European Journal of International Security (2020), **5**, 195–214
doi:10.1017/eis.2020.2

EJIS

RESEARCH ARTICLE

Beyond tit-for-tat in cyberspace: Political warfare and lateral sources of escalation online

Christopher Whyte*

Virginia Commonwealth University, Richmond, Virginia
*Corresponding author. Email: cewhyte@vcu.edu

(Received 12 August 2019; revised 26 March 2020; accepted 27 March 2020)

Abstract
At present, most scholarship on the potential for escalation in cyberspace couches analysis in terms of the technological dynamics of the domain for relative power maneuvering. The result has been a conceptualisation of the logic of operation in cyberspace as one of 'tit-for-tat' exchanges motivated by attribution problems and limited opportunity for strategic gain. This article argues that this dominant perspective overlooks alternative notions of how cyber tools are used to influence. This, in turn, has largely led scholars to ignore second-order effects – meaning follow-on effects triggered by a more direct outcome of an initial cyber action – on domestic conditions, institutions, and individual stakeholders. This article uses the case of cyber-enabled political warfare targeting the United States in 2016 to show how escalation can occur as a second-order effect of cyber operations. Specifically, the episode led to a re-evaluation of foreign cyber strategy on the part of American defence thinkers that motivated an offensive shift in doctrine by 2018. The episode also directly affected both the political positions taken by important domestic actors and the attitude of parts of the electorate towards interference, both of which have reinforced the commitment of military planners towards assertive cyber actions.

Keywords: Escalation; Cyber; Political Warfare

Introduction

In the burgeoning literature on global cyber conflict, International Relations (IR) scholars are increasingly focused on questions of prospective escalation between states that engage one another online.[1] As nation-states have increasingly turned to the use of cyber instruments to better serve their interests, international interactions have been characterised by a broad-scoped intensification and diversification of the ways in which digital action impacts real world security and political dynamics. From cyber-enabled interference in political processes to ongoing

[1] Among others, see Vincent Manzo, 'Deterrence and escalation in cross-domain operations', *Joint Force Quarterly*, 66 (2011), pp. 8–14; Herbert Lin, 'Escalation dynamics and conflict termination in cyberspace', *Strategic Studies Quarterly*, 6:3 (2012), pp. 46–70; Martin C. Libicki, *Crisis and Escalation in Cyberspace* (Santa Monica: Rand Corporation, 2012); David C. Gompert and Martin Libicki, 'Cyber warfare and Sino-American crisis instability', *Survival*, 56:4 (2014), pp. 7–22; Ronald J. Deibert, 'Bounding cyber power: Escalation and restraint in global cyberspace', *Organized Chaos: Reimagining the Internet* (2014), p. 49; Brandon Valeriano and Ryan C. Maness, *Cyber War Versus Cyber Realities: Cyber Conflict in the International System* (New York: Oxford University Press, 2015); Nadiya Kostyuk, Scott Powell, and Matt Skach, 'Determinants of the cyber escalation ladder', *The Cyber Defense Review*, 3:1 (2018), pp. 123–34; Sarah Kreps and Jacquelyn Schneider, 'Escalation firebreaks in the cyber, conventional, and nuclear domains: Moving beyond effects-based logics', *Journal of Cybersecurity*, 5:1 (2019); Mischa Hansel, 'Cyber-attacks and psychological IR perspectives: Explaining misperceptions and escalation risks', *Journal of International Relations and Development*, 21:3 (2018), pp. 523–51; Brandon Valeriano, Benjamin M. Jensen, and Ryan C. Maness, *Cyber Strategy: The Evolving Character of Power and Coercion* (Oxford: Oxford University Press, 2018).

economic warfare efforts and attacks on critical infrastructure, cyber incidents that would unarguably feature on a shortlist of the most significant security incidents affecting Western countries in just the past few years have been manifold in their form.

For policymakers and researchers alike, the problematic and subsequent policy challenges associated with rising tides of cyber conflict are twofold. First, it seems likely that cyber aggression will increasingly be seen as an attractive option for states in responding generally to discord and conflict.[2] Given the manner in which the Internet has opened new space for contention below the threshold of militarised dispute in recent decades, this assumption is perhaps an unsurprising one. Second, the odds seem to be growing – at least according to some, such as Martin Libicki – that the United States and its partners might find themselves embroiled in an escalatory crisis that emerges more specifically from a major cyberattack of one sort or another.[3] Proponents of this position cite heightened public awareness of digital insecurities and the shift in strategic posture by the United States – and undoubtedly, in the future, partner nations – to 'defend forward' and set boundaries for conflict interactions by selectively engaging in limited escalatory offensive operations as bases for concern.[4]

At present, most scholarship on the potential for escalation from the use of cyber instruments couches analysis in terms of the implications of technological dynamics of the domain for relative power calculations and manoeuvring. The result has been a conceptualisation of the logic of operation in cyberspace as one of 'tit-for-tat' exchanges motivated by attribution problems and limited opportunity for strategic gain. This article argues that this dominant perspective overlooks alternative notions of how cyber tools are used to influence. This, in turn, has largely led scholars to ignore second-order effects – meaning follow-on effects triggered by a more direct outcome of an initial cyber action – on domestic conditions, institutions and individual stakeholders. I use the case of cyber-enabled political warfare targeting the United States in 2016 to show how escalation can occur as a second-order effect of cyber operations. Specifically, the episode led to a re-evaluation of foreign cyber strategy on the part of American defence thinkers that motivated an offensive shift in doctrine by 2018. The episode also directly affected both the political positions taken by important domestic actors and the attitude of parts of the electorate towards foreign interference, both of which have reinforced the commitment of military planners towards assertive cyber actions over the prospect of deterrence by defensive denial. The case also highlights other possible mechanisms of lateral escalation beyond what prevailing 'tit-for-tat' thinking on cyber conflict implies.

The remaining sections of this article proceed in several stages. After discussing the broad literature on the utility and risks of escalation in international relations and cyberspace within the IR field, I make the overarching argument that scholars must increasingly look beyond technical and operational factors related to the employment of cyber instruments if they are to effectively explain issues of cyber escalation, coercion, and more. This intercession is a reasonably straightforward one that aligns with the now almost-ubiquitous consensus position that cyber operations most often present in international relations as a modifier of other conflict processes.[5] Surprisingly, however, this article stands virtually alone among recent pieces of scholarship in pointing out that, in spite of such a dynamic, scholars yet rarely consider the logic of conflict processes *augmented* by cyber applications. Rather, despite the increasingly multifaceted and irregular nature of the attack surface of states to cyber-enabled conflict actions encountered in recent years, dominant approaches continue to emphasise technical features of the domain in

[2] Valeriano, Jensen, and Maness, *Cyber Strategy*.

[3] Libicki, *Crisis and Escalation in Cyberspace*.

[4] Alex Wilner, 'Cyber deterrence and critical-infrastructure protection: Expectation, application, and limitation', *Comparative Strategy*, 36:4 (2017), pp. 309–18; Alex S. Wilner, 'US cyber deterrence: Practice guiding theory', *Journal of Strategic Studies*, 43:2 (2020), pp. 245–80.

[5] Erik Gartzke, 'The myth of cyberwar: Bringing war in cyberspace back down to earth', *International Security*, 38:2 (2013), pp. 41–73.

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

analyses of escalation risks, deterrence opportunities, and more, with the result that the cyber conflict studies research programme most commonly frames research around questions of military-to-military engagement and the utility of cyber instruments from the perspective of traditional rationalist perspectives within the IR field. This article pushes back on such practice by arguing that, since cyber operations principally act to modify other conflict processes, method-specific detail cannot entirely drive analyses of strategic dynamics; instead, the underlying logic of those conflict processes should, with new methods thereafter considered in context.[6] To illustrate the validity of the overarching argument, the proceeding sections then examine the phenomenon of cyber-enabled political warfare and flesh out two arguments about analysis of cyber-related conflict and the under-realised escalatory potential of such operations from non-technical factors. First, I argue that the primary focus of existing cyber conflict scholarship on relative power contestation fails to predict avenues towards escalation of interstate hostilities brought on by the employment of cyber instruments. Then, using the case analysis of cyber-enabled political warfare waged in the United States by the Russian Federation during the 2014–16 presidential election season, I demonstrate the real opportunity for escalation given such operations, highlighting at least six opportunities therefor and two clear instances thereof in the case itself.[7]

Escalation: Realities and purpose in cyberspace and out

A relatively new focus of study within the cyber conflict studies field, escalation has received a fairly limited standalone treatment by scholars in IR. The seeming consensus position is, for now, that cyber escalation has not really taken place between states in particularly meaningful ways, either in terms of escalatory behaviour purely within the digital domain or in terms of non-cyber actions prompted by cyberattack. As such, engagements with the subject are almost entirely framed in terms of concern about *future* escalatory behaviour in cyberspace.[8]

To some degree, of course, this overstates the dynamics of the past. Where states hack one another, they regularly respond in kind. It is the exception rather than the rule, however, that cyber incidents prompt rapid and proportional (or escalatory) responses Brandon Valeriano,

[6] This acknowledgement can be found in a range of works on cyber conflict processes (among others, Jon R. Lindsay, 'Stuxnet and the limits of cyber warfare', *Security Studies*, 22:3 (2013), pp. 365–404; Christopher Whyte, 'Ending cyber coercion: Computer network attack, exploitation and the case of North Korea', *Comparative Strategy*, 35:2 (2016), pp. 93–102; Jon R. Lindsay, and Erik Gartzke, 'Coercion through cyberspace: The stability-instability paradox revisited', in K. M. Greenhill and P. J. P. Krause (eds), *The Power to Hurt: Coercion in Theory and in Practice* (New York: Oxford University Press, 2016); Christopher Whyte, 'Dissecting the digital world: A review of the construction and constitution of cyber conflict research', *International Studies Review*, 20:3 (2018), pp. 520–32; Michael Poznansky and Evan Perkoski, 'Rethinking secrecy in cyberspace: The politics of voluntary attribution', *Journal of Global Security Studies*, 3:4 (2018), pp. 402–16. Though the study of Valeriano, Jensen, and Maness (*Cyber Strategy*) stands somewhat apart in that the authors inductively study the contours of escalation in response to cyber incidents and campaigns, their analysis nevertheless only tees up a logical question that has yet to be addressed by scholars: might minor cyber actions that enable other forms of conflict be as likely to provoke an escalatory response as major incidents? Might such minor actions produce disproportionate escalatory risks?

