The European Journal of International Law Vol. 31 no. 3

Beyond Naming and Shaming: Accusations and International Law in Cybersecurity

Martha Finnemore* and Duncan B. Hollis**

Abstract

Accusations of bad state behaviour in cyberspace are proliferating, yet this increase in naming has not obviously produced much shame. Accused states uniformly deny the accusation or decline to comment, without changing behaviour. For international lawyers, the problem is compounded by the absence of international law in these charges. States are not invoking international law when they complain of other states' behaviour, suggesting the law is weak – or worse, irrelevant – in holding states accountable for their cyber operations. In lieu of 'naming and shaming', we introduce and examine the broader concept of 'accusation' as a social, political and legal practice with diverse uses in cyberspace and beyond. Accusers must make strategic choices about how they frame their accusations, and we unpack various elements accusers may manipulate to their advantage. Accusations also have many purposes. They may seek to 'name and shame' an accused into conforming to certain behavioural expectations, but they may also aim at defensive or deterrent effects on both the accused and, crucially, on third parties. Particularly important, accusations may play a constitutive role, constructing new norms, including customary international law, within the international community. In short, accusations offer states and other stakeholders a menu of strategic options beyond those identified by the extant literature on naming and shaming.

* University Professor of Political Science and International Affairs, George Washington University and nonresident scholar, Carnegie Endowment for International Peace. Email: finnemor@gwu.edu. Professor Finnemore would like to thank Amoz Hor for his excellent research assistance and comments in the preparation of this paper.

** Laura H. Carnell Professor of Law, Temple University Beasley School of Law and nonresident scholar, Carnegie Endowment for International Peace. Email: duncan.hollis@temple.edu. In addition to his academic duties, Professor Hollis regularly consults with the Microsoft Corporation on its push to establish new rules and institutions for cyberspace. Professor Hollis would like to thank Corinne Zucker for her excellent research assistance. This article was prepared as part of a project on Protecting Civilian Institutions and Infrastructure from Cyber Operations: Designing International Law and Organizations, carried out by the Center for International Law and Governance at The Fletcher School of Law and Diplomacy, with financial support of the Microsoft Corporation and the Hitachi Center for Technology and International Affairs. All errors or omissions remain the authors' own.

1 Introduction

Once upon a time, hacking victims had little to say about the harms they suffered. Victims might never know they had been hacked and when they did, fears of reputational harm often kept them from disclosing it. In either case, cyberspace's technical architecture meant those responsible could often remain anonymous.[1] Victims had trouble discerning if their adversary was the proverbial basement-dwelling teenager, a shadowy cybercriminal organization or a nation state's intelligence or military services. As the number of states developing offensive cyber capabilities grew, conventional wisdom held that this 'attribution problem' posed serious – and perhaps insuperable – obstacles to enforcement by states of any rules in cyberspace.[2] The attribution problem complicated applying existing international legal regimes (e.g. international humanitarian law) whose operation depends on knowing a perpetrator's identity. More importantly, the attribution problem stymied efforts to clarify what international legal rules apply when cyber operations target civilians and their infrastructure outside of armed conflicts.[3]

Times have changed.[4] Over the last decade, 28 states – including China, Iran, North Korea, Russia, the United Kingdom and the United States – stand accused of conducting or supporting cyber operations with serious impacts on governments, peoples and resources.[5] These accusations 'naming' a state and its cyber operation(s) come

1 See Lindsay, 'Tipping the Scales: The Attribution Problem and the Feasibility of Deterrence Against Cyberattack', 1 *Journal of Cybersecurity* (2015) 53, at 54.

2 See, e.g., Ranger, 'US Intelligence: 30 Countries Building Cyber Attack Capabilities' (2017), available at www.zdnet.com/article/us-intelligence-30-countries-building-cyber-attack-capabilities/; Finnemore and Hollis, 'Constructing Norms for Global Cybersecurity', 110 *American Journal of International Law* (*AJIL*) (2016) 425, at 435–436.

3 We define a 'cyber operation' as the use of information and communication technologies (ICTs) to generate significant losses of confidentiality, integrity and/or access in computer systems. Our definition includes cases of cyber espionage – in which ICTs supplant more traditional spying tools – and more novel forms of cyberattack that degrade, disrupt or damage a computer system and (perhaps) the infrastructure it supports.

4 It is not, however, obvious *why* things changed. Certainly, technology evolved to allow some state (and non-state) actors greater visibility into cyber-attack origins. See, e.g., Edwards et al., 'Strategic aspects of cyberattack, attribution, and blame', 114 *Proceedings of the National Academy of Sciences* (2017) 2825; Davis II et al., *Stateless Attribution: Towards International Accountability for Cyberspace* (2017) at 2, available at www.rand.org/pubs/research_reports/RR2081.html. It is less clear why victims – or victim states – began making their accusations publicly.

5 See Council of Foreign Relations (CFR), *Cyber Operations Tracker* (2020), available at www.cfr.org/interactive/cyber operations ('CFR Tracker'); see also Center for Strategic & International Studies (CSIS), *Significant Cyber Incidents* (2020), available at www.csis.org/programs/technology-policy-program/significant-cyber-incidents; Davis II et al., *supra* note 4; Efrony and Shany, 'A Rule Book on the Shelf? Tallinn Manual 2.0 on Cyberoperations and Subsequent State Practice', 112 *AJIL* (2018) 583, at 594 (11 case studies of state-sponsored cyber operations). The total number of accusations, including those made more privately, is likely much higher.

from a variety of sources, including other states, demonstrating an increased willingness to 'name names'.[6]

All this increased naming, however, has not obviously produced a lot of shame. States accused of conducting or supporting cyber operations uniformly deny the accusation, point the finger at someone else or decline to comment.[7] They show few signs of changing behaviour.[8] The United States and China did reach an understanding in 2015 prohibiting commercial cyber espionage, following the US indictment of five People's Liberation Army (PLA) officers for such behaviour.[9] China's commitment, however, appears to have been more a response to domestic politics, and was – in any case – short-lived.[10] Thus, there is widespread scepticism about the capacity of so called 'naming and shaming' to curb unwanted behaviour in cyberspace.[11]

For international lawyers, the recent spate of accusations is troubling for a different reason – the accusers almost always fail to invoke international law. The international legal system operates, at least in part, via 'protests' – formal objections by states and other subjects 'against a conduct or a claim purported to be contrary to or unfounded in international law'.[12] In other contexts (e.g. human rights, the environment), 'naming and shaming' complaints often take the form of protests, explicitly tied to

[6] See, e.g., Starks, 'Trump Administration Increasingly Calls out Nations' Cyberattacks, But to What End?', *Politico* (6 June 2018), available at www.politico.com/newsletters/morning-cybersecurity/2018/06/06/trump-administration-increasingly-calls-out-nations-cyberattacks-but-to-what-end-243238; Roguski, 'Russian Cyber Attacks Against Georgia, Public Attributions and Sovereignty in Cyberspace' Just Security, 6 March 2020, available at www.justsecurity.org/69019/russian-cyber-attacks-against-georgia-public-attributions-and-sovereignty-in-cyberspace/ (citing 20 states accusing Russia of cyber operations against Georgia as evidence that 'more – especially European – States are willing to adopt public attributions').

[7] See, e.g., Roguski, *supra* note 6 (Russia dismisses accusation of its cyber operations against Georgia as 'unsubstantiated and politically motivated'); Davis II et al., *supra* note 4, at 2; Grove and Simmons, 'Russian Agency at Center of U.S. Hacking Indictment Has Long Operated in the Shadows', *Wall Street Journal* (14 July 2018).

[8] See, e.g., Goldsmith, 'Uncomfortable Questions in the Wake of Russia Indictment 2.0 and Trump's Press Conference With Putin', *Lawfare* (16 July 2018), available at www.lawfareblog.com/uncomfortable-questions-wake-russia-indictment-20-and-trumps-press-conference-putin.

[9] See Office of the Press Secretary, 'Fact Sheet: President Xi Jinping's State Visit to the United States' (25 September 2015), available at https://obamawhitehouse.archives.gov/the-press-office/2015/09/25/fact-sheet-president-xi-jinpings-state-visit-united-states; US Department of Justice, 'U.S. Charges Five Chinese Military Hackers for Cyber Espionage Against U.S. Corporations and a Labor Organization for Commercial Advantage' (19 May 2014), available at www.justice.gov/opa/pr/us-charges-five-chinese-military-hackers-cyber-espionage-against-us-corporations-and-labor ('2014 PLA Indictments').

[10] See, e.g., Laskai, 'A New Old Threat: Countering the Return of Chinese Industrial Cyber Espionage' (6 December 2018), available at www.cfr.org/report/threat-chinese-espionage; FireEye iSIGHT Intelligence, 'Red Line Drawn: China Recalculates Its Use of Cyber Espionage' (21 June 2016), available at www.fireeye.com/blog/threat-research/2016/06/red-line-drawn-china-espionage.html (citing President Xi's desire to reign in freelancing Chinese ministries as a factor in the reduction of Chinese commercial cyber espionage operations).

[11] See, e.g., Starks, *supra* note 6; Goldsmith, 'The DNC Hack and (the Lack of) Deterrence', *Lawfare* (9 October 2016), available at www.lawfareblog.com/dnc-hack-and-lack-deterrence.

[12] Eick, 'Protests', in R. Wolfrum (ed.), *Max Planck Encyclopedias of International Law* (2006), at ¶ 1, available at https://opil.ouplaw.com/view/10.1093/law:epil/9780199231690/law-9780199231690-e1460?rskey=GL9LnQ&result=1&prd=MPIL (by subscription).

violations of treaty terms or customary international law rules.[13] When it comes to state-sponsored cyber operations, however, accusers have studiously avoided invoking international law, let alone assessing if behaviour comports with its rules. Cyber operations are simply labelled as malicious, as irresponsible or as violations of 'international norms'.[14] Efrony and Shany highlight how 'remarkable' it is having 'so little in the practice of victim states to indicate that [their international legal rights] actually guide their conduct when confronted by cyber operations …'.[15]

[13] See, e.g., Human Rights Watch, *Egypt: Al-Sisi Should End Rights Abuses* (10 April 2018), available at www.hrw.org/news/2018/04/10/egypt-al-sisi-should-end-rights-abuses (disclosing irregularities in the Egyptian electoral process and calling on the government to 'comply with its international obligations under the International Covenant on Civil and Political Rights (ICCPR) and the African Charter on Human and Peoples' Rights'). Challenges to even robust norms, like those in human rights, are continuous, and widely studied in social science. See *infra* notes 18–19. Indeed, contestation around norms is often what strengthens them since contestation forces norm adherents to reaffirm those norms.

[14] See, e.g., US Department of the Treasury, 'Treasury Sanctions Russian Cyber Actors for Interference with the 2016 U.S. Elections and Malicious Cyber-Attacks' (15 March 2018), available at https://home.treasury.gov/news/press-releases/sm0312; Office of the Press Secretary, 'Statement by the President on Actions in Response to Russian Malicious Cyber Activity and Harassment' (29 December 2016), available at https://obamawhitehouse.archives.gov/the-press-office/2016/12/29/statement-president-actions-response-russian-malicious-cyber-activity; Kerry, 'Condemning Cyber-Attack by North Korea' (19 December 2014), available at https://2009–2017.state.gov/secretary/remarks/2014/12/235444.htm (describing 2014 Sony hack as a violation of 'international norms').

