting the operational infrastructure that adversaries use to conduct their cyber attacks. Where this operational infrastructure is more isolated from other infrastructure, it may also pose less risk of producing collateral costs. However, where this operational infrastructure overlaps with or connects to other important infrastructure, the risk of escalation and collateral effects may rise with the value of those systems. For instance, if adversaries use military communications networks as part of their operational infrastructure, targeting this may be misinterpreted as an attempt to disable critical communications instead of as a response to a cyber offence. Similarly, if an adversary manoeuvres on third-party servers that are also used by other civil organisations as part of their operational infrastructure, then targeting this brings greater risk of collateral effects.

If forward cyber operations target systems beyond an adversary's operational infrastructure, as with in-kind reciprocation such as the Israel–Iran example above, then risk of escalation is determined by its value to the adversary. Furthermore, the risk of collateral costs depends on its exposure to other systems. It is worth noting here again that because cyber operations take time to develop, there may be a limit to the depth and clarity of reciprocation possible. Some decision-makers may be tempted to use whatever targets are available to respond in a timely manner, but this may come at the cost of increased risk.

84. Fischerkeller, 'Persistent Engagement and Tacit Bargaining: A Strategic Framework for Norms Development in Cyberspace's Agreed Competition', p. 12.
85. George W Downs and David M Rocke, 'Tacit Bargaining and Arms Control', *World Politics* (Vol. 39, No. 3, 1987), pp. 297–325.

Conclusion: Implications and Recommendations

THIS PAPER EXAMINES the utility and risks in using forward cyber operations to generate restraint along different theoretical pathways. It draws from existing international relations literature regarding restraint and cyber competition to conduct an analysis that informs critical policy decisions and advances understanding of the concepts themselves. The current conceptions of restraint were largely developed around the details of the Cold War nuclear competition. Applying them across the new bounds and features of cyber competition helps to expand knowledge on inducing restraint in cyberspace and more generally. For example, when considering cross-domain deterrence, this paper's analysis points to the enabling value of in-domain actions (such as cyber intelligence collection operations) to cross-domain strategies. It also further develops thinking on competitive interactions in cyberspace and how to use them to shape competitor behaviour. Expanding on Persistent Engagement's discussion of finding agreed bounds of behaviour through tacit bargaining,[86] this analysis leverages cooperation theory to develop the underlying mechanics of how that might work.

From a policy perspective, it expands the discussion of using forward cyber operations for defence beyond its focus on deterrence by punishment or simply actively contesting bad behaviour. There is ample cyber literature that details the limits of traditional deterrence by punishment approaches in cyberspace, but there are more paths to restraint than have been meaningfully considered in policy discussions.[87] In closing, this paper offers the three following high-level implications and associated recommendations for policymakers.

First, using forward cyber operations to induce restraint is not a straightforward matter of imposing costs. Instead, forward defence operations can affect costs and gains at multiple parts of an adversary's decision calculus. This includes reducing potential gains from attacks, and raising the costs of offensive operations, as well as imposing retaliatory costs (both in cyberspace and cross-domain). Additionally, the utility, requirements and risks involved vary by the theoretical pathway taken. For instance, the factors that influence the utility of forward cyber operations as a tool of outright deterrence by punishment or compellence, such as level of interest and cost-imposing capability, are distinct in degree and type from those that affect the others. The risk of escalation and unwanted collateral effects also varies by the target and operation type

86. See Fischerkeller and Harknett, 'Persistent Engagement and Tacit Bargaining: A Path Toward Constructing Norms in Cyberspace'; Fischerkeller, 'Persistent Engagement and Tacit Bargaining: A Strategic Framework for Norms Development in Cyberspace's Agreed Competition'.
87. See Lindsay, 'Tipping the Scales'; Schneider, 'Deterrence In and Through Cyberspace'; Michael P Fischerkeller and Richard J Harknett, 'Deterrence Is Not a Credible Strategy for Cyberspace', *Orbis* (Vol. 61, No. 3, 2017), pp. 381–93.

needed to pursue restraint under each theory. In light of this, current cyber strategy development should expand scope to consider multiple paths to restraint, accounting for their distinct utility and risks, to offer policymakers better options that increase flexibility in addressing a broader range of cyber threats (above and below the threshold of an armed attack). In general, the utility–risk analysis here suggests that use of forward cyber operations for restraint should prioritise intelligence collection for deterrence by denial and as an enabler of other restraint pathways, and targeting adversary cyber-operations infrastructure.