[7] This term is often used interchangeably within discussions of cyber conflict alongside others such as active measures, hybrid warfare, irregular warfare, and information warfare. For work on hybrid or political warfare as it intersects with cyber operations in the recent Russian context, see, among others, Andrew Monaghan, 'The "war" in Russia's "hybrid warfare"', *Parameters*, 45:4 (2016), pp. 65–74; Alexander Lanoszka, 'Russian hybrid warfare and extended deterrence in eastern Europe', *International Affairs*, 92:1 (2016), pp. 175–95; Bettina Renz, 'Russia and "hybrid warfare"', *Contemporary Politics*, 22:3 (2016), pp. 283–300; Christopher S. Chivvis, 'Understanding Russian hybrid warfare', *Rand Corporation* (2017); and Benjamin Jensen, 'The cyber character of political warfare', *The Brown Journal of World Affairs*, 24 (2017), p. 159.

[8] Lin, 'Escalation dynamics and conflict termination in cyberspace'; Libicki, *Crisis and Escalation in Cyberspace*; David C. Gompert and Martin Libicki. 'Waging cyber war the American way', *Survival*, 57:4 (2015), pp. 7–28; Deibert, 'Bounding cyber power'; Kreps and Schneider, 'Escalation firebreaks in the cyber, conventional, and nuclear domains'; Hansel, 'Cyber-attacks and psychological IR perspectives'.

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

Benjamin Jensen, and Ryan Maness note the important difference between these two assessments in their eschewing of the term 'cyberattack' as somewhat inappropriate for those who study cyber conflict.[9] Certainly, states engage one another with thousands of individual attacks (that is, 'cyberattacks') online daily. However, incidents of strategic significance are much more relatively rare (they count only 192 in 16 years since the year 2000). In those rare cases of escalation (about 12 per cent of all cases), response actions occurred within sixty days (most within thirty) and were almost entirely incremental, with most of the few clear retaliatory actions taken by states since 2000 being only marginally more severe than was the attack to which they responded.

Valeriano, Jensen, and Maness thus articulate what is so far the prevailing argument about how cyber operations are employed as escalatory instruments by states in positing 'that low-level cyber actions are signaling mechanisms designed to limit future escalation and establish credibility'.[10] The point of burning vulnerabilities in response is not to demonstrate an ability to act so much as it is to show willingness to act. Efforts to credibly breach foreign networks are intended to increase the expected likelihood, in the calculus of an opponent, that the initial victim will continue to successfully intrude into critical systems should the conditions call for it. However, the nature of cyber instruments makes the signal necessarily limited in scope of its ability to deter further escalation. The signal is necessarily ambiguous and must be employed rapidly in order to achieve the needed signalling effect. Thus, the escalatory dynamic between states that engage online is invariably one of 'tit-for-tat' wherein cyber intrusions can be employed to limit future action but have an extremely short half-life.

Despite a general lack of escalatory behaviour linked with the fifth domain to date, increasing use of cyber instruments across all types of international conflict coupled with the nature of such tools as both prospectively low-cost/high-gain and inherently prone to ambiguity naturally nevertheless implies significant opportunities therefor. Given this, it is necessary to consider what IR scholars have said about the value proposition of escalatory behaviour, the risks involved and how cyber conflict might specifically impact such dynamics.

Why states do and do not escalate

Under certain circumstances, a state may choose to respond to a foreign provocation with an escalation specifically because it believes it stands to gain a military advantage of some kind.[11] Though there are obvious problems with this motivation for escalating conflict, sometimes escalation nevertheless can be beneficial wherein a strategic advantage – geostrategic positioning, force posture, new technology, etc. – for the initial victim state has gone unrealised. An aggressor state might, for instance, simply believes that some held advantage is unrealised by the opposing power and seek to act first.

Another cause of escalation is the desire to demonstrate military might.[12] This kind of escalatory posture is often thought of as a 'swaggering' posture.[13] Here, the point is to show off national strength and, more specifically, demonstrate capabilities in some detail. Vulnerabilities

[9] Valeriano, Jensen, and Maness, *Cyber Strategy*, p. 117.

[10] Ibid., p. 49.

[11] Daniel S. Geller, 'Nuclear weapons, deterrence, and crisis escalation', *Journal of Conflict Resolution*, 34:2 (1990), pp. 291–310; Lisa J. Carlson, 'A theory of escalation and international conflict', *Journal of Conflict Resolution*, 39:3 (1995), pp. 511–34; Forrest E. Morgan, Karl P. Mueller, Evan S. Medeiros, Kevin L. Pollpeter, and Roger Cliff, *Dangerous Thresholds: Managing Escalation in the 21st Century* (Santa Monica: RAND Corporation, 2008); Libicki, *Crisis and Escalation in Cyberspace*.

[12] Perhaps the best discussion of this is found in Thomas C. Schelling, 'Arms and influence', in *Strategic Studies* (New Haven, CT: Yale University Press, 2008), pp. 96–114 and in Herman Kahn, *On Escalation: Metaphors and Scenarios* (Abingdon: Routledge, 2017).

[13] Craig Neuman, and Michael Poznansky, 'Swaggering in cyberspace: Busting the conventional wisdom and cyber coercion', *War on the Rocks*, 28 (2016); Borghard and Lonergan, 'The logic of coercion in cyberspace'; Valeriano, Jensen, and Maness, *Cyber Strategy*; Max Smeets and Herbert S. Lin, 'Offensive Cyber Capabilities: To What Ends?', 10th International Conference on Cyber Conflict (CyCon) (IEEE, 2018), pp. 55–72.

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

are burned and a certain amount of strategic capital is spent so as to make clear to the original attackers the full extent of the victim's capacity to respond. Though there are risks, states have historically escalated conflict in this fashion under the assumption that a dramatic escalation of the stakes involved in a dispute will make the other side blink. A related aim of escalatory behaviour is attempting to test the extent of an opponent's knowledge.[14] If an opponent responds to escalation in a way that the victim deems to be within a range of reasonable responses based on the victim's strategic calculus, then subsequent steps can be taken with greater certainty.

Perhaps the most common purpose of one escalatory signal or another is the signalling of resolve and of serious intent towards future conflict.[15] As is the case with an attempt to gain military advantage, the manner in which escalation occurs matters a great deal because there needs to be nuance in the messages being sent. Resolve can be employed to a range of potential outcomes. One of the most common is deterrence by punishment wherein a state attempts to dissuade future deviation from whatever the status quo behaviour has been by clearly signalling an unwillingness permit conflict above a certain threshold. The point, in short, is to signal that a boundary has been reached beyond which there will be no tolerance of further conflictual behaviour.

Of course, escalation is risky for several reasons. These might broadly be divided into two categories relating to (1) the process of assessment and decision-making surrounding the direction of conflict actions and (2) the domestic context of decision-making and posturing. On the former, a major school of thought places much explanatory value on the nature of state force structures and military strategies, enshrined in doctrine and functionally shaped by both institutional experience and technological priors. Security planners and military leaders often hold specific versions of future conflict in their head.[16] In this way, particular conflictual actions on the part of opposing actors are more likely to lead to escalation than others regardless of the results of a rational security calculus. Some scenarios simply 'fit the model' and encourage an ignorance of context that might otherwise dissuade the victim state from responding aggressively. And there is often a knowledge gap between leaders' understanding of broad strategic posture and the details of battlefield or supply chain possibility. The result can be a lack of appreciation for the dangers involved in escalation.

More generally, decision-makers can suffer from a set of heuristic conditions that skew the results of the otherwise rational security calculus being produced in crisis situations.[17] Some are the result of institutional process. Others stem from natural psychological outputs of crisis conditions. Leaders of victim states will invariably assume that foreign attack is an indication of aggressive intentions *vis-à-vis* that state's security, as opposed to signalling linked with alliance posturing or simply an effort to meddle below the threshold of armed conflict. At the same time, leaders of victim states often do not consider their position in the same terms. And there is often an overstated centrality assumption in the way that states engaging in conflictual activities think of one another.[18] Both assume that aggression emerging from the other is the result of centralised, controlled decision-making processes and that the other state views the other as a core element of what determines foreign policy, resulting in over-interpretation of foreign intentions, underestimation of one's influence on others and even overconfidence.

[14] Libicki, *Crisis and Escalation in Cyberspace*, p. 75.