[15] Efrony and Shany, *supra* note 5, at 654; see also Roguski, *supra* note 6 (none of the states accusing Russia of cyber operations against Georgia in 2020 invoked international law). In one notable exception, in October 2018, five states (Australia, Canada, the Netherlands, New Zealand and the United Kingdom) coordinated accusations that the GRU – Russia's military intelligence arm – was responsible for a series of cyber operations, including those targeting the Organization for the Prohibition of Chemical Weapons (OPCW) and the World Anti-Doping Agency (WADA). The UK Foreign Secretary suggested that Russia had a 'desire to operate without regard to international law or established norms'. See Foreign & Commonwealth Office, 'UK Exposes Russian Cyber Attacks' (4 October 2018), available at www.gov.uk/government/news/uk-exposes-russian-cyber-attacks; National Cyber Security Center (NCSC), 'Reckless Campaign of Cyber Attacks by Russian Military Intelligence Service Exposed' (4 October 2018), available at www.ncsc.gov.uk/news/reckless-campaign-cyber-attacks-russian-military-intelligence-service-exposed. The Netherlands suggested that these activities 'undermine the international rule of law', while Canada's accusation incorporated both formulations. See Ministry of Defense, 'Netherlands Defence Intelligence and Security Service Disrupts Russian Cyber Operation Targeting OPCW' (4 October 2018), available at https://english.defensie.nl/latest/news/2018/10/04/netherlands-defence-intelligence-and-security-service-disrupts-russian-cyber operation-targeting-opcw; Global Affairs Canada, 'Canada Identifies Malicious Cyber-Activity by Russia' (4 October 2018), available at www.canada.ca/en/global-affairs/news/2018/10/canada-identifies-malicious-cyber-activity-by-russia.html (Russian activity demonstrates 'a disregard for international law and undermine[s] the rules-based international order'). None of these accusations delineated whether all of the GRU's alleged operations violated international law or if only some did; nor did they elaborate which international laws were violated. In contrast, Australia and New Zealand accused Russia of 'malicious cyber activity' without referencing international law at all. See, e.g., Government Communications Security Bureau (GCSB), 'Malicious Cyber Activity Attributed to Russia' (4 October 2018), available at www.gcsb.govt.nz/news/malicious-cyber-activity-attributed-to-russia/; Prime Minister of Australia, 'Attribution of a Pattern of Malicious Cyber Activity to Russia' (4 October 2018), available at www.pm.gov.au/media/attribution-pattern-malicious-cyber-activity-russia.

This reluctance to invoke international law might suggest that law is weak – or worse, irrelevant – in holding state actors accountable for their cyber operations.[16] However, focusing on international law's absence in accusations risks missing the larger effects that accusations may have on both compliance, and on the law itself. Certainly, social science suggests accusations *can* change an accused's behaviour.[17] International relations scholars have spent years studying the phenomenon of 'naming and shaming' states.[18] That research shows that public accusations of international law violations, most often in the human rights context, led certain accused states to conform with – or at least reduce their deviation from – international law. However, much of this scholarship (like the state practice it studies) assumes that shaming is the only effect of naming and that it occurs unproblematically.[19] It neglects both the other, independent effects that naming can have and the varied forms naming may take. Moreover, it presupposes a norm already in place to generate shame, rarely a warranted assumption for cybersecurity. Thus, we need a better analytic tool to investigate the diverse forms and potential effects that follow claims of state responsibility for cyber operations.

In lieu of 'naming and shaming' or 'protests', we develop the concept of 'accusation' to better capture the variation in the legal and political processes at work in these claims. Accusations make no prior assumptions about the goal of the 'accuser' (i.e. not always to 'shame') or about an accusation's effects (i.e. not always to enforce existing international law). Unpacking the concept, we understand accusations to be comprised of at least two of three discrete processes:

[16] If there are references to law, they usually involve domestic legal standards like the US indictments of foreign government agents for participating in various cyber operations. See, e.g., Department of Justice, 'Seven Iranians Working for Islamic Revolutionary Guard Corps-Affiliated Entities Charged for Conducting Coordinated Campaign of Cyber Attacks Against U.S. Financial Sector' (24 March 2016), available at www.justice.gov/opa/pr/seven-iranians-working-islamic-revolutionary-guard-corps-affiliated-entities-charged ('2016 Iranian Indictment'); Mazzetti and Benner, '12 Russian Agents Indicted in Mueller Investigation', *N.Y. Times* (*NYT*) (13 July 2018); 2014 PLA Indictments, *supra* note 9.

[17] See Pawson, 'Evidence and Policy and Naming and Shaming', 23 *Policy Studies* (2002) 211.

[18] See, e.g., Franklin, 'Shame on You: The Impact of Human Rights Criticism on Political Repression in Latin America', 52 *International Studies Quarterly* (*ISQ*) (2008) 187, at 204–207; Hafner-Burton, 'Sticks and Stones: Naming and Shaming the Human Rights Enforcement Problem', 62 *International Organization* (*Int'l Org.*) (2008) 689; Krain, 'J'accuse! Does Naming and Shaming Perpetrators Reduce the Severity of Genocides or Politicides?', 56 *ISQ* (2012) 574.

[19] Most scholarship in international relations conceives of naming and shaming as a unitary concept. See, e.g., Friman, 'Introduction: Unpacking the Mobilization of Shame', in H. Friman (ed.), *The Politics of Leverage in International Relations* (2015) 1, at 3 ('unpacking naming and shaming' by examining what 'exactly the *concept* means') (emphasis added). It is, moreover, almost always associated with altering the accused's behaviour. See, e.g., Beutz Land, 'Networked Activism', 22 *Harvard Human Rights Journal* (2009) 205, at 208 (defining 'naming and shaming' as 'the process of gathering information about a country's human rights record and publicizing that information in an effort to pressure or shame the government into changing its conduct'). See also *infra* note 20.

(i) *attribution* (the process of ascribing what happened to a particular actor or territory);
(ii) *exposure* (the process of disclosing what happened to third parties);
(iii) *condemnation* (the process of signalling disapproval of what happened).[20]

Accusations can include all three processes, as when the United States accused the Russian Federation of interference in its 2016 presidential election.[21] Other accusations may feature only two elements. Accusers can choose to attribute and condemn what happened without exposing it – i.e. making accusations via private or diplomatic channels. Or, accusers can expose and condemn what happened without disclosing (or even knowing) to whom it may be attributed.

Our elaboration of the concept of accusations in cybersecurity proceeds in five main parts. First, we develop the 'accusation' concept and survey its appearance in recent cyber practice (Section 2). Second, we identify different functions that accusations may serve for states and other actors based on cyber accusations made to date (Section 3). While accusations may aim at enforcement of the accuser's preferred rules and norms, accusations can also deter, aid defence and, importantly, contribute to the emergence of new norms and international law. Third, we describe an accusation's different components and show how these may be manipulated in various ways to achieve the desired goal(s) (Section 4). Fourth, we discuss external conditions that may influence an accusation's efficacy (Section 5). We conclude by examining implications of cybersecurity's accusation dynamics for international law (Section 6). Accusations are not merely vehicles for enforcing international law; they can also serve as building blocks for its creation. Accusations – and responses to them – may comprise the requisite evidence of state practice and/or *opinio juris* for the construction of customary international law. States and other actors need to understand the broader potential of their accusations, and shape their accusations – and responses – accordingly.

Appreciating the forms and functions of accusations enriches our understanding of both international relations and international law. Accusations are bread and butter in politics of many types, not just in cybersecurity. But international relations scholars have paid little attention to how they are deployed, what effects they create and the

[20] Those who have studied naming and shaming to date envision it as either a unitary mechanism or a compilation of two processes. See, e.g., Koliev, '*The Politics of Leverage in International Relations: Name, Shame, And Sanction*. Edited by H. Richard Friman', 91 *International Affairs* (2015) 1168, at 1169 (praising the volume for 'its conceptual distinction between public exposure (naming) and public condemnation (shaming)'). For his part, Friman defines naming and shaming as '[p]ublic exposure and condemnation'. See Friman, *supra* note 19, at 5, 203. As discussed in Section 4, however, we do not view accusations as requiring exposure and believe attribution is a distinct – albeit optional – component of accusations.

[21] See DHS Press Office, 'Joint Statement From the Department of Homeland Security and Office of the Director of National Intelligence on Election Security' (7 October 2016), available at www.dhs.gov/news/2016/10/07/joint-statement-department-homeland-security-and-office-director-national; Intelligence Community Assessment (ICA), 'Assessing Russian Activities and Intentions in Recent US Elections', Doc. ICA 2017-01D (6 January 2017), available at www.intelligence.senate.gov/sites/default/files/documents/ICA_2017_01.pdf.

strategic choices entailed in their creation and use. For international lawyers, the cybersecurity context provides a valuable case-study of how international law may be constituted in the shadows. State silence on international law in the cyber context does not signal the law's desuetude. Accusations may play a crucial role, not just in enforcing international law, but also in constituting it. This more detailed understanding of how accusations can be constructed and deployed to various ends will help states and other stakeholders better manage cybersecurity challenges and the development of law on this issue.

2 Accusations: Defining the Concept in Theory and Practice

Accusations are a regular feature of all social interactions. A parent may accuse her child of causing a sibling to cry; a non-governmental organization (NGO) may accuse a company of using child labour; shareholders may accuse CEOs of mismanagement. In the context of this article on global cybersecurity, we define an accusation as the process by which one or more actors allege that a state bears responsibility for a cyber operation.[22]

Our accusation concept captures more behaviour than fits under the 'naming and shaming' umbrella and avoids some of its problematic presumptions.[23] For example, 'naming' presupposes that charges are public, but that is not required for accusations.[24] Accusers can, and often do, communicate their charges privately to an accused or make accusations in closed settings. The nature and scope of 'shaming' also requires more specification. It might refer to the accuser's acts (e.g. 'X shamed Y into action'), but whether Y felt shame (or if shame is even possible in institutional actors like states) is ambiguous in the 'naming and shaming' construct. An accusation, in contrast, focuses our attention on behaviour by the accuser, alone. It makes no presumptions about effects upon the accused.

Notions of 'naming and shaming' also implicitly assume (and often conflate) both the normative virtue of the demanded action and the accuracy of the claims made against the 'shamed'. Our concept of accusations makes no such assumptions. Accusations may prove inaccurate or false under scrutiny, but scrutiny takes time and sometimes creates confusion that can be strategically useful. Perpetrators of cyber

[22] See *supra* note 3 (defining 'cyber operation'); H. Lin, 'Attribution of Malicious Cyber Incidents: From Soup to Nuts', *Columbia SIPA Journal of International Affairs* (2016). For a discussion of responsibility, see *infra* note 81.

[23] Our concept, unlike protests, does not limit itself to complaints expressed in legal terms. See Eick, *supra* note 12, at ¶ 10 ('A protest can only be made against a violation of international law or a conduct or claim that has no basis in international law. A deviation from comity alone, or other unfriendly acts, cannot give rise to a protest in the legal sense').

[24] Krain, *supra* note 18, at 575 (defining 'naming' as 'publiciz[ing] rights violations and their perpetrators'); Pawson, *supra* note 17, at 212 (characterizing naming and shaming as a strategy of '"public disclosure" to overcome recalcitrant behavior').

operations may attempt to 'false flag' their origins. Alternatively, victims might knowingly accuse an innocent party for various reasons, whether to cause the accused state reputational harm, to sow confusion or to avoid appearing incompetent before domestic audiences with respect to the capacity to identify a cyber operation's actual provenance.[25]

Given these issues, it is not surprising that global cybersecurity has avoided the 'naming-and-shaming' moniker. States and scholars have instead increasingly focused their attention on the concept of 'attribution'.[26] Unfortunately, that term has multiple distinct meanings, some technical, some legal and some political.[27] 'Technical' attribution may refer to identifying (i) the machine from which a cyber operation arises, (ii) the operator of that machine or (iii) the person or entity who directed the operator to act.[28] All three of these are distinct, one from another, and also distinct from legal or political attribution, which seeks to assign responsibility for ordering a cyberattack to its authors or to identify the location from which it originates.[29]

We agree that attribution is an important task for global cybersecurity. Many accusations include an aspect of attribution. Yet, the concepts are not synonymous. Accusations can occur without attribution (i.e. when accusers say, 'we do not know who did this, but it happened, and it was bad'). And where the concepts do overlap, accusations highlight additional issues for attention beyond where the malware originated or who launched it.

For this paper, we constrain our accusation definition to cases involving state or state-sponsored cyber operations for three reasons. First, accusations against non-state actors have existed since *The New York Times* accused Robert Morris of authoring the first computer virus.[30] Today, these accusations are ubiquitous, encompassing, for example, all cybercrime charges in every state. As a practical matter, assessing hacking accusations *in toto* is simply too unwieldy for an initial conceptual analysis such as

[25] On false flags, see US National Security Agency and UK National Cyber Security Centre, 'Cybersecurity Advisory: Turla Group Exploits Iranian APT to Expand Coverage of Victims' (21 October 2019), available at https://media.defense.gov/2019/Oct/18/2002197242/-1/-1/0/NSA_CSA_TURLA_20191021%20VER%203%20-%20COPY.PDF ('US–UK Turla Group Alert') (describing Russian attempts to disguise exploits as Iranian in origin).