Second, a state's ability to leverage partnerships, both internationally and domestically between government and industry, influences its ability to affect an adversary's offensive decision calculus. The majority of theoretical pathways to restraint are partnership dependent or, at least, partnership amplified. States with existing partnerships or the ability to develop new trusted partnerships that enable information exchange and collaboration possess a strategic advantage over those that do not. As such, it is essential to develop new or enhance existing trusted partnerships that can exchange and leverage forward-derived information, thus expanding the power of entanglement, normative and denial paths to restraint. This involves government partnerships with industries that own and operate the vast majority of cyberspace, including the critical infrastructure being defended. A tighter partnership with industry is required to better understand defended terrain and to share sensitive information to bolster defences. This points to developing a trusted and dynamic operational-level partnership with critical industry elements that aligns a country's disparate public and private capabilities. Public–private partnerships for cybersecurity are certainly not a new idea, but the depth of cooperation this calls for is a more recent consideration and likely demands a reconfiguration of partnership approaches. For instance, among a number of requirements, tailoring partnership initiatives to leverage and account for the idiosyncrasies of sectors' market dynamics and existing relationship structures is essential.[88] Some might argue that closer government–industry interaction will open the door to potential civil liberties compromises and there is reason for concern here. However, states can address these concerns by building in protections that ensure their cybersecurity partnerships are consistent with their values.[89]

Third, effective forward defence involves more than intelligence or military organisations. Despite being the primary focus of the current discourse on forward cyber defence, these are not the only entities equipped with the tools and authorities to perform them. Other areas of government possess distinct expertise, relationships and capabilities that can produce powerful forward effects. The example of the FBI partnering with capable private sector entities such as Microsoft for large-scale botnet takedowns highlights this. Accordingly, it is critical to meaningfully integrate a broader range of capable actors (beyond the intelligence and military communities) into forward cyber operations and strategies. As several scholars have highlighted,

88. Sean Atkins and Chappell Lawson, 'An Improvised Patchwork: Success and Failure in Cybersecurity Policy for Critical Infrastructure', *Public Administration Review* (2020), pp. 1–15.
89. Sean Atkins and Chappell Lawson, 'Regulation or Integration? Cybersecurity for Critical Infrastructure and its Implications for Business-Government Relations' (currently under review for publication).

possessing capability alone is not the full measure of state power.[90] Being able to organise and employ it to effect is just as important. Considerable organisational and political impediments to achieving this exist, however, and overcoming these will likely require greater centralised control of the disparate organisations and activities involved.[91] The US, for instance, recently established a White House-level Office of the National Cyber Director, but it remains to be seen whether the authorities and resources aligned will be sufficient.

As offensive cyber activities continue to trend in a dangerous direction, some states are turning to defence strategies that employ forward cyber operations. An oft-cited aim of this approach is to induce restraint in adversaries but there is a lack of clarity regarding how exactly this is to be achieved. Because of the potential risks associated with this approach, policymakers should redouble efforts to explore and clearly identify desired end-states and the logics that link cyber means to strategic objectives. There are a number of potential pathways and, without doing this, cyber operations in pursuit of restraint risk disconnection from a strategic logic that guides their employment, potentially steering states in undesired directions.

90. See, for example, Stephen Biddle, *Military Power: Explaining Victory and Defeat in Modern Battle* (Princeton, NJ: Princeton University Press, 2004). Separately, Slayton also discusses this when considering the offence–defence balance in cyberspace, though on a micro level (individual organisations or operations). See Slayton, 'What Is the Cyber Offense-Defense Balance?'.
91. John Costello and Mark Montgomery, 'How the National Cyber Director Position Is Going to Work: Frequently Asked Questions', *Lawfare*, 24 February 2021.

About the Author

Sean Atkins is a political science PhD candidate at the Massachusetts Institute of Technology and an active duty US Air Force officer. His research focuses on national defence and state competition in cyberspace. His military service includes experience in cyber operations and national cyber policy.