[15] See, among others, Paul Huth and Bruce Russett, 'Deterrence failure and crisis escalation', *International Studies Quarterly*, 32:1 (1988), pp. 29–45; Geller, 'Nuclear weapons, deterrence, and crisis escalation'; James D. Fearon, 'Signaling versus the balance of power and interests: An empirical test of a crisis bargaining model', *Journal of Conflict Resolution*, 38:2 (1994), pp. 236–69; David Kinsella and Bruce Russett, 'Conflict emergence and escalation in interactive international dyads', *The Journal of Politics*, 64:4 (2002), pp. 1045–68; Kahn, *On Escalation*.

[16] Barry R. Posen, *Inadvertent Escalation: Conventional War and Nuclear Risks* (Ithaca: Cornell University Press, 2014).

[17] See, for instance, Ithiel de Sola Pool and Allan Kessler, 'The Kaiser, the Tsar, and the computer: Information processing in a crisis', *American Behavioral Scientist*, 8:9 (1965), pp. 31–8.

[18] Robert Jervis, *Perception and Misperception in International Politics: New Edition* (Princeton: Princeton University Press, 2017).

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

With regards to domestic conditions, a wide-ranging literature has highlighted the manner in which public sentiment can affect leaders' interpretation of events and options available to them during a crisis.[19] In some cases, broad and excessive public confidence – often developed by a military-government complex intent on selling one or other security solution to a democratic population – in a particular national capability for dealing with outside threats might undermine the nature of the cost-benefit calculus carried out by leaders, particularly insofar as foreign states are unlikely to assess such capabilities via the same tinted lens. And, of course, leaders sometimes choose to incur audience costs, taking upon themselves the risk of political loss so as to more credibly signal seriousness during a crisis.[20]

The utility and risks of cyber escalation

Off the bat, it's easy to see why state leaders, operational decision-makers, and strategic planners might favour the use of cyber instruments of statecraft over others for managing escalation to some benefit. In a broad sense, the most recent information revolution centred on Internet-enabled computer technologies has opened up significant new room for contestation beneath the threshold of armed conflict. Moreover, digital antagonism can function as a useful toolkit for those interested in introducing some measure of ambiguity into escalatory dynamics.[21] Wherein there is confusion about the proportional value of, say, a denial of service attack to a kinetic disruption of an opponent's supply chain, there is arguably less chance of automatic escalation based on preconfigured protocols therefore.[22]

Within this argument, it is often also noted that cyber instruments are relatively cheap to produce and deploy.[23] There is, naturally, some nuance to had here: sophisticated digital weapons take time and resources to develop, and also lose value once revealed.[24] Thus, the use of some costly instruments will often be undesirable because of promised future efficiency. Even here, though, there is value to be had in the sending of clearer, stronger signals of resolve and intent. And where other signals occasionally lack nuance, cyberattack offers opportunities for measured assault on a foreign asset. In this, the attraction for decision-makers of an often-cheap, highly

[19] Such as the literatures on the marketplace of ideas (Jack Snyder and Karen Ballentine, 'Nationalism and the marketplace of ideas', *International Security*, 21:2 (1996), pp. 5–40; Chaim Kaufmann, 'Threat inflation and the failure of the marketplace of ideas: The selling of the Iraq War', *International Security*, 29:1 (2004), pp. 5–48; Trevor A. Thrall, 'A bear in the woods? Threat framing and the marketplace of values', *Security Studies*, 16:3 (2007), pp. 452–88) and the democratic peace (John M. Owen, 'How liberalism produces democratic peace', *International Security*, 19:2 (1994), pp. 87–125; Christopher Layne, 'Kant or cant: The myth of the democratic peace', *International Security*, 19:2 (1994), pp. 5–49; David Kinsella and Bruce Russett, 'Conflict emergence and escalation in interactive international dyads', *The Journal of Politics*, 64:4 (2002), pp. 1045–68.

[20] See, for instance, Kenneth A. Schultz, 'Looking for audience costs', *Journal of Conflict Resolution*, 45:1 (2001), pp. 32–60; Branislav L. Slantchev, 'Politicians, the media, and domestic audience costs', *International Studies Quarterly*, 50:2 (2006), pp. 445–77; Michael Tomz, 'Domestic audience costs in international relations: An experimental approach', *International Organization*, 61:4 (2007), pp. 821–40.

[21] Valeriano, Jensen, and Maness, *Cyber Strategy*, as well as Ben Buchanan, *The Cybersecurity Dilemma: Hacking, Trust, and Fear between Nations* (Oxford: Oxford University Press, 2016); Whyte, 'Ending cyber coercion'; Erik Gartzke and Jon R. Lindsay, 'Weaving tangled webs: Offense, defense, and deception in cyberspace', *Security Studies*, 24:2 (2015), pp. 316–48; Rebecca Slayton, 'What is the cyber offense-defense balance? Conceptions, causes, and assessment', *International Security*, 41:3 (2017), pp. 72–109; Benjamin Jensen, Brandon Valeriano, and Ryan Maness, 'Fancy bears and digital trolls: Cyber strategy with a Russian twist', *Journal of Strategic Studies*, 42:2 (2019), pp. 212–34.

[22] Kostyuk, Powell, and Skach, 'Determinants of the cyber escalation ladder'.

[23] Adam P. Liff, 'Cyberwar: A new "absolute weapon"? The proliferation of cyberwarfare capabilities and interstate war', *Journal of Strategic Studies*, 35:3 (2012), pp. 401–28; Thomas Rid and Ben Buchanan, 'Hacking democracy', *SAIS Review of International Affairs*, 38:1 (2018), pp. 3–16; Gartzke and Lindsay, 'Weaving tangled webs'.

[24] Erica D. Borghard and Shawn W. Lonergan, 'The logic of coercion in cyberspace', *Security Studies*, 26:3 (2017), pp. 452–81.

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

targetable, non-violent instrument that is seemingly well-suited to proportional responsive actions is clear.

Cyber instruments might also be attractive go-to options for political leaders and their adjuncts wherein contestation is possible beyond the public eye. For leaders concerned about the commitment required by more visible retaliatory options, relative secrecy is prospectively a godsend, particularly when one considers the domain's attribution dynamics. Even if opponents were to attempt to name and shame an escalatory political leader, it is sometimes possible to maintain plausible deniability and avoid audience costs. Indeed, cyber instruments might be an attractive option even for leaders that are willing to incur such costs, as the delayed commitment involved provides flexibility in the early stages of a confrontation.

These prospective benefits aside, retaliatory cyber engagement carries distinct risks. The domain's attribution challenges – particularly in bridging the gap between technical detail and sociopolitical context – can exacerbate the psychological problems outlined above, prompting use of assumptions not backed by evidence and often fitting an expected model of interaction. This is particularly problematic for conflict in cyberspace because of the manner in which the security dilemma manifests. Briefly, is extremely difficult to know if an offensive cyber action constitutes (1) a major assault; (2) a reconnaissance prelude thereto; (3) a simple signalling attempt; or (4) an everyday intelligence-gathering probe.[25] From the attacking side, it is remarkably difficult to predict with any certainty how signals might land on the other end.[26] Assumptions of executive authority risk overestimating the value of what the victim considers to be a clear, declaratory responsive action. On the flip side, supposing that cyber interactions are a form of inter-institutional signalling wherein security forces are taking the measure of one another risks the miscommunication of political or geopolitical consequence in an escalatory action.

Beyond the psychological elements involved, there are also more fundamental issues that emerge from cyber capabilities to signal and escalate conflict. With cyber, the meaning of specific actions is not a particularly generalisable thing. In-domain approaches do not look the same across national defence establishments. Particularly outside the West, where cyber capacity is often housed within intelligence units or flavoured by the experiences of quelling internal dissent, different operational red lines mean that some actions will inevitably carry more or less weight than they do across Europe or North America. And context matters on two fronts. First, probing attacks may be an everyday occurrence but patterns matter. Targeted probes that depart from the norm in some fashion may be seen with great alarm. This means, second, that cyber strategy is not inherently static. Particular context shapes operational assumptions and objectives that are then enshrined in doctrine over time.

The use of cyber instruments for escalation also contains added risks for misstep, notably emerging from unintended consequences in usage. This can look like a few things. An intended effect may ultimately produce some other outcome, causing clear potential for misperception of intentions. The result may also be a fizzle. Code may not execute as intended with the result being that an attack is not noticed. From the defender's perspective, such consequences may combine with imperfect view into enemy systems and again prompt decision-makers to make assumptions about an opponent's intentions and perception of events.

Finally, added complexity with cyber interactions certainly means more opportunities for non-violent signalling and for precision strike. But there is also the possibility that an attacker or counterattacker will miscalculate and push over the thin line between effective signalling and the loss of an advantage. Since, again, the value in cyberweapons is so often part and parcel of the secrecy in which they are developed and readied, an over-commitment for the purposes of effective signalling risks burning the wrong vulnerability and ceding a strategic advantage to an opponent.

[25] Buchanan, *The Cybersecurity Dilemma*.

[26] Gompert and Libicki, 'Cyber warfare and Sino-American crisis instability'.