[26] See, e.g., CFR Tracker, *supra* note 5; Hakimi, 'Introduction to the Symposium on Cyber Attribution', 113 *AJIL Unbound* (2019) 189; Eichensehr, 'The Law & Politics of Cyberattack Attribution', 67 *UCLA Law Review* (*UCLA L. Rev.*) (forthcoming 2020).

[27] Eichensehr, *supra* note 26, at 5–6.

[28] *Ibid.* Note that the first, and possibly the second, of these attribution forms may not involve 'naming' at all. Identifying the machine may reveal little about the names of the humans (or other actors) actually responsible for the cyber operation.

[29] *Ibid.*, at 1 ('Cyberattack attribution is the process of assigning responsibility for carrying out a cyber attack'); Banks, 'The Bumpy Road to a Meaningful International Law of Cyber Attribution', 113 *AJIL Unbound* (2019) 191, at 192 ('Attribution is defined as "identifying the agent responsible for the action"'); Clark and Landau, 'Untangling Attribution', 2 *Harvard National Security Journal* (2011) 531, at 531–532 (defining attribution as 'determining the identity or location of an attacker or an attacker's intermediary').

[30] See Markoff, 'Author of Computer "Virus" is Son of N.S.A. Expert on Data Security', *NYT* (5 November 1988).

this. Second, by focusing on accusations against *states* we work with the primary unit of analysis for both international law and international relations. While we believe our framework could be employed by researchers to investigate accusations by other international actors in other international contexts, we do not do so here.

Third, our concept captures a new – and expanding – behavioural phenomenon. In 2007, Estonia's Foreign Minister accused Russia of being behind directed denial of service (DDoS) attacks that significantly disrupted that country for three weeks.[31] Since then, researchers have catalogued an increasing and diverse practice of state-sponsored cyber operations. Accusations about these operations include allegations against individuals or 'threat groups' reportedly affiliated with a state as well as against states themselves.[32] Accusations may be made by non-governmental actors, including information and technology communication (ICT) companies, cybersecurity vendors or academic institutions.[33] Alternatively, states themselves may make accusations against other states. These can take different forms, including (i) criminal law indictments of individuals affiliated with a state, (ii) economic sanctions, (iii) technical warnings and (iv) press releases.[34] In some cases, states appear to rely on non-governmental proxies to make their accusations for them.[35] Most recently, states have begun to make collective accusations, issued contemporaneously or even jointly.[36] Taken together, accusations are a new and important aspect of the geopolitics of cybersecurity; they require attention from states and scholars alike.

[31] See Davis, 'Hackers Take Down the Most Wired Country in Europe', *Wired* (21 August 2007), available at www.wired.com/2007/08/ff-estonia/.

[32] See Mueller et al., 'Cyber Attribution', 4 *Cyber Defense Review* (2019) 107, at 112.

[33] See, e.g., Burt, 'Recent Cyberattacks Require Us All to Be Vigilant', *Microsoft on the Issues* (4 October 2019), available at https://blogs.microsoft.com/on-the-issues/2019/10/04/recent-cyberattacks-require-us-all-to-be-,vigilant/ (accusing a threat group linked to Iran of significant cyber activity); Marczak et al., 'Hide and Seek – Tracking NSO Group's Pegasus Spyware to Operations in 45 Countries', *Citizen Lab* (18 September 2018), available at https://citizenlab.ca/2018/09/hide-and-seek-tracking-nso-groups-pegasus-spyware-to-operations-in-45-countries/ (accusing six countries – Bahrain, Kazakhstan, Mexico, Morocco, Saudi Arabia and the United Arab Emirates – of using spyware to target civil society); Alperovitch, 'Bears in the Midst: Intrusion into the Democratic National Committee', *CrowdStrike Blog* (15 June 2016), available at www.crowdstrike.com/blog/bears-midst-intrusion-democratic-national-committee/ (accusing Russia of hacking the Democratic National Committee).

[34] See Eichensehr, *supra* note 26, at 10. For more on the practice of making accusations (and attributions) via indictments, see Hinck and Maurer, 'What's the Point of Charging Foreign State-Linked Hackers?', *Lawfare* (24 May 2019), available at www.lawfareblog.com/whats-point-charging-foreign-state-linked-hackers; Keitner, 'Attribution by Indictment', 113 *AJIL Unbound* (2019) 207.

[35] See *infra* note 95 and accompanying text.

[36] See, e.g., Roguski, *supra* note 6; Stubbs, Menn and Bing, 'Special Report: Inside the West's Failed Fight Against China's "Cloud Hopper" Hackers', *Reuters* (26 June 2019), available at www.reuters.com/article/us-china-cyber-cloudhopper-special-repor/special-report-inside-the-wests-failed-fight-against-chinas-cloud-hopper-hackers-idUSKCN1TR1DK (US links hackers, known as APT10, to China's Ministry of State Security, an attribution that led to statements by Germany, New Zealand, Canada, Britain, Australia and other allies backing the US accusations); Stilgherrian, 'Blaming Russia for NotPetya was Coordinated Diplomatic Action', *ZNet* (12 April 2018), available at www.zdnet.com/article/blaming-russia-for-notpetya-was-coordinated-diplomatic-action/; US–UK Turla Group Alert, *supra* note 25.

3 What Can Accusations Achieve?

What purpose do accusations serve? We focus here on four reasons an accuser may deploy an accusation: (i) enforcement; (ii) defence; (iii) deterrence; or (iv) constitution.[37] Some accusations may focus on achieving only one of these purposes; others may pursue multiple purposes sequentially or simultaneously. In every case, however, accusations are provocative, seeking to launch a broader chain of political, social, or legally significant events.

A *Enforcement*

Enforcement is the function most often associated with the 'naming and shaming' literature. Accusations often call out undesirable behaviour as a means to alter that behaviour in line with the accuser's behavioural expectations. The basic logic of such accusations is straightforward. Bad actors usually seek to hide their bad actions. Polluting firms would prefer we not know about their activities.[38] Companies engaged in questionable financial practices may not welcome public scrutiny.[39] Human rights-violating governments usually prefer to torture and 'disappear' their opponents in secret.[40] Public exposure or revelation of the bad behaviour ('naming') seeks to impose reputational damage and/or moral discomfort ('shaming') on the bad actor, thereby inducing a change in that behaviour.

The enforcement logic lies behind a number of accusations in the global cybersecurity context, especially those involving states as the accuser. It was the rationale behind President Obama's accusation that North Korea was responsible for the Sony Pictures hack and of subsequent US charges and sanctions against a named Pyongyang operative, Park Jin Hyok.[41] US indictments of specific Chinese and Iranian individuals

[37] This is not an exhaustive list. Accusations by private cybersecurity companies, for example, may serve an economic function. As Mandiant's financial success after first accusing China of cyber-espionage shows, credible accusations by cybersecurity companies may boost client sales or profitability. See Finkel, 'Mandiant Goes Viral After China Hacking Report', *Reuters* (22 February 2013), available at www.reuters.com/article/net-us-hackers-virus-china-mandiant/mandiant-goes-viral-after-china-hacking-report-idUSBRE91M02P20130223. Alternatively, as mentioned above, accusations might be deployed falsely (or otherwise) for their disruptive value – i.e. to sow confusion or create chaos.

[38] See, e.g., J. Hamilton, *Regulation Through Revelation: The Origin, Politics, and Impacts of the Toxics Release Inventory Program* (2005).

[39] van Erp, 'Naming Without Shaming: The Publication of Sanctions in the Dutch Financial Market', 5 *Regulation and Governance* (2011) 287.

[40] Murdie and Peksen, 'Women's rights INGO shaming and the government respect for women's rights', 10 *Review of International Organizations* (2015) 1; Murdie and Davis, 'Shaming and Blaming: Using Events Data to Assess the Impact of Human Rights INGOs', 56 *ISQ* (2012) 1; Hafner-Burton, *supra* note 18.

[41] Office of the Press Secretary, Statement, 'Imposing Additional Sanctions with Respect to North Korea' (2 January 2015), available at https://obamawhitehouse.archives.gov/the-press-office/2015/01/02/statement-press-secretary-executive-order-entitled-imposing-additional-s; Sullivan, 'Obama: North Korea Hack "Cyber-Vandalism," Not "Act of War"', *Washington Post* (*Wash. Post*) (21 December 2014); Nakashima and Barrett, 'U.S. Charges North Korean Operative in Conspiracy to Hack Sony Pictures, Banks', *Wash. Post* (6 September 2018), available at www.washingtonpost.com/world/national-security/justice-department-to-announce-hacking-charges-against-north-korean-operative-the-charge--stemming-from-the-2014-sony-pictures-case--is-the-first-against-a-pyongyang-spy/2018/09/06/f477bfb2-b1d0-11e8-9a6a-565d92a3585d_story.html.

affiliated with their respective governments had a similar purpose in the absence of agreed mechanisms to bring them before US courts.[42]

Accusations may thus seek to force compliance by a state merely by making the accusation. In other cases, accusations may pursue enforcement more indirectly. Accusers may, for example, issue accusations to 'persuade a set of third-party actors to generate support for sanctions', with those sanctions triggering the desired enforcement.[43] Accusations are also required to deploy domestic criminal penalties, including those against individuals acting at the direction of, or on behalf of, states who control them. Since first indicting five PLA officers, the United States has pursued indictments with increased frequency.[44] And, although most of the accused have escaped law enforcement, the United States did arrest a Chinese national in 2017 on charges of participating in hacks on the US Office of Personnel Management (OPM) detected in 2014 and 2015.[45] More recently, the United States succeeded in extraditing from Belgium Yanjun Xu, a deputy division director in China's main spy agency, the Ministry of State Security, for allegedly committing cyber espionage against US suppliers of commercial and military aircraft.[46]

Accusations can also be deployed to enforce international law via its two traditional vehicles for obtaining the cessation of wrongful behaviour: retorsion and counter-measures. Acts of retorsion are unfriendly – but lawful – acts (e.g. the expulsion of diplomats) designed to respond to an unlawful act.[47] Counter-measures are non-forceful acts that would otherwise be illegal, but which international law permits when conducted by a state in response to another state's prior wrongful act(s).[48] For a state to engage in either retorsion or counter-measures, however, requires some accusation articulating the requisite wrongful acts that form the basis for it to pursue the enforcement of its legal rights.[49]

[42] See, e.g., 2016 Iranian Indictment, *supra* note 16; 2014 PLA Indictments, *supra* note 9.

[43] Davis II et al., *supra* note 4, at 17.

[44] Cimpanu, 'DOJ Explains Recent Wave of Cyber-Espionage-Related Indictments', *ZDNet* (5 October 2018), available at www.zdnet.com/article/doj-explains-recent-wave-of-cyber-espionage-related-indictments/ (cataloguing post-2014 US indictments of three Chinese, nine Iranian, one North Korean and groups comprised of 12 and seven Russian hackers, all of whom were associated with their nation's governments).

[45] Menn, 'Chinese National Arrested in Los Angeles on U.S. Hacking Charge', *Reuters* (24 August 2017), available at www.reuters.com/article/us-usa-cyber-opm/chinese-national-arrested-in-los-angeles-on-u-s-hacking-charge-idUSKCN1B42RM.

[46] Benner, 'Chinese Officer Is Extradited to U.S. to Face Charges of Economic Espionage', *NYT* (10 October 2018).

[47] Crootof, 'International Cybertorts: Expanding State Accountability in Cyberspace', 103 *Cornell Law Review* (2018) 565, at 579.

[48] See International Law Commission (ILC), 'Draft Articles on the Responsibility of States for Internationally Wrongful Acts', Report of the International Law Commission on the Work of its 53rd session ('ASR'), UN Doc A/56/10 55, 23 April–1 June and 2 July–10 August 2001, Art. 22, at 75.

[49] *Ibid.*, Art. 43, at 119 ('An injured State which invokes the responsibility of another State shall give notice of its claim to that State'); *ibid.*, Art. 52, at 135 ('1. Before taking countermeasures, an injured State shall: (a) call upon the responsible State, in accordance with article 43, to fulfil its obligations …'). The lawfulness of counter-measures is measured in part by its proportionality to the originally wrongful act. *Ibid.*, Art. 51, at 134.

The Trump Administration has recently touted enforcement as the core of its 'naming and shaming' strategy. In describing the increasing number of US accusations of state-sponsored cyber operations, Jeanette Manfra, then the Department of Homeland Security's Assistant Secretary for Cybersecurity and Communications, made clear their purpose: 'The U.S. … wants to alter the behavior of nations that are carrying out attacks …. The broader policy purpose still remains [that] we need to be able to hold bad actors accountable.'[50]

B *Defence*

Manfra, however, also articulated a second function that accusations can serve: defence.[51] Accusations provide information on what happened that can have great utility to third parties. This is especially important for cybersecurity where an accusation 'may encourage victims or other vulnerable populations to bolster network defenses'.[52] Thus, a number of accusations regarding cybersecurity operations have included technical indicators of compromise (IOCs) to assist other potential victims in identifying and defending against the malware in question (or future manifestations of it). Accusations about the Trisis/TRITON malware – which could result in loss of life by disrupting emergency shutdown systems within industrial plants – focused on detailing the nature of the threat without identifying its specific authors.[53] Similar defence-oriented contents have accompanied other accusations, including those associated with Russia's 2016 electoral interference and the malware that targeted Ukraine's power grid in 2015.[54]

C *Deterrence*

Accusations may do more than assist third parties in defences; they may seek to deter potential perpetrators from ever engaging in the unwanted activity in the first place.[55] By exposing a state's cyber operations, accusers signal that others cannot engage

[50] Starks, *supra* note 6.

[51] *Ibid.* (Manfra 'said the move toward more direct and public attribution is about giving the private sector as much information as possible so it can safeguard their networks. That means being direct about who carried out the attack and announcing it publicly to reach the most people').

[52] Davis II et al., *supra* note 4, at 17.

[53] See, e.g., Johnson et al., 'Attackers Deploy New ICS Attack Framework "TRITON" and Cause Operational Disruption to Critical Infrastructure', *FireEye Blog* (14 December 2017), available at www.fireeye.com/blog/threat-research/2017/12/attackers-deploy-new-ics-attack-framework-triton.html; see also Bing, 'Trisis Has the Security World Spooked, Stumped and Searching for Answers', *Cyberscoop* (16 January 2018), available at www.cyberscoop.com/trisis-ics-malware-saudi-arabia/.

[54] See, e.g., National Cybersecurity and Communications Integration Center (NCCIC) and Federal Bureau of Investigation (FBI), 'Grizzly Steppe – Russian Malicious Cyber Activity', Ref. No. JAR-16-20296A (29 December 2016), available at www.us-cert.gov/sites/default/files/publications/JAR_16-20296A_GRIZZLY%20STEPPE-2016-1229.pdf ('Joint Analysis Report'); Cybersecurity & Infrastructure Security Agency (CISA), *ICS Alert: Cyber-Attack Against Ukrainian Critical Infrastructure* (25 February 2016), available at www.us-cert.gov/ics/alerts/IR-ALERT-H-16-056-01.

[55] Deterrence – stopping an actor from engaging in behaviour that has yet to occur – thus differs from enforcement, which involves material or social mechanisms of coercion to stop bad behaviour that has already occurred or is ongoing.

in similar behaviour without public attention. Cyber operations are often attractive precisely because perpetrating states think that they can be deployed anonymously – i.e. the operation will be undetected, the perpetrator can keep its own role unclear or perhaps it can foist blame onto another state or non-state party (a 'false flag operation').[56] Diminishing anonymity may change the cost–benefit calculus of states contemplating cyber operations; in some cases, it could deter them from acting at all. Thus, UK Foreign Secretary Jack Straw explained his government's motivations for accusations about cyber operations by Russia's military intelligence agency, the GRU, as follows: to 'expose and respond to the GRU's attempts to undermine international stability'.[57] Deterrence was also likely among the reasons that seven states – Australia, Canada, Denmark, Lithuania, New Zealand, the United Kingdom and the United States – accused the Russian Federation of launching the NotPetya ransomware.[58] Similarly, deterrence interests may explain accusations discrediting false flag cyber operations, including reports that Russia – not ISIS – conducted a cyberattack knocking TV5Monde off the air in France, and that Russia – not North Korea – disrupted the information infrastructure associated with the 2018 Winter Olympic Games.[59]

D *Constitution*

Finally, accusations may be constitutive of new norms and law, or new interpretations of their meanings. In many cases, an accusation 'sends a public message about correct and appropriate behavior'.[60] In the human rights context, accusations often invoke well-established legal norms of behaviour (e.g. prohibitions on torture or genocide; freedoms of expression or religion) against which the accused's behaviour is measured.[61] In such cases, the norm's existence is already widely acknowledged and the constitutive role of accusations lies in elaborating its meaning with respect to new circumstances or actors. A similar process could occur within cybersecurity whereby an

[56] See, e.g., US–UK Turla Group Alert, *supra* note 25.

[57] See Foreign & Commonwealth Office, *supra* note 15.

[58] See, e.g., Office of the Press Secretary, 'Statement from the Press Secretary' (15 February 2018), available at www.whitehouse.gov/briefings-statements/statement-press-secretary-25/; see also Stilgherrian, *supra* note 36. Ukraine also blamed Russia. Security Service of Ukraine, 'SBU Establishes Involvement of the RF Special Services into Petya: A Virus-Extorter Attack' (1 July 2017), available at https://ssu.gov.ua/en/news/1/category/2/view/3660#.eXBAf7Sa.dpbs. NotPetya was a ransomware attack designed to target Ukraine and significantly disrupted its hospitals, power companies, airports and central bank. But it also affected 64 other countries, and companies such as FedEx, Maersk and Merck sustained losses of hundreds of millions of dollars. See Forrest, 'NotPetya Ransomware Outbreak Cost Merck More Than $300M Per Quarter', *TechRepublic* (30 October 2017), available at www.techrepublic.com/article/notpetya-ransomware-outbreak-cost-merck-more-than-300m-per-quarter/.

[59] Frenkel, 'Experts Say Russians May Have Posed as ISIS to Hack French TV Channel', *Buzzfeed News* (9 June 2015), available at www.buzzfeednews.com/article/sheerafrenkel/experts-say-russians-may-have-posed-as-isis-to-hack-french-t; Nakashima, 'Russian Spies Hacked the Olympics and Tried to Make It Look Like North Korea Did It, U.S. Officials Say', *Wash. Post* (24 February 2018).

[60] van Erp, *supra* note 39.

[61] See, e.g., International Covenant on Civil and Political Rights (ICCPR), 999 UNTS 171, 16 December 1966; Convention on the Prevention and Punishment of the Crime of Genocide, 78 UNTS 277, 9 December 1948.

accusation references pre-existing norms, offering an interpretation that other actors (e.g. the accused, third-party states) could accept, reject or ignore. These interactions may thus interpret and articulate the meaning of the norm in ways that clarify future expectations for state behaviour.

But accusations may also play a key role in constructing new norms from scratch. They can do this in several ways. The most prominent cyber operations (Estonia, Stuxnet, WannaCry) are defined by their novelty; they did things never seen before or on a scale not previously thought possible.[62] It was often unclear if *any* norm existed to govern states engaging in these operations.[63] In such cases, an accusation serves as an opening bid, aimed at a particular community, indicating not just the accuser's disapproval of the cited operation, but often, too, its proposal (perhaps implicit) that all such conduct should be barred, i.e. that there should be a norm against such conduct.[64] Accusations may thus lay out the contours of 'bad behaviour' along with an argument about why, exactly, the behaviour is undesirable. Other actors may then respond to the accusation. They may accept some of it; they may accept all of it; they may accept it in some situations but not others; or they may reject it entirely. It is these interactions between the accuser, the accused and third party audiences that – over time – may result in the creation of a new norm (or its failure).[65]

The United States may have employed such a constitutive strategy in suggesting that certain cyber operations (e.g. the Sony Hack, 2016 election interference) violated 'established international norms'.[66] Ambiguity in the US statements leaves open which norms it believes were violated, and the accused have denied the US charges.[67] Nonetheless, the US accusations also served as an invitation to other like-minded states to express similar views on the appropriate norms of behaviour. In the case of US accusations about election interference, foreign and security ministers from the G7 subsequently issued a joint statement denouncing foreign attempts to interfere in

[62] See, e.g., K. Zetter, *Countdown to Zero Day: Stuxnet and the Launch of the World's First Digital Weapon* (2014); Graham, 'NHS Cyber Attack: Everything You Need To Know About "Biggest Ransomware" Offensive in History', *The Telegraph* (20 May 2017), available at www.telegraph.co.uk/news/2017/05/13/nhs-cyber-attack-everything-need-know-biggest-ransomware-offensive/.

[63] See Katzenstein, 'Introduction: Alternative Perspectives on National Security', in P. J. Katzenstein (ed.), *The Culture of National Security: Norms and Identity in World Politics* (1996) 1, at 5 (defining norms as 'collective expectations for the proper behavior of actors with a given identity').

[64] Where the opening bid comes from a private – as opposed to a public – actor, the constitutive function of the accusation may be more limited. That said, even private company accusations may have constitutive value if done at the behest of a state or where other states accept and adopt the behavioural lines drawn by the accusation as an existing or developing customary legal norm.

[65] Finnemore and Hollis, *supra* note 2, at 475–477.

[66] See Office of the Press Secretary, 'Statement by the President on Actions in Response to Russian Malicious Cyber Activity and Harassment' (29 December 2016), available at https://obamawhitehouse.archives.gov/the-press-office/2016/12/29/statement-president-actions-response-russian-malicious-cyber-activity (opposing 'Russia's efforts to undermine established international norms of behavior and interfere with democratic governance'); Kerry, *supra* note 14 (the US Secretary of State condemns North Korea for the Sony hack as 'lawless acts of intimidation' that 'demonstrate North Korea's flagrant disregard for international norms').

[67] See, e.g., Davis II et al., *supra* note 4, at 2; Grove and Simmons, *supra* note 7.

democratic processes, including 'through cyber-enabled activities'.[68] That norm was then endorsed by 1000+ governments, firms, universities and civil society institutions who have signed the French Government-led, *Paris Call for Trust and Security in Cyberspace*.[69]

Accusations may also help construct new norms by supporting extant norm proposals. For example, in 2015, the consensus report of the UN Group of Governmental Experts (GGE) on Information Security identified 11 'voluntary' norm candidates for responsible state behaviour in peacetime.[70] States have endorsed these subsequently in various fora, yet concerns about operationalizing them remain.[71] Accusations offer a way to do this; states might consider incorporating references to GGE norms in their accusations to signal to the accused (and the international community as a whole) that these norms are more than words on paper, adding clarity to expectations for appropriate state behaviour going forward. Similarly, accusations might complain about non-conformance with best practices or confidence-building measures like those promulgated by the Organization of Security and Cooperation in Europe to help constitute them as normative expectations.[72] Cyber accusations could help build out legal norms in similar ways. Several scholars have already examined accusations of cyber operations such as WannaCry and Russia's 2016 election interference in terms of their (non)conformance with existing rules of international law.[73] Although states have not done so to date, they could take similar steps to incorporate legal claims in their accusations.

It would be a mistake, however, to assume that a state's silence on the international legal implications of its accusation means that the accusation has none. Customary international law does not emerge immediately and fully formed. It

[68] G7, Joint Statement of Foreign and Security Ministers, 'Defending Democracy: Addressing Foreign Threats' (April 2018), available at www.g8.utoronto.ca/foreign/180423-democracy.html.

[69] Ministère de l'Europe et des Affaires Étrangères, 'Cybersecurity: Paris Call of 12 November 2018 for Trust and Security in Cyberspace' (12 November 2018), available at www.diplomatie.gouv.fr/en/french-foreign-policy/digital-diplomacy/france-and-cyber-security/article/cybersecurity-paris-call-of-12-november-2018-for-trust-and-security-in (includes list of stakeholder signatories).