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

Cyber operations as an enabler of other conflict

The remainder of this article forwards the argument that prevailing perspectives on cyber escalation miss some of the needed discussion to be had about the potential for future retaliatory dynamics within the domain. Here, I specifically take up the subject of cyber-enabled political warfare aimed at 'hacking' democratic political processes as a means of describing the ways in which cyber operations often enable disruption or manipulation beyond the immediate purview of computer network intrusions. Though there are exceptions (for example, Haggard and Lindsay 2015; Whyte 2018), studies of coercion and escalation within the domain tend to tacitly eschew any assumption that cyber operations might be employed without consideration of at least some direct impact on existing security relationships.[27] I begin by discussing the logic of anti-democratic information warfare seen in recent years.

The strategic logic of anti-democratic information warfare

At first blush, there are several different perspectives that we might adopt so as to develop context and meaning with regards to the anti-democratic interference campaigns conducted against Western states by the Russian Federation and other belligerent world powers over the past decade. The personal vendetta that American intelligence assessed as significant in the direction given by Vladimir Putin to his government to interfere with Hillary Clinton's bid for the presidency in 2016 is one such perspective. Here, I follow emerging work that suggests the digital campaign was *subversive* in nature.[28] Such a perspective makes great sense, and yet also conceptually offers a challenge to IR scholars that presume that major cyber operations are generally employed in support of efforts to directly contest the power of a foreign state. By contrast with, the nature of subversion is of normative disempowerment in aid of material goals – of an effort to diminish the sources of a competitor's power so as to alter the field of competition upon which war, coercion, and diplomacy will take place in the future.[29] In democratic societies, of course, that power in part emerges from the legitimacy of the status quo, an intangible fact of free societies with tangible contributors thereto that political warfare campaigns can target. Thinking about political warfare in this way is significant because a bargaining perspective on power in interstate relations does not make space for belligerent action that is not intended to address prevailing power dynamics. Given that the question addressed herein is of the potential for escalation from cyber operations employed to support grey zone activities, however, the relevant perspective is inevitably that of target vulnerabilities. What pressure points and vulnerabilities dictate the utility of cyber operations and, subsequently, the shape of potential escalation?

[27] Stephen Haggard and Jon R. Lindsay, 'North Korea and the Sony hack: Exporting instability through cyberspace', *AsiaPacific Issues*, 117 (2015); Whyte, 'Dissecting the digital world'.

[28] Henry Farrell and Bruce Schneier, 'Common-Knowledge Attacks on Democracy', Berkman Klein Center Research Publication 2018–7 (2018); Henrik Breitenbauch and Niels Byrjalsen, 'Subversion, statecraft and liberal democracy', *Survival*, 61:4 (2019), pp. 31–41; Samuel Zilincik, Michael Myklin, and Petr Kovanda, 'Cyber power and control: A perspective from strategic theory', *Journal of Cyber Policy*, 4:2 (2019), pp. 290–301.

[29] See Joseph Pierce and Olivia R. Williams, 'Against power? Distinguishing between acquisitive resistance and subversion', *Geografiska Annaler: Series B, Human Geography*, 98:3 (2016), pp. 171–88. Subversion is a concept that has received some significant focus in the past two years by scholars seeking to better understand how both state and non-state actors have employed cyber instruments in their efforts to spread or gain influence. Subversion differs from coercion insofar as it is a strategy that does not aim to degrade power in direct relational power terms, but rather by the challenging and derailing of authority. Subversion involves disruption of process and the normative degradation of symbols of the status quo so as to *dis*empower a foe and create opportunities for more conventional forms of contestation. For more work on subversion, see Audrey Kahin and George McTurnan Kahin, *Subversion as Foreign Policy: The Secret Eisenhower and Dulles Debacle in Indonesia* (Seattle: University of Washington Press, 1997); Laurence W. Beilenson, *Power through Subversion* (Washington, DC: Public Affairs Press, 1972); William Rosenau, *Subversion and Insurgency*, Vol. II (Santa Monica: RAND Corporation, 2007); Christopher Whyte, 'Out of the Shadows: Subversion and Counterculture in the Digital Age' (PhD dissertation, George Mason University, Fairfax, VA, 2017); Breitenbauch and Byrjalsen, 'Subversion, statecraft and liberal democracy'; and Zilincik, Myklin, and Kovanda, 'Cyber power and control'.

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

Democracies are information systems.[30] Thus, thinking about cyber operations conducted to support the disruption of democracies should consider more than just the outcomes of individual operations; it should consider how such national information systems work and might be disrupted. Here, the informational function of democracies lies with a moderating set of effects that emerge from broad-scoped discourse and sociopolitical contestation.[31] To be clear, democratic discourse and process does not lead public opinion and resultant policy towards truth or even fact. It does, however, centre and moderate perspectives on significant social, political, and economic issues via a complex updating process.[32]

As with all information systems, democracies depend on a series of mechanisms for assuring the proper condition of information as it is handled and transmitted. In particular, for the output of democratic discourse to be one of moderation, such conditions include an ability to reasonably determine the origination of information, the capacity to maintain credible discourse, high quality of available information, and a general belief that participation is fair and allowed. The origination criterion is clear – if discursive outcomes are to be reflective of the interests of democratic stakeholders, then citizens must be reasonably able to determine the source of information as it is handled and assessed in various forms of discourse. The credibility criterion leads on from the implication that commonplace misattribution of information – whether presented as reporting or opinion – fundamentally disenfranchises democratic process. Simply put, for democratic processes to function, there must be a continuing trust that discourse *is* discourse rather than simply a façade of some kind. This credibility requirement manifests in a secondary fashion insofar as citizens in a democracy must genuinely believe that no speech is prohibited, lest discourse be deemed artificial. And, finally, there must exist a reasonably broad base from which information is drawn and interpreted such that diverse opinions and investigative outcomes might be considered.

Deconstructing the function of democracies as information systems in this ways suggests an attack surface constituted of related mechanisms for information assurance that we might then use to better understand the threat of cyber-enabled political warfare. And indeed, such an assessment suggests that cyber operations play a relatively limited role in producing disruptive effects in this vein. Rather, as the case presented below suggests, the prospects for disruption of democratic process in the case of Russian efforts from at least 2014 onwards appear to have been principally shaped by: (1) the manner in which social media and social information aggregation platforms are minimally regulated and prone to manipulation based on understanding of underlying algorithmic design; (2) increased access to private information that might be employed to great disruptive effect; (3) enhanced opportunities for masking attribution of various online activities, including but not exclusively those linked to breaches; (4) a risk aversion among stakeholders emerging from the unprecedented shape of the threat; (5) a willingness to embrace foreign interference among some domestic actors; and (6) a willingness on the part of Russia to take action in an as yet unrealised area of contestation.

The utility of cyber operations for political warfare

Before considering the case of Russian interference in the United States through 2016 and assessing escalatory prospects, it is necessary to parse apart what about recent cyber-enabled political

[30] This is a common assertion of policy theory on democratic process, particularly focused on the marketplace of ideas. See, among others, Kaufmann, 'Threat inflation and the failure of the marketplace of ideas' and Thrall, 'A bear in the woods?'. For a recent assessment of democratic vulnerability to 'hacking' that makes the same overarching assertion, see Farrell and Schneier, 'Common-knowledge attacks on democracy'.

[31] Maxwell E. McCombs and Donald L. Shaw, 'The evolution of agenda-setting research: Twenty-five years in the marketplace of ideas', *Journal of Communication*, 43:2 (1993), pp. 58–67.

[32] See also John M. Owen, 'How liberalism produces democratic peace', *International Security*, 19:2 (1994), pp. 87–125; Layne, 'Kant or cant'; Snyder and Ballentine, 'Nationalism and the marketplace of ideas'; Ronald R. Krebs and Chaim Kaufmann, 'Selling the market short? The marketplace of ideas and the Iraq war', *International Security*, 29:4 (2005), pp. 196–207; and Tim Dunne, 'Liberalism, international terrorism, and democratic wars', *International Relations*, 23:1 (2009), pp. 107–14.

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

warfare campaigns has been truly 'cyber' (that is, directly linked to offensive computer network operations). As noted above, scholars interested in the dynamics of cyber conflict waged between state actors and their proxies have thus far vested much explanatory power in principles of technical function. Such an approach certainly enhances our ability to understand dynamics of engagement within the domain. But it also overlooks the reality that cyber operations are often employed as a derivative element of state strategy, with the only clear exception lying in the now-commonplace assumption that cyber instruments might be used to signal opponents within the context of strategic bargaining. Even there, however, some commentators are increasingly reluctant to adopt such traditional IR frameworks. For instance, recent work on the nature of engagement in the grey zone points out that much recent global conflict is puzzling exactly because it appears to be a suboptimal employment of state capabilities beyond what rationalist theories might predict.[33] As such, it is not appropriate to exclusively think about issues of cyber escalation or deterrence within the confines of rationalist assumptions.

The case outlined below of cyber-enabled political warfare prosecuted by the Russian Federation against the United States in recent years effectively demonstrates the shortcomings in thinking about escalation only in the context of cyber instruments deployed to coerce. Assuming that cyber operations were not employed obtain some particular compellent outcome but rather were intended to enhance an effort to disorient and disempower the target nation, it seems reasonable that analysis of the threat should begin with the viability of the campaign as a whole. As I argue above, this involves recognising the vulnerabilities that democratic processes face bound up in both the function of national information environments and time-and-place incentives for important stakeholders to deviate from their normal roles. Within this, cyber operations play only a limited role. Specifically, cyber operations stand to enhance political warfare campaigns in four ways.