[70] See United Nations General Assembly (UNGA), Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security, UN Doc. A/70/174 (22 July 2015), at ¶ 13 ('2015 GGE Report').

[71] See, e.g., GA Res. 73/27, 5 December 2018, at ¶ 1 ('welcom[ing]' norms enshrined in the 2013 and 2015 GGE Reports); Association of Southeast Asian Nations (ASEAN), ASEAN–United States Leaders' Statement on Cybersecurity Cooperation (16 November 2018), at ¶ 5, available at https://asean.org/asean-united-states-leaders-statement-cybersecurity-cooperation/; G7 Declaration on Responsible State Behavior in Cyberspace (11 April 2017), at 3–4, available at www.mofa.go.jp/files/000246367.pdf.

[72] See, e.g., Organization for Security and Co-operation in Europe (OSCE), Decision No. 1202 OSCE Confidence-Building Measures to Reduce the Risk of Conflict Stemming from the Use of Information and Communication Technologies, PC.DEC/1202, *PC Journal* 1092 (10 March 2016), available at www.osce.org/files/f/documents/d/a/227281.pdf.

[73] See, e.g., Schmitt and Fahey, 'WannaCry and the International Law of Cyberspace', *Just Security* (22 December 2017), available at www.justsecurity.org/50038/wannacry-international-law-cyberspace/; Ohlin, 'Did Russian Cyber Interference in the 2016 Election Violate International Law?', 95 *Texas Law Review* (2017) 1579.

is the product of interactions and iterations over time, where a sufficiently uniform practice is generally (although not universally) accepted as *opinio juris* (i.e. recognized as being legally obligatory).[74] Today's accusations may serve as early evidence of a 'usage' – that is, a habitual practice followed without any sense of legal obligation. If such accusations persist and spread over time, states may come to assume that these accusations are evidence of *opinio juris*, delineating which acts are either appropriate or wrongful as a matter of international law.[75] In other words, accusations can directly contribute to the formation of customary international law.

While accusations may help construct or elaborate some new international law rules, they can also do the opposite: they can help undermine the development of other, potentially permissive, customary international laws.[76] By objecting and making accusations of wrongdoing, states and other actors can limit the potential for the accused's behaviour to become legally accepted. The UN International Law Commission emphasized this point in its recent *Draft Conclusions on Identifying Customary International Law*, noting how a failure to react to behaviour can constitute evidence that it is lawful.[77] In other words, 'tolerance of a certain practice may indeed serve as evidence of acceptance as law (*opinio juris*) when it represents concurrence in that practice'.[78] Thus, whether or not states currently characterize their cyber accusations in explicitly legal terms, by signalling disapproval of certain cyber acts (as President Obama did with respect to Chinese cyber espionage), these accusations counteract claims that the accused state's operations are (or are becoming) permitted by international law.[79]

[74] Many of the constitutive elements of custom are ambiguous (How many states must engage in a practice for it to be sufficiently general?) or contested (Can states engage in 'practice' by words rather than deeds? Can *opinio juris* be presumed or must it take an express form?). See, e.g., Norman and Trachtman, 'The Customary International Law Game', 99 *AJIL* (2017) 541, at 542; Guzman, 'Saving Customary International Law', 27 *Michigan Journal of International Law* (2005) 115, at 122; Roberts, 'Traditional and Modern Approaches to Customary International Law: A Reconciliation', 95 *AJIL* (2002) 757, at 757–758.

[75] Not all customary international law originates from a usage; it is possible for state practice to develop where the acts (or inaction) are accompanied by *opinio juris* from the outset.

[76] For certain international lawyers, this is *the* critical question given the theory that what international law does not prohibit, it permits. See, e.g., *S.S. Lotus (France v. Turkey)*, Judgment, 1927 PCIJ (ser. A) No. 10, 17 September 1927, at 18–19.

[77] See ILC, Identification of Customary International Law: Text of the draft conclusions as adopted by the Drafting Committee on second reading, UN Doc. A/CN.4/L.908, 17 May 2018 (Conclusion 10(3): 'Failure to react over time to a practice may serve as evidence of acceptance as law (*opinio juris*), provided that States were in a position to react and the circumstances called for some reaction').

[78] UNGA, Report of the International Law Commission, UN Doc. A/73/10, 30 April–1 June and 2 July–10 August 2018, at 140–41.

[79] Alternatively, if other states accept or acquiesce in the legality of certain state or state-sponsored cyber operations, the accusing state may be able to employ its accusation to claim the status of a persistent objector. *See* J. Crawford, *Brownlie's Principles of Public International Law* (8th ed., 2012), at 28.

4 Disaggregating Accusations: Attribution, Exposure, Condemnation

Successful accusations require factual knowledge of cybersecurity incidents. Such information is not always cheap or easy to obtain, but assembling the corroborating details of malicious activity to construct accusations can be an important legal tool for an array of savvy actors seeking to enforce, deter, defend or delineate bad behaviour online. But just as accusations may differ in *why* they are made, they may also differ in *how* they are formulated. Broadly conceived, accusations of malicious cyber activity share some or all of three common features: (i) attribution, (ii) exposure and (iii) condemnation.

A *Attribution*

Attribution is the process of answering the age-old question of who did what, exactly.[80] In international politics, efforts to attribute actions to named actors can take many forms, including individual investigations, fact-finding missions, truth and reconciliation commissions and the decisions of international courts and tribunals.

For our purposes, attribution is the identification or assignment of responsibility for a cyber operation.[81] Unlike physical and static identifiers used in other contexts (e.g. fingerprints), digital attribution involves very different technical indicators and patterns that may complicate the process.[82] Much of the cybersecurity literature focuses extensively on these technical aspects of attributing responsibility for cyber incidents.[83] Yet, as Herb Lin emphasizes, cyber attributions may require more than a technical analysis, depending on their goal(s). Does attribution seek to identify (i) the machine that enabled intrusion into the victim's systems; (ii) the human perpetrator that set the intrusion in motion; or (iii) the adversary (e.g. a state) directing that human and ultimately responsible for the incident?[84] The latter two efforts will usually require other sources of intelligence beyond the technical indicators that point to a particular IP address or network.

Whatever the goal, cyber attribution is not binary – possible or impossible. Rather, as Rid and Buchanan explain, cyber attributions vary in both confidence

[80] See, e.g., Rid and Buchanan, 'Attributing Cyber-Attacks', 38 *Journal of Strategic Studies* (2015) 4, at 4.

[81] We use the term 'responsibility' in a broad sense to include not only circumstances satisfying the legal standards for 'state responsibility' under international law, but also circumstances holding a state politically accountable for certain behaviour. In other words, our definition of responsibility is not tied exclusively to the evidentiary standards or control requirements subject to so much attention in international law (although we recognize that broad claims of responsibility may themselves impact how existing international legal standards and requirements manifest themselves vis-à-vis cyber operations). *See infra* note 151 and accompanying text.

[82] See Davis II et al., *supra* note 4, at 9–10.

[83] See, e.g., Wheeler and Larsen, 'Techniques for Cyber Attack Attribution', *Institute for Defense Analyses* (October 2003), available at https://apps.dtic.mil/dtic/tr/fulltext/u2/a468859.pdf.

[84] Lin, *supra* note 22, at 8–19; Davis II et al., *supra* note 4, at 9.

and specificity.[85] Thus, when the University of Toronto's Citizen Lab uncovered the 'Ghostnet' cyber espionage network targeting Tibetan institutions, its analysis circumstantially pointed to China as the culprit, but never formally named the identity of the attackers.[86] In contrast, the UK Foreign Ministry indicated that it was 'highly likely' that 'North Korean actors known as the Lazarus Group were behind the WannaCry ransomware campaign'.[87] A US cybersecurity company, Mandiant, concluded that Unit 61398 of China's PLA was the source of a long-standing commercial cyber espionage campaign, baring

> A secret, resourced organization full of mainland Chinese speakers with direct access to Shanghai-based telecommunications infrastructure . . . engaged in a multi-year, enterprise scale computer espionage campaign right outside of Unit 61398's gates, performing tasks similar to Unit 61398's known mission.[88]

Accusers can always attribute with less certainty or specificity than their actual knowledge, and often do so to protect sources and methods. It is also possible to have attribution at one level (e.g. to a machine, to a territory, to a person, to a state) but not others. Jason Healey, for example, shows it is possible to attribute responsibility for a cyber operation to a particular state even without evidence permitting attribution to particular individuals.[89]

B *Exposure*

Exposure refers to the publicity an accusation receives. Accusers face an array of strategic choices about how, and how much, to expose about a cyber operation. Some accusations are communicated privately between the accuser and the accused.[90] Other accusations may be more public, communicated among members of a specific and limited community. Still others may be shared widely with the public at large. Of course, we have multiple examples of the latter in the cybersecurity context, from Estonia's public claims of Russian responsibility for the 2007 DDoS

[85] Rid and Buchanan, *supra* note 80, at 7.

[86] Information Warfare Monitor, *Tracking GhostNet: Investigating a Cyber Espionage Network* (29 March 2009), at 12–13, available at www.nartv.org/mirror/ghostnet.pdf.

[87] Foreign & Commonwealth Office, *Foreign Office Minister Condemns North Korean Actor for WannaCry Attacks* (19 December 2017), available at www.gov.uk/government/news/foreign-office-minister-condemns-north-korean-actor-for-wannacry-attacks; see also NCSC, *supra* note 15 (NCSC had 'high confidence that the GRU was almost certainly responsible' for cyber operations); Global Affairs Canada, *supra* note 15 (assessing Russian responsibility for cyber operations with 'high confidence'); GCSB, *supra* note 15 (New Zealand assessing it 'highly likely' that the GRU was behind certain cyber campaigns).

[88] Mandiant Intelligence Center, 'APT1: Exposing One of China's Cyber Espionage Units' (19 February 2013), available at www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf ('Mandiant, *APT1*').

[89] *See* Healey, 'Beyond Attribution: Seeking National Responsibility for Cyber Attacks' (22 February 2012), available at www.atlanticcouncil.org/in-depth-research-reports/issue-brief/beyond-attribution-seeking-national-responsibility-in-cyberspace/.

[90] Much of the work of the International Committee of the Red Cross operates this way. See, e.g., MacInnis, 'International Red Cross Issues Rare Myanmar Censure; (29 June 2007), available at www.reuters.com/article/idUSL28287051 (ICRC 'normally deals under a cloak of confidentiality').

attacks against its systems, to the US, UK and Australian accusations that North Korea launched WannaCry.[91]

The content of private or semi-private accusations is, by design, harder to discern, but their use is clearly widespread. Transnational technical communities (e.g. FIRST) or industry collectives (e.g. the Cybersecurity Tech Accord) certainly have information that could underpin accusations, but may be reluctant to share it publicly because of concerns that doing so might harm individual members of these groups.[92] Similarly, private accusations by states, perhaps through diplomatic channels, may be a useful first step in an escalatory ladder. After formally accusing Russia of using cyber means to interfere in the 2016 US presidential election, for example, President Obama revealed that he had first privately conveyed the accusation to President Putin directly.[93]

Accusers interested in exposure of a cyber operation must choose what vehicle they will use and what evidence to share. Accusers may proffer accusations, themselves, directly. States can use press releases and speeches to make accusations; private cybersecurity companies issue reports detailing their claims.[94] Alternatively, accusers may use proxies to expose information about a cyber operation. Proxies, as agents, might be engaged by and answerable to a variety of principals. States may use proxies, as the USA used Mandiant and its APT1 report as part of a larger effort to accuse China of acts of commercial cyber espionage.[95] CrowdStrike was authorized by its client – the Democratic National Committee (DNC) – to make public its accusation that Russia had hacked the DNC's systems.[96] Media reports may perform a similar function, using 'anonymous' government sources to advance or confirm the existence of an accusation. Although they were unwilling at the time to accuse Iran directly, US officials used media outlets in 2012 to publicize their views that Iran had launched a series of cyberattacks against US banks.[97]

[91] See, e.g., Bickers, 'UK and US Blame "WannaCry" Cyber Attack on North Korea' (20 December 2017), available at www.news.com.au/technology/online/security/uk-and-us-blame-wannacry-cyber-attack-on-north-korea/news-story/fe2218525eb04875a92a4479e2580d2f; Associated Press, 'Estonia Links Moscow to Internet Attack', *NYT* (18 May 2007).