First, there is an obvious advantage ceded to modern political warfare efforts in the ability to access, exfiltrate, and repurpose large volumes of private information for purposes of manipulation. Regardless of the extent of new abilities to manipulate target populations through careful navigation of the realities of new media platform design and management dynamics, the 2016 case makes it abundantly clear that possession of scandalous information stands to magnify the impact of meddling efforts. Specifically, such volumes make it easier to shape conspiratorial narratives and to release information in the most strategically beneficial manner (for example, in timed releases that prevent the orientation of discursive markets on one or few issues over long periods of time). A secondary part of such magnifying efforts was also seen, in the 2016 case, in the form of malware pushed via social media to vulnerable populations with the goal of artificially inflating the visibility of fake news content – that is, of infecting computers with malware that would then direct browsers to such content for purposes of 'upvoting' without the knowledge of the user. Without such cyber operations conducted to build arsenals of private information and to manipulate the media infrastructure used by targeted populations, political warfare efforts would be unable to as effectively blend factual (if private) information with altered, fake content. Resultantly, they would likely have a significantly harder time foiling those mechanisms of democratic systems that ensure abilities to attribute information.

Second, cyber operations divert both specialist and practitioner attention from other elements of political warfare, including attempts to manipulate virtual discourse, efforts made to subvert domestic stakeholders and the distribution of propaganda. Such distraction naturally takes several forms. On one hand, cyber operations provide a focal point for media coverage of apparent threats to national security and suggest – because of the relatively less abstract shape of networks, servers, and underlying infrastructure as targets – a somewhat different threat profile than what is actually being faced. Often this is because cyber operations are directed against symbolic targets with some social or political significance and because the median consumer of news media reporting is prone to

[33] Andres Gannon, Erik Gartzke, and Jon Lindsay, Working Paper (San Diego: University of California, 2018).

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

seizing and freezing on pre-existing notions of the shape of a threat.[34] With cyber-enabled political warfare as a relatively new threat type, this means a form of cognitive dissonance where conflict actions are perceived to indicate the occurrence of some pre-envisioned model of threat to national security, most likely that of digital insecurity centred on the oft-cited scenario of cyber doom perpetuated by pundits, scholars, and policymakers alike. On the other hand, cyberattacks can galvanise support for certain political narratives wherein proponents of transparency about one or all political actors feel that the violation of information privacy is justified by parochial political concerns.

Third, cyber operations necessarily consume the resources of institutions set up to counter digital insecurity. Where cyber intrusions occur as a component part of a campaign characterised by a hybrid deployment of various conventional and non-traditional instruments of statecraft, this can mean a mismatch between the roles assigned to different institutions and the challenge of countering a multifaceted threat. Particularly in democratic states and particularly where Western countries have increasingly stood up units with discrete cyberwarfare, there exists a natural challenge of coordination. Though surmountable, previously under-realised threat types – as the threat of cyber-enabled political warfare was before the past few years – demand rapid catch-up and virtually necessitate extended periods of adaptation.

Finally, and perhaps of greatest concern, cyber operations open space for the coercion of actors beneath the level of the state. Though it seems overly parsimonious to label such a multifaceted campaign as Russia's from 2014 onwards as 'cyber coercion', there is undoubtedly room for coercion bound up in the implications of network intrusions and subsequent manipulations of domestic discourse for certain non-state actors. The suggestion of sensitive information released to the public either by the leaking of other data or simply by the incidence of intrusions themselves presents as a motivator of conservatism – meaning of an inhibited inclination to adopt normal sociopolitical modes of operation – among domestic stakeholders. All else equal, evidence of compromise based on a foreign force's machinations seems likely to produce incentives to avoid antagonism on the part of domestic stakeholders concerned with similar compromise. I discuss this and other points introduced in this section below as I turn towards questions of prospective escalation.

The case of Russian interference in the 2016 American election

In part due to the work of government and media investigative efforts focused on the consequences thereof, few incidents provide such a clear view into the technical and operational unfolding of a cyber-enabled political warfare[35] campaign as do Russian efforts to manipulate social and political processes in the lead up to the American presidential election in 2016.[36] Though there is evidence of efforts to manipulate discourse in minor ways via thousands of fake accounts on social media sites as far back as 2014, the episode began in 2015 in the form

[34] Miguel Alberto N. Gomez, 'Sound the alarm! Updating beliefs and degradative cyber operations', *European Journal of International Security*, 4:2 (2019), pp. 190–208.

[35] It is important to recognise here that political warfare is a term often used interchangeably and with some degree of conceptual confusion along with terms like active measures, hybrid warfare, irregular warfare, and information warfare. Political warfare is seminally described as efforts that 'range from such covert actions as political alliances, economic measures, and white propaganda to such covert operations as clandestine support to friendly foreign elements, black psychological warfare and even encouragement of underground resistance in hostile states'; see Policy Planning Staff Memorandum, available at: {https://history.state.gov/historicaldocuments/frus1945-50Intel/d269}. For George Kennan, political warfare simply represented states' attempts to politick in conflictual fashion outside times of declared hostilities. Given the maintaining peace in such times is paramount, this means using all available mechanisms of approach beyond (though sometimes involving) military ones and avoiding actions that might be seen as provocative. For work on political warfare prosecuted by the Russian Federation, as is the main focus of this section, see, among others, Monaghan, 'The "war" in Russia's "hybrid warfare"'; Lanoszka, 'Russian hybrid warfare and extended deterrence in eastern Europe'; Renz, 'Russia and "hybrid warfare"'; and Chivvis, 'Understanding Russian hybrid warfare'.

[36] For one of the most up to date summaries of the episode to date, see Scott Shane and Mark Mazetti, 'The plot to subvert an election: Unraveling the Russia story so far', *New York Times* (2018).

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

of cyber instruments deployed to gain access to proprietary data.[37] Two different Russian government-affiliated threat actors, known to most in the cybersecurity industry and intelligence community for their malicious hacking work in the former Soviet sphere in recent years, launched attacks against several different domestic political and government targets off the back of a broad-scoped spearphishing campaign. The phishing effort involved thousands of emails that targeted staffers in an effort to gain personal access credentials. About one in forty such phishing emails were opened, leading to the collection of the private information of a range of political operatives and federal employees.[38]

Initial intrusions, which went unnoticed by many for as much as a year, allowed for access to unclassified systems across parts of the US executive branch. Likewise, phishing-enabled intrusions allowed both threat actors – known commonly, among other names, as CozyDuke and Fancy Bear – to acquire various private materials of political organisations. The most significant of these was the Democratic National Committee (though the Republican counterpart organisation was also attacked) and, relatedly, the campaign of then-presidential candidate Hillary Clinton.[39] APT28 and APT29 – the formal classifications of the threat actors in question – managed to steal emails from the political campaign alongside internal documents detailing various plans and designs held by the Democratic Party on upcoming political events.[40]

What followed was many months of timed information releases and dumps, some of which was coordinated via the Internet Research Agency (IRA), a Russian firm owned and operated by an oligarch with close ties to the Kremlin.[41] The IRA utilised thousands of bot accounts on social media platforms and online communities like those on the social news aggregator site Reddit to release snippets of information to the Western world over the course of the election season.[42] Elsewhere, WikiLeaks publicly sponsored the release of yet larger troves of data about the Democratic Party and Hillary Clinton, often timed – possibly strategically – so as to maximise the ensuing disruption to regular political processes.[43]

The IRA and a number of individuals affiliated with Kremlin-aligned oligarchs also acted thereafter to amplify the effect of information disruptions in 2015 and 2016. In particular, bots controlled by the IRA inserted themselves extensively into virtual conversations taking place on, primarily, Twitter and Reddit.[44] On Facebook, it was revealed in 2017 that Russian actors had purchased some thousands of dollars' worth of ad space and had published content targeted at demographic groups and localities where divisive rhetoric might have the most effect.[45] Recently, evidence has emerged that click fraud, a common feature of information manipulation on social community

[37] Valeriano, Jensen, and Maness, *Cyber Strategy*; M. Mazzetti and K. Benner, '12 Russian agents indicted in Mueller investigation', *The New York Times* (13 July 2018), available at: {https://www.nytimes.com/2018/07/13/us/politics/mueller-indictment-russian- intelligence-hacking.html}.

[38] Thomas Rid, 'Disinformation: A primer in Russian active measures and influence campaigns', *Hearings before the Select Committee on Intelligence, United States Senate, One Hundred Fifteenth Congress*, 30 (2017).

[39] Eric Lipton, David E. Sanger, and Scott Shane, 'The perfect weapon: How Russian cyberpower invaded the US', *The New York Times* (2016).

[40] David E. Sanger and Rick Corasaniti, 'DNC says Russian hackers penetrated its files, including dossier on Donald Trump', *The New York Times* (2016).

[41] Johan Farkas and Marco Bastos, 'IRA Propaganda on Twitter: Stoking Antagonism and Tweeting Local News', Proceedings of the 9th International Conference on Social Media and Society (2018), pp. 281–5; Brandon C. Boatwright, Darren L. Linvill, and Patrick L. Warren, 'Troll Factories: The Internet Research Agency and State-Sponsored Agenda Building', Resource Centre on Media Freedom in Europe (2018).

[42] Farkas and Bastos, 'IRA Propaganda on Twitter'; James P. Farwell, 'Countering Russian meddling in US political processes', *Parameters*, 48:1 (2018), pp. 37–47. Also see Christopher Whyte, 'Of commissars, cults and conspiratorial communities: The role of countercultural spaces in "democracy hacking" campaigns', *First Monday*, 25:4 (2020).