[92] See, e.g., The Forum of Incident Response and Security Teams (FIRST), available at https://first.org/; The CyberSecurity Tech Accord, available at https://cybertechaccord.org/.

[93] *See* Landler and Sanger, 'Obama Says He Told Putin: "Cut It Out" on Hacking', *NYT* (16 December 2016).

[94] See, e.g., NCSC, *supra* note 15; 2014 PLA Indictments, *supra* note 16; Kerry, *supra* note 14; Mandiant, *APT1*, *supra* note 88; Johnson, 'SWIFT Attackers' Malware Linked to More Financial Attacks', *Broadcom* (26 May 2016), available at https://community.broadcom.com/symantecenterprise/communities/community-home/librarydocuments/viewdocument?DocumentKey=8ae1ff71-e440-4b79-9943-199d0adb43fc&CommunityKey=1ecf5f55-9545-44d6-b0f4-4e4a7f5f5e68&tab=librarydocuments (accusing the Lazarus group affiliated with North Korea of hacking the Bangladesh Central Bank).

[95] See Mandiant, *APT1*, *supra* note 88.

[96] Alperovitch, *supra* note 33.

[97] See, e.g., Mount, 'U.S. Officials Believe Iran Behind Recent Cyber Attacks', *CNN* (16 October 2012), available at www.cnn.com/2012/10/15/world/iran-cyber/index.html. Several Iranians were later indicted for their participation in these operations. See 2016 Iranian Indictment, *supra* note 16.

As with attribution, accusers must also make strategic choices about how much documentation to employ when exposing a cyber incident. Detailing – and documenting – what happened bolsters an accusation's credibility.[98] Part of what has made accusations from the likes of Mandiant (now FireEye) or the University of Toronto's Citizen Lab so powerful are the technical details employed to support their claims.[99] But documenting accusations also comes with costs and risks. Hacking victims – both states and firms – are often reluctant to reveal the extent of intrusion, exfiltration or damage. Neither states nor firms want to appear weak or vulnerable. Firms often fear drops in share price or loss of customer confidence.

The means and methods by which accusers investigate a cyber incident may also be proprietary to companies or classified for states. Documenting the accusation thus risks giving the accused or third parties information that can be used to degrade future investigative efforts. They may even create new opportunities for offensive cyber operations. Although they were not disclosed in an accusation, the theft and leak of certain US National Security Agency (NSA) surveillance tools demonstrates just how much harm can follow the disclosure of means and methods: the NSA's tools provided the foundation for both the WannaCry and NotPetya ransomware attacks.[100]

Consequently, cyber accusations vary in terms of the exposure of underlying documentation. When the United States originally pointed the finger at North Korea for the Sony Pictures hack, it did not document what support it had for the accusation.[101] This led to some disagreement about the accuracy of US charges.[102] In contrast, the United States accusation of Russia hacking the DNC included details that allowed the accused and third parties to evaluate the claim.[103] Reputation and credibility matter greatly in the latitude an accuser has in disclosing supporting details when making accusations. If the accuser has a record of veracity and has technical capacity for sophisticated forensics and good intelligence, accusations with less detail may still be widely credible. As accusations of cyber operations become more common, we expect demands for documentation to rise, along with efforts to normalize how much substantiation should accompany an accusation.[104]

[98] Beutz Land, *supra* note 19, at 208 (discussing how the quality of the 'naming evidence' matters).

[99] See Mandiant, *APT1*, *supra* note 88; Information Warfare Monitor, *supra* note 86.

[100] See, e.g., Newman, 'The Leaked NSA Spy Tool that Hacked the World', *Wired* (7 March 2018), available at www.wired.com/story/eternalblue-leaked-nsa-spy-tool-hacked-world/.

[101] Office of the Press Secretary, *supra* note 41.

[102] Compare Leupp, 'A Chronology of the Sony Hacking Incident', *CounterPunch* (29 December 2014), available at www.counterpunch.org/2014/12/29/a-chronology-of-the-sony-hacking-incident/ and Roberts, 'New Clues in Sony Hack Point to Insiders, Away from DPRK', *Security Ledger* (28 December 2014), available at https://securityledger.com/2014/12/new-clues-in-sony-hack-point-to-insiders-away-from-dprk/ with Novetta, 'Operation Blockbuster: Unraveling the Long Thread of the Sony Attack' (24 February 2016), available at www.novetta.com/2016/02/operation-blockbuster-unraveling-the-long-thread-of-the-sony-attack/.

[103] Joint Analysis Report, *supra* note 54.

[104] In 2015, a UN GGE reached consensus on a voluntary norm under which states would not make unsubstantiated accusations with respect to another state's purported cyber operations. See 2015 GGE Report, *supra* note 70, ¶ 28(f).

C *Condemnation*

Condemnation refers to an expression of disapproval.[105] This term remedies two problems with the extant concept of 'shaming'. It removes ambiguity. 'Shaming' might refer either to actions of an accuser or to the emotional state of the accused; condemnation is unambiguously an action of accusers and is separate from the accused's feelings. Second, and related, it avoids anthropomorphizing states.

Implicitly or explicitly, condemnations usually have a reference point – a normative standard from which the accused's behaviour supposedly diverged. Condemnations can vary in the specificity with which they reference the normative standard. In some cases, the standard may be left unstated, and the accused's behaviour is simply labelled as 'bad'. At other times, the normative standard may be referenced explicitly.[106] In cyberspace, accusations to date have condemned the accused's behaviour in general terms (e.g. as 'malicious').[107] In a few cases, such as the Sony Hack and WannaCry, the condemnation suggested that the accused had violated 'international norms', albeit without identifying which norms specifically.[108] President Obama referred to the Sony Pictures hack as an act of 'cyber vandalism', but that was a novel phrase without any clear international normative antecedents. [109]

Such unspecific condemnation is not due to an absence of normative candidates. In 2015, a UN GGE reached consensus on a list of 'voluntary' norms of responsible state behaviour in peacetime.[110] Moreover, as the two *Tallinn Manuals* demonstrate, international law may offer a range of rules to both constrain and facilitate state cyber operations.[111] Both sources thus purport to offer normative guidance for states' cyber activity outside of armed conflicts and short of the use of force. Yet, states have not used the GGE's language (e.g. its prohibition on targeting critical infrastructure in peacetime) to condemn other states' cyber operations, even though Russia purportedly targeted Ukrainian power grids.[112] Moreover, as Efrony and Shany's survey

[105] Although the term 'shaming' also suggests opprobrium, we do not use it here because it suggests a capacity for the accused to have an 'emotional' response to the accusation that is disputed. See Friman, *supra* note 19, at 18 ('the extent to which targets actually feel ashamed on their actions being revealed may be more wishful thinking on the part of advocacy networks than reality'). We prefer to reserve our position on whether states can feel shame and employ the term 'condemnation' instead to capture the accuser's disapproval of the conduct in question.

[106] Condemnations may, moreover, invoke norms that have different bases of propriety. Norms can delineate appropriate behaviour by reference to culture, politics, religion or law (whether domestic or international). See Finnemore and Hollis, *supra* note 2, at 441–442.

[107] See US Department of the Treasury, *supra* note 14; Office of the Press Secretary, *supra* note 14 and accompanying text.

[108] *Ibid*.

[109] See Office of the Press Secretary, *supra* note 41; Sullivan, *supra* note 41.

[110] 2015 GGE Report, *supra* note 70.

[111] *See* M. Schmitt (ed.), *Tallinn Manual 2.0 on the International Law Applicable to Cyber operations* (2nd ed., 2017) ('*Tallinn 2.0*'); M. Schmitt (ed.), *Tallinn Manual on the International Law Applicable to Cyber Warfare* (2013).

[112] This may be because not all states believe that the 2015 GGE Report – the product of consultations among 20 states – reflects global norms. Or, it may be because states believe that Ukraine and Russia were in a state of international armed conflict at the time of the power grid hack, meaning that the GGE's peacetime norms were inapplicable.

reveals, states have, to date, usually refrained from condemning cyber operations with reference to the *Tallinn Manuals* or the international law they purport to codify.[113] In one of the only accusations referencing international law to date, the United Kingdom accused the GRU of a 'flagrant violation' of international law in a series of cyber operations – including those targeting the Ukraine transport system, OPCW, WADA, the US DNC and various businesses and users – albeit without explaining which laws were violated or which operations did so.[114]

D *Constructing Accusations*

Accusers face a number of trade-offs and strategic choices when constructing their accusations. Depending upon their goals and the function(s) they want the accusation to serve, accusers might combine some or all features – *attribution*, *exposure* and *condemnation* – and do so in different ways. In what follows, we examine some of these choices, trade-offs and possibilities for accusation construction by states.[115]

1 Attribution

Not all features are required in every type of accusation. Attribution, for example, may not be necessary if the accusation aims primarily at shoring up defences. Simply sharing the vulnerability and technical indicators of the malware may be enough to both prompt and enable defensive measures by diverse parties. The original accusations surrounding the TRITON/Trisis malware did not identify its authors, but cybersecurity firms nonetheless alerted relevant communities to defend against the threat posed.[116] Similarly, attribution will not always be relevant for accusations designed to constitute new norms. An accuser may identify unwanted behaviour without identifying its perpetrator, and then call on other members of the relevant community to join in condemning such behaviour, thus contributing to the development of a norm that prohibits it.

[113] Efrony and Shany, *supra* note 5, at 73.

[114] See NCSC *supra* note 15, and accompanying text (NCSC report contains the quoted language); see also EP Resolution of 13 June 2018, 2018/2004(INI) (indicating some 'malicious cyber activities' by state actors like 'Russia, China and North Korea among others, but also non-state actors (including organised crime groups) … disregard and violate international law … ').

[115] Non-state actors may also make accusations strategically, although their impact on enforcement, and certainly constitution, may be more indirect. Further research is needed to examine the trade-offs between state and non-state accusations, as well as any relevant differences among non-state actor accusations (i.e. do accusations by NGOs operate differently from those by commercial actors?).

[116] See Groll, 'Cyberattack Targets Safety System at Saudi Aramco', *Foreign Policy* (21 December 2017), available at https://foreignpolicy.com/2017/12/21/cyber-attack-targets-safety-system-at-saudi-aramco/. This is not to suggest that attribution is not relevant to defending against cyber threats; there may be added value in knowing where it came from. Our point is simply that attribution is not a necessary condition for accusations to have defensive value. With respect to the TRITON/Trisis malware, moreover, at least one cybersecurity company later attributed it to the Russian Federation. See, e.g., FireEye Intelligence, 'TRITON Attribution: Russian Government-Owned Lab Most Likely Built Custom Intrusion Tools for TRITON Attackers' (23 October 2018), available at www.fireeye.com/blog/threat-research/2018/10/triton-attribution-russian-government-owned-lab-most-likely-built-tools.html.

Some attribution appears necessary, however, for accusations seeking enforcement or deterrence. Enforcement requires identifying which actor(s) must change their behaviour or suffer punishment.[117] Similarly, an accusation's deterrent value lies in showing third parties that they, too, could be identified, accused and punished if they engage in the cited behaviour.