[43] Raphael Satter, Jeff Donn, and Chad Day, 'Inside story: How Russians hacked the Democrats' emails', *APNews. com* (2017).

[44] Darren L. Linvill and Patrick L. Warren, 'Troll factories: Manufacturing specialized disinformation on Twitter', *Political Communication* (2020), pp. 1–21.

[45] Dan Keating, Kevin Schaul, and Leslie Shapiro, 'The Facebook ads Russians targeted at different groups', *Washington Post* (2017).

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

sites and in ad monetisation schemes, was used to promote these topics above others viewers might otherwise see. As one newspaper estimates, a minimum of 23 million Americans were exposed to this content over almost a two-year period; this number may have been much higher.[46]

Russia's attempt to meddle in the political machinations of the American democracy was broad-scoped and sophisticated, though the redundant nature of actions taken by multiple known threat groups does suggest a lack of coordination in the early months of the effort. Nevertheless, the campaign was virtually unprecedented in Western experience insofar as cyber operations were employed in order to help scale and optimise opportunities for widespread disruption on a range of fronts. Notionally, the campaign's objectives seem to have been simple. The intelligence community in 2017 assessed that the point of Russian efforts was to 'undermine public faith in the US democratic process, denigrate Secretary [Hillary] Clinton, and harm her electability and potential presidency'.[47] The motivation for the years-long attack that intensified as electoral happenings ramped up in early 2016 appears to have in large part been the personal animosity felt by Russian leader Vladimir Putin against Secretary Clinton for the endorsement of the Panama Papers documents and for encouraging anti-Putin government protests in 2011–12.

And yet, for any discussion of the utility of cyber operations among other elements of Russia's interference toolkit it is not simply enough to acknowledge such appraisals of initial intent. If scholars are to understand where the potential for cyber escalation – or for that matter the prospects for deterrence or dynamics of compellence – lies with any given cyber-enabled incident, it is first necessary to nest knowledge about the uses of cyber instruments in an understanding of the strategic logic of one or other mode of contestation. Here, that means thinking about Russian interference in the context of the attack surface of democratic political systems outlined above.

Disruption of the discursive functions of democratic process rests increasingly on vulnerabilities inherent in an emergent information environment where the gatekeepers of information are algorithms and technology firms, rather than traditional news media organisations. Twitter, Facebook and other platforms are increasingly primary means of consuming information globally and Russia's effort to meddle in social and political conversations took full advantage of the fact that such firms are relatively less accountable to their customers than media outlets are to the public interest. Operators linked to the IRA not only purchased ad space on Facebook to promote anti-factual narratives; they specifically targeted content using marketer-friendly tools provided by the firm.[48] This allowed for targeting dispersion of propaganda via the specification of keywords ('Dixie', 'Black Lives Matter', etc.), region, and political leanings. Likewise, bot accounts on Twitter used sentiment tags and the retweeting function to align themselves with distinct virtual communities.[49] In both cases, platform functions subsequently did little to control for such targeting messaging, instead suggesting further content and connections based on such targeted circulation.

Likewise, though access to greater volumes of private information for weaponisation – discussed further below – is a major boon to political warfare efforts today, by far the more significant element of the threat lies with enhanced opportunities to inject scandalous, modified, or falsified content into regular democratic discourse. Foreign operators benefited during the 2016 election cycle from attribution difficulties on several fronts. For the average citizen, determining with some certainty the source and subsequent value of information was difficult because the IRA relied on a diverse portfolio of delivery methods. These might be organised into two categories. First, Russian operative constructed a three-pronged set of mechanisms for delivering

[46] Sarah Posner, 'What Facebook can tell us about Russian sabotage of our election' (2017), available at: {https://www.washingtonpost.com/blogs/plum-line/wp/2017/09/27/what-facebook-can-tell-us-about-russian-sabotage-of-our-election/}.

[47] Valeriano, Benjamin M. Jensen, and Ryan C. Maness, *Cyber Strategy*.

[48] Karoun Demirjian, 'Republican lawmakers move to restrain President-Elect Trump on Russia', *The Washington Post* (16 November 2016).

[49] Michael Katina, 'Bots trending now: Disinformation and calculated manipulation of the masses', *IEEE Technology and Society Magazine*, 36:2 (2017), pp. 6–11; Tobias R. Keller and Ulrike Klinger, 'Social bots in election campaigns: Theoretical, empirical, and methodological implications', *Political Communication*, 36:1 (2019), pp. 171–89.

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

stolen information to the American information environment via WikiLeaks, DCLeaks.com, and a hacker persona, 'Guccifer 2.0').[50] This diversification was clearly designed to amplify existing attribution challenges for the average engaged citizen on the topic of leaked information. Second, a series of efforts were made to shape conspiratorial communities on platforms like Reddit, 9gag, and 4chan.[51] The purpose of such communities appears to have been to establish the existence of conspiratorial narratives that would impact upon national discourse – most often laterally via fringe media outlets that would recirculate a limited volume of such perspectives to conservative commentators – via processes distinct from those surrounding leaked information. This two-pronged effort made use of fake content and provided a foundation from which Russian operators might enable the leak of information and sensational reporting for purposes of limiting time in which citizens might be able to consider data and update preferences.[52]

Finally, the particular format of the cyber-enabled threat faced by the United States up to and beyond 2016 emerges in large part from the unprecedented nature of conditions faced by domestic stakeholders. Because of the context of a presidential election that did not involve an incumbent candidate, media outlets found themselves particularly prone to recirculating candidate rhetoric. Coupled with the widespread politicisation of the race into the early months of 2016 and the embracement of numerous extreme positions by then-candidate Donald Trump, this amounted to a broad-scoped muting of the watchdog function of such entities. This happened not so much because criticism of important political candidates was not commonplace, but rather because the basis of that criticism was a news cycle increasingly driven by the tempo and content of political elements favoured (though not principally driven by) and tacitly supported by Russia's interference effort. Likewise, other important domestic stakeholders embraced conservative positions with regards their traditional corrective roles.[53] Where expert opinion found less-than-solid footing in the tumult of a fast-moving issues landscape, Republican officials and legislators found themselves uncertain about their responsibility to speak given the success of their party's adopted platform. And the executive, Obama, was reluctant to intercede in any significant way for fear of compromising the integrity of the domestic political process and uncertain about the value of steps that might be taken to deter further foreign action.[54]

Cyber operations, other conflict processes, and escalation

Though much emphasis is placed on the unique characteristics of the fifth domain by scholars interested in unpacking potential dynamics of escalation, it cannot be overlooked that cyber operations are augmentative more than they are strategically useful in their own right. With political warfare, though the uniqueness of the threat associated with Russia's efforts came from a combination of factors linked with the development of new informational conditions in Western democracies, cyber operations enable an expansion of operational scope and had second-order effects on how both the public and domestic stakeholders responded to the threat. In this section I argue that three primary escalatory effects are visible with the 2016 case and several others seem relevant. These effects must be taken into account when thinking about the potential for either cyber or conventional escalation emerging from such uses of digital instruments in the future.

[50] Satter, Donn, and Day, 'Inside story'.

[51] Niko Heikkilä, 'Online antagonism of the alt-right in the 2016 election', *European Journal of American Studies*, 12:12–2 (2017).

[52] Kathleen Hall Jamieson, *Cyberwar: How Russian Hackers and Trolls Helped Elect a President: What We Don't, Can't, and Do Know* (Oxford: Oxford University Press, 2018).

[53] See, for instance Sanger and Corasaniti, 'DNC Says Russian Hackers Penetrated Its Files', with the description of the Obama administration's paralysis given uncertain circumstances determined by both unprecedented assault and the context of domestic political processes unfolding.

[54] Ibid.

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

Cumulative effects and the subjectivity of hybrid threats

Perhaps the most obvious consideration emerging from the case overview above is the idea that effects brought about by state action across multiple domains of conflict pose unique escalation risks. Among the sources of escalation risk discussed in this section, this is the one that has received some reasonable measure of attention from IR scholars in the form of initial examinations of cyber conflict and prospects for cross-domain deterrence.[55] Indeed, the notion of multi-domain complexity causing problems for actors interested in signalling about escalation thresholds effectively is far from new. Nevertheless, it has to be the starting point here because of the degree to which the second-order effects discussed below add dimensionality to the problem of complexity and cumulative effects already assessed by scholars.[56]

Much recent scholarship has focused on cross-domain deterrence and the potential for conventional escalation thresholds to be breached by interstate engagement across other domains of contestation. Though there is much nuance to be considered in such work, the main concern is a relatively simple one to understand. As Figure 1 illustrates, the distance between status quo conditions and the threshold of armed interstate conventional or even nuclear conflict might be dramatically lessened due to the breaching of softer thresholds for escalation in other domains of engagement. In the example in Figure 1, a low-level escalation threshold might be crossed as unconventional conflict over a minor maritime territorial dispute enters a new phase. Part and parcel of this escalation may include retaliation against proxy elements of opponents' toolkit of contestation options, such as the interdiction of merchant ships tasked with harassing routine military operations. Over several stages, harassment in cyberspace and further escalation in the way that proxy and space-based communications assets are employed to frustrate one or both sides may open up new opportunities for more and more disruptive retaliation. At each threshold, the sides involve must choose whether or not to escalate the situation or maintain the status quo. However, the concern about cross-domain conflict dynamics is quite simply that no two thresholds are created equal in terms of the cost-benefit calculus associated with retaliation. Likewise, even though decision-makers should ideally be aware of the multifaceted nature of the challenge before them, attribution difficulties and issues with determining intention that is variable across domains presents information challenges therefor. And heightened complexity inevitably implies increased difficulty in overcoming a range of cognitive challenges associated with clarifying the significance of certain acts.

Herein lies a major source of potential escalation risk from the prosecution of cyber operations in support of political warfare efforts. Political warfare augmented by cyber operations is a cross-domain threat that carries with it the basic challenge of multifaceted efforts states undertake to contest and subvert the power of others. As such, decisions made at escalation thresholds unique to engagement in cyberspace may be shaped by the broader context of political warfare's effects. Though cyber belligerence may be limited or appear relatively unremarkable when considered in domain-specific isolation, the relative significance and perceived success of a broader political warfare effort will likely weigh in the balance and impact the decision to retaliate or not. Indeed, recent changes to national cyber doctrine in the United States reflect just this lateral impact in the assertion that the military should increasingly 'defend forward' to set new thresholds for restraining future conflict off the back of holistic analysis of the impact of foreign cyber operations on American national security.

[55] For instance, Paul M. Nakasone and Charlie Lewis, 'Cyberspace in multi-domain battle', *The Cyber Defense Review*, 2:1 (2017), pp. 15–26; Erik Gartzke and Jon R. Lindsay (eds), *Cross-domain Deterrence: Strategy in an Era of Complexity* (Oxford: Oxford University Press, 2019); Jacquelyn Schneider, 'Cyber and Cross Domain Deterrence: Deterring Within and From Cyberspace' (Naval War College, 2019).

[56] Chart inspired by that found in King Mallory, *New Challenges in Cross-Domain Deterrence* (Santa Monica: RAND Corporation, 2018).

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

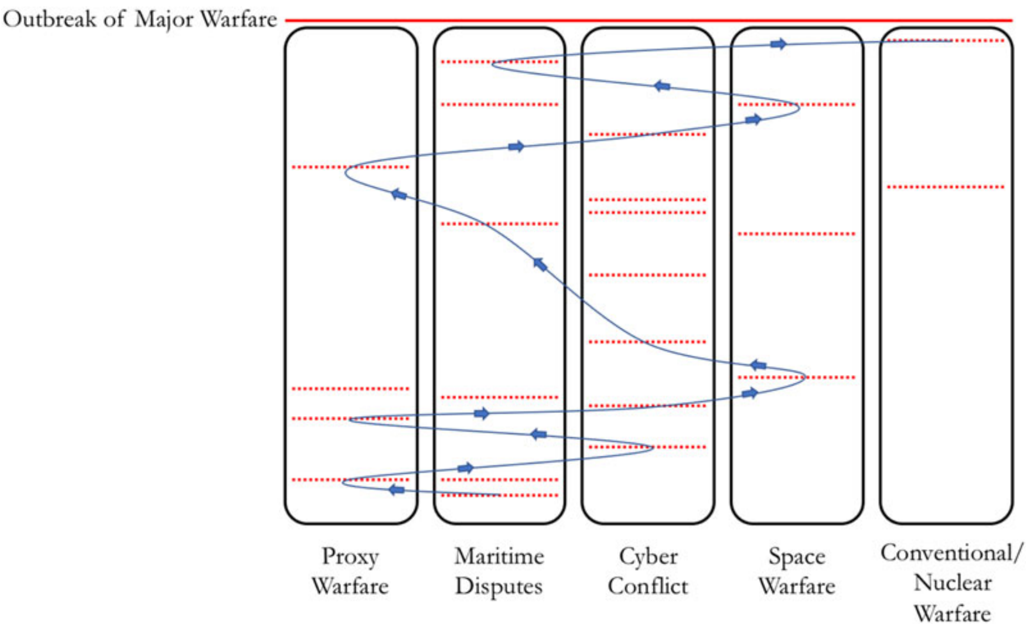

Figure 1. Escalation across multiple domains.

Perhaps more importantly, escalation dynamics seem not only likely to emerge from subjective assessments of hybrid threats with cyber characteristics, they have clearly done so already. The case of Russian interference in 2016 reflects an ideal-type circumstance for escalation along such lines. After all, particularly where the fullest extent of the connection between cyber intrusions and broader information operations is either unrealised or deemed to fall beyond the scope of the mission of core defence entities, limited response might prompt opponents to intensify the broader campaign in the belief that a threshold has not yet been reached. Conversely, early stage retaliation designed to deter further cyber aggression without a parallel effort to explicitly communicate the intention that the punishing action emerges from attribution of the broader campaign may be received ambiguously and have limited effect on other pillars of the interference effort. And such calculations naturally face further challenge insofar as the value proposition of the threat being made and interpreted might differ – perhaps even radically – between actors. Where the target is some element of democratic functionality, for instance, an authoritarian attacker may be more likely to discount non-violent interference as a cheap means of either signalling or attempting to subvert the bases of foreign power. For the democratic defender, however, the question may be much more substantially existential. This might particularly be the case if distributed effects have laid bare excessive vulnerability to foreign meddling on some fronts, thus incentivising a rapid or even panicked response wherein decision-makers are traditionally prone to being influenced by crisis heuristics and incentives to compress information processing to fit within a response schedule.

In the case of interference in 2016, perhaps the most substantial domestic shift in the practice of cyber conflict came in the form of a dramatic shift in strategy for the US military – promulgated in 2018 – from a posture of deterrence by denial to one that more closely reflects the precepts of deterrence by punishment. In this case, the strategy emphasises persistent engagement in the domain and the need to increasingly launch operations that consistently provoke enemies beyond America's own networks. Numerous commentators linked with development of the strategy note the significance of Russian political warfare in 2016 as a realisation moment prompting

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

the development of the new, more assertive strategy, despite the ironic fact that the new posture only addresses the cyber conflict elements of the threat.[57] Indeed, reference to the episode as a motivation of the new posture is made in the official strategy document itself just ten sentences in. This strategic shift resulted in at least one known retaliatory counterstrike against Russian assets linked with political warfare efforts in 2018.

Tying in the public

The nature of the targets of cyber-enabled political warfare as often being democratic states additionally impacts upon escalation dynamics and the pressures of appropriate decision-making in such cases. Subjectivity is not only a concern within the relatively dyadic game played by states interested in manageable contestation; it also manifests in the two level nature of foreign policy decision-making within any state that links the public voice to political outcomes. In particular, it is likely that the calculus of escalation versus maintaining the status quo might be impacted by such campaigns in one of two ways.

First, leaders may either choose or be forced by circumstance to incur audience costs associated with interference efforts. On the former possibility, domestic conditions may incentivise a leader to distract public attention from ongoing political or social turmoil by focusing on foreign threats. In such a situation, state leaders may either take advantage of a perceived threat from cyber-enabled information operations to stoke such domestic interest in international justice or construct one of their own trappings. The latter option, though it might seem unusual to suggest that leaders might regularly be a position to fabricate a foreign attack in all or some detail, is uniquely possible in the case of cyber-enabled political warfare efforts. Even where the public is exposed to some element of foreign interference, the full picture is – at least in incidents encountered to date – something that does not emerge for some months or years after the fact and following substantial investigation. As such, democratic leadership – whose membership traditionally holds enormous information advantages over other democratic stakeholders owing to their access to intelligence products and the use of the bully pulpit – are uniquely positioned to interpret the extent, scope, and intentions behind such attacks in order to fit their political needs. Conversely, of course, leaders may be forced into escalatory responses by organic public outrage. In the 2016 case, there is little doubt that interference has muted the responsiveness of US President Donald Trump to claims of Russian intransigence, as the politician has eagerly engaged the belligerent party instead of taking punitive actions that may be interpreted as admittance of foreign aid in the 2016 election.

It is also possible that leaders may find themselves constrained from pursuing otherwise rational courses of action to escalate for the purposes of deterring future attack by the prior existence or development of domestic division. As seen in the 2016 election season, the unique time-and-place circumstances of democratic process underway disincentivised action on the part of the Obama administration not only because of uncertainty about the value of available responses in an under-realised space of engagement, but also – indeed, primarily – due to concern that action taken would result in some measure of damage to credibility. The second-order effect of political circumstances was, in short, a muting of escalatory inclinations. It's easy to see how such dynamics might play out in the future to impact upon decisions made to escalate or do nothing based on either (1) coincidental internal circumstances or (2) manufactured turmoil. On the latter point, successful subversion of narratives and even specific democratic stakeholders by foreign meddlers may heighten the fear felt by the executive that retaliation will be seen as a political ploy. Conversely, such subversion might manifest as manufactured outrage and calls for escalation, at which point decision-makers must choose between responsiveness of popular concern and unpopular restraint for fear of authorising an unwise retaliation.

[57] Nina Kollars and Jacquelyn Schneider, 'Defending forward: The 2018 cyber strategy is here', *War on the Rocks* (20 September 2018), available at: {https://warontherocks.com/2018/09/defending-forward-the-2018-cyber-strategy-is-here/}; Jeff Kosseff, 'The Contours of "Defend Forward" under International Law', 11th International Conference on Cyber Conflict (CyCon) (IEEE, 2019), pp. 1–13.