For accusations that require attribution, accusers must weigh how much specificity and certainty to convey. Enforcement actions premised on punishing a state's agents will require as much detail as the accuser can muster – indictments, after all, must not only name the person accused, they must also detail the factual bases for the criminal charges against that person. In contrast, accusations targeting the state itself may successfully change behaviour even with imprecise charges and incomplete evidence. Indeed, strategic ambiguity in the framing of accusations may be diplomatically useful and create face-saving opportunities for enforcement. States sensitive to stigma may respond better to more obliquely framed accusations that refrain from implicating the government directly but simply state that the cyber incident emanated from their territory. This gives the accused government opportunities and incentives to respond positively (e.g. through domestic prosecutions or cessation of the operation) without conceding complicity in the first place. This was then-Secretary of State Hillary Clinton's approach with respect to 'Operation Aurora', where Google's source code was lost as a result of intrusions from China.[118] China followed a similar path in response to media reports linking it to the hacking of the US OPM – i.e. rather than admitting its responsibility, China identified and arrested several Chinese citizens, claiming they were the real culprits (charges many US officials regarded as suspect).[119]

2 *Exposure*

Exposure and publicity also may not be necessary components of effective accusations. Private accusations may work well (or better) than their public counterparts. Enforcement may be pursued publicly or privately; a state taking counter-measures may be obligated to communicate its intentions to the accused, but it has no obligation to communicate them more broadly.[120]

[117] The naming and shaming literature has already recognized a version of this problem. While naming and shaming may be an effective tool with respect to certain types of civil and political rights, it has proven more difficult to apply to economic and social rights where violations are not attributable to a particular actor. Who is to blame for hunger or poverty or lack of shelter and medicine in poor countries? Even activists do not agree. States may technically be the 'duty bearers' for fulfilment of economic and social rights, but if citizens, activists and other states do not see poor state governments as the cause of violations (i.e. governments are not intentionally starving or impoverishing their people) then they are unlikely to change their behaviour. See M. Jurkovich, *Feeding the Hungry* (2020).

[118] McGreal and Johnson, 'Hillary Clinton Criticises Beijing Over Internet Censorship', *The Guardian* (21 January 2010), available at www.theguardian.com/world/2010/jan/21/hillary-clinton-china-internet-censorship.

[119] Nakashima, 'China: Hackers Arrested', *Wash. Post* (3 December 2015).

[120] See ASR, *supra* note 48, Art. 52, at 135.

Exposure is essential, however, in accusations designed to shore up defences, to deter or to constitute new norms. If third parties do not know about a cyber operation, they cannot defend against it, and nor are they likely to be deterred from engaging in similar behaviour. Similarly, the construction of norms involves public communications directed at (or among) the community of actors to which the norm should apply. When it comes to customary international law, for example, there must be some observable 'practice' that states can join or resist and to which the requisite *opinio juris* attaches.

3 Condemnation

Condemnation plays a key role in accusations pursuing deterrence, clarifying which behaviour is to be avoided. Condemnations may have less purchase in accusations that emphasize defence. Reports like Mandiant's on the APT1 attack show that it may be enough to expose and attribute 'malicious' acts; little by way of explicit condemnation was needed.[121]

The relationship between condemnation and enforcement is more complicated. Strong condemnation may, indeed, cause the accused to change its behaviour, as the 'naming and shaming' logic suggests. But condemnation – particularly public condemnation – also risks stigmatization that may lead an accused to 'dig in', retrench and repeat the condemned behaviour.[122] Those involved in truth and reconciliation commissions are often at some pains to avoid alienating key parties negotiating a future peace for exactly this reason.[123] Similarly, when regulatory authorities try to move companies towards better behaviours, they must balance condemnation via fines or public sanctions (which may be a useful deterrent) with potential stigmatization that could spawn evasive behaviour, make crime worse and create adversarial relationships between regulator and companies that are counterproductive.

Where accusers fear a backlash, they may substitute technical assistance for condemnation in a process known as 'reintegrative shaming'.[124] The logic here is to couple stigmatization with opportunities and help in adopting more prosocial behaviours. Social science research suggests this approach can produce better results, especially in situations where there is ambiguity about the relevant rules of behaviour and culpability may not be felt.[125] In cyber contexts, this approach might be useful where otherwise lawful perpetrators failed to act or acted negligently (e.g. by failing to be diligent in ensuring an otherwise lawful cyber operation stayed within its expected

[121] See Mandiant, *APT1*, *supra* note 88.

[122] See Adler-Nissen, 'Stigma Management in International Relations: Transgressive Identities, Norms, and Order in International Society', 68 *Int'l Org.* (2014) 143.

[123] Wiebelhaus-Brahm, 'Promoting Accountability, Undermining Peace? Naming and Shaming in Transitional Justice Processes', in Friman, *supra* note 19, at 86.

[124] van Erp, *supra* note 39, at 288.

[125] *Ibid.*, at 288, 290–291; see also Makkai and Braithwaite, 'Reintegrative Shaming and Compliance with Regulatory Standards', 32 *Criminology* (1994) 361.

parameters). Reintegrative shaming might also work where a state fails to operate with due diligence after it learns other actors have used its territory to launch a cyber operation. However, it is less likely to be useful for purposefully malicious actors. We also suspect it may have little effect when the accused operates outside – or at some distance from – the relevant community. Reintegrative shaming thus seems unlikely when dealing with rogue states like North Korea.

When it comes to constituting new norms, condemnation plays an obvious role. Condemnation articulates publicly what 'bad' (or unlawful) behaviour looks like and, perhaps implicitly, what 'good' (or lawful) behaviour might be. These articulations, often done through accusations, can then form the basis for a new norm or legal rule.

Perhaps counterintuitively, norm construction may sometimes occur without condemnation. Consider Stuxnet. On 1 June 2012, *New York Times* reporter David Sanger published a story that assigned responsibility for the virus (which destroyed up to 1,000 centrifuges in Iran's nuclear programme) to the United States and Israel.[126] Far from condemning the US and Israeli actions, Sanger presented the operation positively, as giving the accused states a new, non-lethal mechanism to oppose nuclear proliferation. Thus, one could interpret Stuxnet's exposure as an effort – in this case by media actors – to establish, not the maliciousness, but the propriety of using cyber capacities to thwart proliferation instead of more traditional kinetic means (with their attendant death and destruction).[127] The international community has not, however, embraced that idea. When and where such operations are appropriate remains unclear and contested. Subsequent reverse-engineering of Stuxnet into the Shamoon and BlackEnergy malware also suggests that the benefits of such cyber tools must be weighed against some significant costs.[128]

The fact that Stuxnet was celebrated in some circles and condemned in others reveals that accusations may work differently with different audiences. A condemnation may resonate with one audience but not another. The OPM hack was condemned within a US domestic law framework as a breach of national security, but the US Director of National Intelligence, James Clapper, indicated that such behaviour was acceptable among states: '"You have to kind of salute the Chinese for what they did," adding the US would have done the same thing if it could.'[129] The efficacy of an accusation thus depends on more than just alignment of its specific contents – attribution, exposure and condemnation – with its intended purpose(s). Context and audience will always affect an accusation's effects.

[126] Sanger, 'Obama Order Sped Up Wave of Cyberattacks Against Iran', *NYT* (1 June 2012).

[127] See Albright et al., *Did Stuxnet Take out 1,000 Centrifuges at the Natanz Enrichment Plant?* (22 December 2010), available at https://isis-online.org/isis-reports/detail/did-stuxnet-take-out-1000-centrifuges-at-the-natanz-enrichment-plant/.

[128] See Perlroth, 'In Cyberattack on Saudi Firm, U.S. Sees Iran Firing Back', *NYT* (23 October 2012).

[129] Sciutto, 'Director of National Intelligence blames China for OPM hack', *CNN* (25 June 2015), available at www.cnn.com/2015/06/25/politics/james-clapper-china-opm-hacking/index.html.

5 Under What Conditions Do Accusations Work?

Accusations will not work at all times or in all conditions. Context matters in many ways. Two contextual features of particular consequence are: the *normative environment* in which the accusation is made, and the *relationships* among accused, accuser and third parties.

A *Normative Environment: Is There a Norm?*

As an enforcement tool, accusations require a norm or some standard of propriety against which the accused's behaviour, and deviation from the norm, can be measured. The existing naming and shaming literature has not emphasized this condition, perhaps because some core shared norms are relatively clear in the areas featured in that literature (e.g. human rights, the environment). States – including accused violators – do not contest norms prohibiting torture, genocide or significant transboundary pollution.[130] Instead, accused states deny what the accused says happened or offer a different interpretation or application of the norm from that proffered by the accuser.[131] By contrast, the norms (and international law) governing online behaviour are not always clear and well-entrenched. Ongoing contestation about their existence and meaning make their enforcement via accusations tricky.

Consider, for example, recent debates about whether sovereignty is violated when cyber operations by one state create unwanted effects in another state's territory. *Tallinn Manual 2.0* says it is, as do the Dutch and French governments.[132] This view has been contested, however, by those who question whether sovereignty is a rule governing state behaviour or a background principle that informs the content of other rules (such as the duty of non-intervention).[133] The UK Attorney General, for example, has firmly placed the United Kingdom in the sovereignty-as-background-principle camp.[134] Accusations

[130] See *supra* note 61 and accompanying text. This pattern of denial and reinterpretation is central to Risse, Ropp and Sikkink's well-known 'spiral model' of human rights compliance. See Risse, Ropp and Sikkink (eds), *The Power of Human Rights: International Norms and Domestic Change* (1999) at 1, 17–35.

[131] See ICCPR, *supra* note 61 and accompanying text.

[132] *Tallinn 2.0*, *supra* note 111, at 17 (Rule 4); Ministère des Armées, 'Droit international appliqué aux operations dans le cyberspace' (9 September 2019), at sec. I.1.1; Dutch Minister of Foreign Affairs, 'Letter to the Parliament on the International Legal Order in Cyberspace' (5 July 2019) Appendix, at 2, available at www.government.nl/ministries/ministry-of-foreign-affairs/documents/parliamentary-documents/2019/09/26/letter-to-the-parliament-on-the-international-legal-order-in-cyberspace.

[133] See, e.g., Corn, 'Tallinn Manual 2.0 – Advancing the Conversation', *Just Security* (15 February 2017), available at www.justsecurity.org/37812/tallinn-manual-2-0-advancing-conversation/.

[134] Wright, 'Cyber and International Law in the 21st Century' (23 May 2018), available at www.gov.uk/government/speeches/cyber-and-international-law-in-the-21st-century. The General Counsel of the US Department of Defense has expressed a similar view: see Ney, 'DOD General Counsel Remarks at U.S. Cyber Command Legal Conference' (2 March 2020) available at www.defense.gov/Newsroom/Speeches/Speech/Article/2099378/dod-general-counsel-remarks-at-us-cyber-command-legal-conference/:

> For cyber operations that would not constitute a prohibited intervention or use-of-force [i.e. those that might be covered by a rule of sovereignty], the Department believes there is not sufficiently widespread and consistent State practice resulting from a sense of legal obligation to conclude that customary international law generally prohibits such non-consensual cyber operations in another State's territory.

that one state has violated another's sovereignty can thus get bogged down in an existential debate about sovereignty's legal status, diverting attention from the more focused question about what cyber operations are permissible.[135] Of course, it is exactly these sorts of existential debates that may iterate over time to constitute a new norm (or block a permissive rule from forming). As such, the existence of a norm is less of a pre-condition for the constitutive function of an accusation than it is for an accusation's enforcement efforts.

B *Relationships*

An accusation is a social practice whose meaning and effects are shaped by relationships among the accuser, the accused and the third-party audiences or communities in which those actors are embedded.

The relationship between accuser and accused will often influence the efficacy of an accusation. The more an accused values its relationship (whether political, economic or social) with the accuser, the greater the likelihood the accusation may prove effective. We would expect, for example, that accusations that the United Kingdom inappropriately targeted a European ally's critical infrastructure (e.g. a telecommunications carrier like Belgium's Belgacom) are more likely to lead to the cessation of that cyber operation than accusations that the United Kingdom targeted a similarly situated Russian company.[136]

Relationships between the accused and the larger community may also affect an accusation's result. Accusations seeking behavioural changes assume that perpetrators have pro-social reputations they want to protect and/or a moral compass of some kind.[137] This may not always be a good assumption. In cyberspace, for example, some actors (e.g. hacktivists with only loose ties to a state) may actually value a reputation for destructive cyber operations.[138] Indeed, they may seek to profit from it on the Dark Web or in other nefarious corners of the Internet. Rogue states, like North Korea, also may not care much about community opinion.

But in many instances, pro-social reputation matters to accused parties because they care about these relationships. Existing research has explored the likelihood that social ties may generate norm compliance. Goodman and Jinks find that the likelihood of a positive response to an accusation about a state depends on the strength, immediacy

[135] On the different implications of existential arguments in international law, see Hollis, 'The Existential Function of Interpretation in International Law', in A. Bianchi et al. (eds), *Interpretation in International Law* (2015) 78, at 78–79.