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

And, of course, a leader who is motivated not to address interference, as Trump has been throughout his first term in office, will be incentivised to consistently cite division on the reality of a given threat. In the 2016 case, a substantial portion of all Americans have consistently answered polling that Russia either did not interfere or did but to no effect. Indeed, the figure has grown over time from approximately 35 per cent to more than half of the population. Such a dynamic not only enables domestic stakeholders who are tacitly coerced by cyber-enabled interference towards or away from escalatory behaviour; it has also led defence officials to double down on support for more assertive policy towards cyber conflict response, with several officials noting through 2019 that a lack of cohesive defensive efforts by civilian government and civil society actors ups the ante on the success of military efforts to punish foreign intransigence.

The other domestic concern

Domestic conditions further play into escalation dynamics emerging from the prosecution of cyber operations in support of broader political warfare efforts insofar as potential retaliation is not the sole domain of the state. The challenge of hack-back as a prevalent activity that might affect escalation risks manifests on several fronts. States may actually be inclined to encourage some forms of hack back by private domestic actors due to a belief that proxy action might be valuable for signalling while still maintaining distance necessary to avoid meaningful escalation with a foreign power. The risks here, however, are numerous. Not least of these is the risk that substate actors may not know what they're doing in returning fire, leading to confusion about the intentions and consequences of such actions. Second, in allowing for broad-scoped retaliation at the level of civil society actors, states cede much control over determining the proportionality of the retaliatory action. Third, allowing domestic actors to actively defend themselves in this manner further constrains the ability of states to signal effectively because it adds variables to the interpretive task set before foreign adversaries. Adversaries may be unclear as to whether any one signal is intended to reflect substate actor desires, state-set norms of engagement, or merely a knee jerk reaction not anchored in strategic intent. In this way, it is the norms surrounding hack-back that contribute to escalation risks where the minimal stickiness thereof limits the value of the action, particularly where the targets of responsive actions are themselves mere proxies of foreign powers. Fourth, public declaration of hack-back by one or more companies might further place demands on political leadership to respond to foreign aggression for reasons of national pride, even where the scope of the interference might not otherwise demand it. And finally, diffusion of the capability beyond the trappings of state power may simply encourage further cyber-enabled attempts at foreign interference, as the effects and context of substate actor response are perceived to have a interfering effect on the ability of states to effectively signal their intentions. In this scenario, escalation both in cyberspace and beyond becomes more likely as states seek a way to compensate for the loss of signalling control they suffer by allowing degrees of substate actor autonomy.

The coercion of domestic stakeholders

Discussed in part in the sections above, the possibility of successful subversion or coercion of domestic stakeholders looms large as a potential source of escalation risk related to cyber augmentation of political warfare campaigns. In a vacuum, it is easy to imagine the influence of compromised individuals or of narratives covertly inserted into domestic media environments having an effect on the trajectory of national discourse on a given issues. By far the more pointed version of the threat, however, lies in the coercion of critical stakeholders whose normal operation is required for national political systems to effectively function. Where cyber operations are leveraged to dramatically enhance disinformation and doxing efforts, they send a clear signal about the capabilities and intentions of foreign forces attempting to meddle in domestic processes. For social, economic and political operators whose voice might otherwise play a role in helping

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

moderate democratic discourse – by offering expert analysis or platform-specific opinions, even in the extreme – the example of such foreign capability suggests an imminent threat to reputation or position. This is dependent, of course, on the empirical basis for potential compromise. Nevertheless, the result is still likely to be a second-order effect wherein domestic stakeholders develop heuristics of response that favour conservative (that is, risk-averse) speech based on an altered understanding of the potential value of future actions. Such second-order cognitive effects, alongside the occasional direct coercion of domestic stakeholders, further complicates the calculus of escalation for state leaders by introducing the possibility of imperfect information about how domestic systems function over and above what exists in the circumstances of popular debate and the conduct of state forces – both government and otherwise – online.

The danger of Byzantine failures

In this final point is an overarching concern that should be articulated by itself. The spectre of disinformation, propaganda, and corruption-generating campaigns supported and enlarged by cyber operations suggests a set of conditions not only in which particular elements of democratic functionality could break down, but where such failure may be hard to perceive. In computer science, this class of failure is referred to as being about *Byzantine faults*. The name takes point from a game often played to describe the challenges involved in developing fault tolerant systems (that is, information systems that are resilient in the face of efforts to disrupt their various functions). The game involves generals at the head of several armies surrounding a city. In order to successfully take the city, they must coordinate their attack plans and attack all at once. Anything less than the commitment of all will end in failure. The problem before the generals is actually unsolvable, though not because of the inherent issues involved in developing systems to effectively secure communications against enemy interdiction. Rather, the problem is unsolvable because of the problem of other minds. Simply put, a general can never fully know the intention of his peers and so can never fully trust whatever system is put in place. No matter how secure, a fault may occur beyond the scope of the methods of communication involved. The trick for information technology designers then is to try and develop as robust a system as possible to make sure that a *Byzantine fault*, where failure might happen but in such a way that it would be hard to observe, won't happen.

With cyber-enabled political warfare, there is presently significant opportunity for failure of societal processes wherein attribution of the underlying causes is likely to be extremely difficult. Particularly where cyber operations enable second-order effects via information operations that themselves incur third and fourth order reverberations around the architecture of policymaking in democracies, scholars and policymakers must consider the potential for escalation that emerges from misguided understanding – even if it makes sense in context – of the intentions behind overlapping conflict actions. In short, that the attack surface of democracies is so diverse means that cyber effects deemed to be produced in support of broader subversive or coercive objectives should *always* be considered as distinct from attacks seemingly undertaken for narrower purpose.

Implications for scholarship and policy

Escalation risks emerging from the strategic use of cyber instruments are as diverse in shape as they are numerous. This article argues that, in thinking about the challenges inherent in governing strategic interactions that increasingly include digital aggression, scholars and practitioners alike must look beyond the characteristics of the methods involved and the operational implications thereof to consider lateral effects. When cyber instruments enhance other, more traditional forms of contestation – both non-coercive and coercive, conventional and otherwise – there is invariably the creation of second-order effects that themselves have the potential to impact upon the decision-making of state leaders.

The arguments made above carry an implicit criticism of much scholarly thinking about the dynamics of cyber conflict that adopts the framework of rationalist perspectives on conflict in

https://doi.org/10.1017/eis.2020.2 Published online by Cambridge University Press

international relations. Not all cyber antagonism is an attempt at coercion aimed at state decision-makers. In many instances, for instance, cyber instruments are used to support subversion, which constitutes a set of strategies aimed at disempowering the productive bases of actor capabilities rather than trying to actively degrade or neutralise their utility. Indeed, this conclusion is particularly meaningful because it suggests that national security practitioners should consider an alternative attack surface than the one traditionally articulated and addressed by national defence establishments. If subversive attacks on democratic process involve foreign attempts to degrade the normative strength of symbols of the status quo so as to reduce the credibility of discourse, then effective defence policy must take a more holistic approach in the future.

It is, of course, difficult to make such an argument about the scope of the bases of national power without considering implications for deterrence. At present, cyber conflict strategy in the West has shifted towards the belief that punitive conduct and constant engagement in the domain is required in order to shape favourable adversary behaviour. Prevailing thought holds that such conduct can avoid unwanted escalation via the use of reversible cyber effects and the tactical decision to retaliate only when operational gain can be scaled to strategically meaningful outcomes. The arguments made in the preceding sections do not walk back the underlying premises of such assumptions, per se. However, the reality that second-order effects might set the stage for escalation beyond the assumed tit-for-tat scope of interstate interactions does further inform circumstances as to when we might expect current strategy to have the desired effects. First, policymakers would do well to recognise that the use of cyber instruments for ends other than coercion or coercive signalling implies the existence of use cases that cannot readily be deterred by persistent shaping of the battlefield. Such usage and the low-intensity interference context of such usage adds further uncertainty to dynamics of perception such that adversaries (1) may be inclined to assume low likelihood of discovery or (2) may rely on an inability on the part of target nations to assign meaning thereto. Second, strategic planners would do well to better study the degree to which defensive efforts informed by the distinct conflict processes that cyber operations augment might enhance deterrent attempts to signal and set red lines. With the case of cyber-enabled political warfare, such a deterrent effort might include not only counteroffensive attacks against entities like the IRA, but also public-private partnership action on building verifiable social media systems.

Finally, for scholars, future work would also do well to consider the implications of such a point. In studying cyber conflict and other forms of conflict enabled by the Internet and related technologies, scholars should not shy away from applying existing frameworks and theories as intervening processes that help explain the relationship between digital actions and security outcomes. In the sections above, in particular, analysis presented suggests robust opportunity for better understanding of states' use of cyber instruments and the effects thereof in areas of IR research from audience costs and democratic peace theory to the study of power as a productive or structural element of world politics.

Acknowledgements. The author would like to thank the excellent comments and guidance of two anonymous reviewers, as well as the thoughtful remarks of colleagues linked with the Digital Issues Discussion Group for this project.

Christopher Whyte (PhD) is an Assistant Professor in the programme on Homeland Security and Emergency Preparedness at the L. Douglas Wilder School of Government and Public Affairs, Virginia Commonwealth University. His teaching and research interests encapsulate a range of topics pertaining to dynamics of global cyber conflict, crisis decision-making, information warfare, and emerging information technologies.

Cite this article: Whyte, C. 2020. Beyond tit-for-tat in cyberspace: Political warfare and lateral sources of escalation online. *European Journal of International Security* **5**, 195–214. https://doi.org/10.1017/eis.2020.2