[136] See, e.g., Gallagher, 'How U.K. Spies Hacked a European Ally and Got Away with It', *The Intercept* (17 February 2018), available at https://theintercept.com/2018/02/17/gchq-belgacom-investigation-europe-hack/.

[137] The 'naming and shaming' literature has emphasized that the efficacy of accusations depends on the accused's sensitivity to communal pressure – i.e. how much it cares about belonging 'to a normative community of nations' and the international reputation that accompanies such status. See, e.g., M. E. Keck and K. Sikkink, *Activists Beyond Borders: Advocacy Networks in International Politics* (1998), at 29, 208.

[138] See Adler-Nissen, *supra* note 122, at 170.

and size of the group with which the accused shares an identity.[139] Interestingly, the social science research on which they rely suggests that the most effective groups have three to eight members, with the efficacy of compliance for larger groups dropping off rapidly.[140] At first glance, that fact does not bode well for international law and the nearly 200 nation states subject to it.[141] On the other hand, market concentration among big tech companies and high network concentration in a few states mean cultivating normative agreement among a relatively small number of actors can address many behaviour coordination challenges in cyberspace. Thus, accusations aimed at norm construction for a smaller audience of like-minded states may be a valuable first step in constructing broader norms.

An accuser's own reputation and relationships will also shape the efficacy of accusations generally, and norm construction specifically. Accusers often aim to change, not just the accused's behaviour, but all similar behaviour: they may want to create a new norm around the undesired behaviour (or to apply an existing norm in some new way). And it is the community – not the accused – that will decide whether such norm development bears fruit. The community's view of the accuser may therefore matter more to a proposed norm or law's reception than the community's view of the accused. States are more likely to accommodate well-documented normative claims on the (im)propriety or (un)lawfulness of WannaCry coming from high-status states like the United Kingdom (the accuser) rather than reactions from North Korea (the accused).[142]

For international law purposes, the importance of the accuser's identity is reinforced by the idea of a 'specially-affected state' – i.e. one 'who either engages in a practice that some states do not or is distinctively affected by a practice – directly or indirectly – in a manner that distinguishes it from other states'.[143] Where states are specially affected – either because they possess cyber operation capabilities that others do not, or because they have been the victim of cyber operations – international law may actually require the community of states to pay particular attention to their views on the state of customary international law. There are even recent suggestions that international law requires a majority of specially-affected states to support the formation of any new customary rule.[144]

[139] R. Goodman and D. Jinks, *Socializing States: Promoting Human Rights Through International Law* (2013), at 28. 'Strength' refers to the importance of the group to the accused; 'immediacy' to the accused's awareness of and interactions with that group; and 'size' to the number of members in the group. See *ibid.*

[140] *Ibid.* On the other hand, the 'cascade' effect by which norms form suggests that small groups can evolve into much broader coalitions. See generally Finnemore and Sikkink, 'International Norm Dynamics and Political Change', 52 *Int'l Org.* (1998) 887.

[141] Goodman and Jinks, *supra* note 139, at 28.

[142] See *supra* note 87 and accompanying text.

[143] See, e.g., Heller, 'Specially-Affected States and the Formation of Custom', 112 *AJIL* (2018) 191, at 193. Exactly which states qualify as specially affected remains a matter of some dispute, as does their role in the formation of customary international law. *Ibid.*, at 192–193.

[144] See *ibid.*, at 193.

Powerful accusations may also be deployed by both less powerful states – like Estonia – and non-governmental organizations who lack state authority and must rely on their reputation and credibility within the community. In these cases, the accuser's reputation matters more than material power. Is the accuser a trusted actor? Have its previous accusations been corroborated and accepted? Or, is the accuser perceived to have a self-interested agenda or motives that do not benefit the community as a whole?

Taken together, certain external conditions may be more or less relevant to the capacity of an accusation to achieve its desired function(s). Those functions may also be more or less achievable depending on the contents of the accusation itself; different elements of an accusation may be mandatory (or not) for different functions. Table 1 offers a tentative menu for how these considerations may align in constructing accusations for global cybersecurity.

6 Implications for International Law: Obstacles and Opportunities

A *The Absence of International Law in Existing Accusations?*

How do accusations interact with international law? Most obviously, they could be a source for its enforcement. If conditions are favourable, accused parties may become more compliant in response to the accuser's condemnation of an internationally wrongful act (or a failure to act). As noted, however, cyber accusations have been slow to take advantage of this possibility.[145] What explains this reluctance to invoke international law?

One reason may be contested views on legality and propriety. At least some of the accusations to date involve behaviour currently regarded as legally acceptable. The OPM hack, for example, may have severely undermined US national security at a scale not seen previously. Yet, from the perspective of international law, this was an act of espionage, which international law either fails to regulate or affirmatively permits.[146] As such, it is not surprising to see accusations against China avoid condemnation for the OPM hack in international legal terms.

Similar contestation may explain the reluctance to invoke other international legal rules that have divided states at the GGE and elsewhere.[147] The 2017 UN

[145] One exception is the accusations surrounding GRU cyber operations against the OPCW, WADA and other targets. See *supra* notes 15, 114, and accompanying text. In the last year, moreover, several states have issued statements offering general views on how international law applies to cyberspace. See, e.g., Ministère des Armées, *supra* note 132; Dutch Minister of Foreign Affairs, *supra* note 132; Kaljulaid, *President of the Republic at the opening of CyCon 2019* (29 May 2019), available at www.president.ee/en/official-duties/speeches/15241-president-of-the-republic-at-the-opening-of-cycon-2019/index.html. None of these, however, make accusations nor reference specific instances of international law violations.

[146] See Deeks, 'An International Legal Framework for Surveillance', 55 *Virginia Journal of International Law* (2015) 291, at 300.

[147] See Tikk, 'Will Cyber Consequences Deepen Disagreement on International Law?', 32 *Temple International and Comparative Law Journal* (2018) 185, at 187–189.

Table 1: Constructing accusations for global cybersecurity

Function	Elements		External Conditions	
	Required	**Optional**	**More Relevant**	**Less Relevant**
Enforcement	Attribution Condemnation	Exposure	Existing Norm Relations between Accuser & Accused Relations between Accused & Audience	Relations between Accuser & Audience
Defence	Exposure	Attribution Condemnation	Existing Norm Relations between Accuser & Audience	Relations between Accuser & Accused Relations between Accused & Audience
Deterrence	Attribution Exposure Condemnation	None	Existing Norm Relations between Accuser & Accused Relations between Accused & Audience	Relations between Accuser and Audience
Constitution	Exposure Condemnation	Attribution	Relations between Accuser & Audience	Existing Norm Relations between Accuser & Accused Relations between Accused & Audience

Group of Governmental Experts failed to achieve consensus reportedly because states were divided over whether certain international law rules that apply elsewhere (e.g. self-defence, international humanitarian law, sovereignty and due diligence) also apply in cyberspace.[148] In other cases (e.g. the duty of non-intervention) there may be agreement on the rule's existence but different interpretations of its meaning (i.e. in what cyber-areas is intervention prohibited? What constitutes the requisite 'coercion' to violate the duty in a cyber

[148] See Sukumar, 'The UN GGE Failed: Is International Law in Cyberspace Doomed as Well?', *Lawfare* (4 July 2017), available at www.lawfareblog.com/un-gge-failed-international-law-cyberspace-doomed-well; Schmitt and Vihul, 'International Cyber Law Politicized: The UN GGE's Failure to Advance Cyber Norms', *Just Security* (30 June 2017), available at www.justsecurity.org/42768/international-cyber-law-politicized-gges-failure-advance-cyber-norms/; see also *supra* notes 132–135 (debates over whether sovereignty is a rule or just a background principle for state cyber operations).

operation?). Such disparate interpretations make it difficult to expect enforcement to flow from an accusation of the rule's violation.[149] Consequently, some states may avoid accusing another state of acts they, themselves, believe violate a rule of international law (e.g. sovereignty) because they are unsure if the community as a whole would agree.

Another reason states may decline to invoke international legal rules is that they do not want the accusation to have constitutive effects. Some states may value the lack of clear rules. Ambiguous rules – or no rules – may provide more flexibility to engage in cyber operations that states value. Reciprocity concerns may operate along similar lines. Iran, for example, never challenged the US and Israeli role in Stuxnet as a use of force or even an armed attack (triggering a right of self-defence), preferring instead to deploy its own cyber operations against US financial targets without any legal rhetoric or justification at all.[150]

Even if an accuser perceives consensus around the existence of an international legal norm, documentation issues may serve as a barrier to referencing it. International legal accusations pose particular evidentiary challenges. Accusers must tie the accused state to the actual hackers, demonstrating that those hackers were government officials, affiliated with a non-state actor operating under the state's control or affiliated with a non-state actor's operations that were later adopted by the state.[151] International legal claims also require a particular standard of proof, and the accuser may not have sufficient evidence to meet that standard (or may resist burning the sources and methods to produce such evidence). Fear of reckless or spurious accusations is widespread and, indeed, among the norms agreed to by the 2015 UN GGE was: 'the accusations of organizing and implementing wrongful acts brought against States should be substantiated'.[152]

[149] The principle of non-intervention has a long, well-established pedigree from UNGA resolutions like the Declaration on the Principles of International Law concerning Friendly Relations, to its repeated endorsement by the ICJ. See, e.g., *Case Concerning Military and Paramilitary Activities in and Against Nicaragua (Nicaragua v. United States)*, Judgment, 27 June 1986, ICJ Reports (1986) 14, at 205; *Case Concerning Armed Activities on the Territory of the Congo (Democratic Republic of the Congo v. Uganda)*, Judgment, 19 December 2005, ICJ Reports (2005) 168; Declaration on Principles of International Law concerning Friendly Relations & Co-operation among States, GA Res. 2625 (XXV), 24 October 1970, UN Doc. A/RES/25/2625. *Tallinn 2.0*'s Independent Group of Experts incorporated the prohibition as Rule 66: 'A State may not intervene, including by cyber means, in the internal or external affairs of another State.' See *Tallinn 2.0*, *supra* note 111, at 312. As the Experts noted, however, 'the precise contours and application of the prohibition of intervention are unclear', with outstanding debates about (a) what falls within a state's 'internal' affairs? and (b) what constitutes a prohibited 'intervention'. *Ibid.*

[150] See *supra* note 97 and accompanying text.

[151] Of course, the standards of control required to establish a state's responsibility for acts of non-state actors are disputed, with the ICJ favouring a rule of 'effective control', in contrast to the rule of 'overall control' advocated by the International Criminal Tribunal for the Former Yugoslavia. Compare *Nicaragua v. United States*, *supra* note 149, at ¶ 115; *Case Concerning Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v. Serbia and Montenegro)*, Judgment, 26 February 2007, ICJ Reports (1997) 43, at ¶¶ 399–401; with Judgment, *Prosecutor v. Dusko Tadic aka 'Dule'* (ICTY-94-1-A), Appeals Chamber, 15 July 1999, at ¶¶ 131, 145.

[152] 2015 GGE Report, *supra* note 70, ¶ 28(f).

Other challenges in using international law may have little to do with norm constitution or enforcement. Jack Goldsmith and Stuart Russell emphasize that '[u]nless a nation is able to effectively redress a cyber intrusion, it can be harmful or self-defeating to publicize it, since public knowledge of loss and the failure to respond effectively invite more attacks'.[153] This may be true for all accusations, but it certainly resonates with respect to international law accusations specifically. States may be reluctant to make international legal claims where they lack available and effective enforcement remedies to bring the accused into compliance with their view of the law. And to the extent the accused are rogue actors, states may not find much added utility in invoking an international legal regime that the accused has demonstrated a willingness to flout in other contexts.

B *Deploying Accusations Strategically to Advance International Law*

A more nuanced understanding of how accusations work could help states better construct them for desired ends, including improved enforcement of international law and constituting the contents of the law itself. To date, cyber accusations have emphasized the former, with little to no attention to the latter possibility. However, careful crafting of accusations, with attention to their structure and function, could enhance their effectiveness for either end.

Levels of exposure (or publicity) can be strategically varied and ratcheted up to achieve desired goals. Very specific accusations made in private may open different diplomatic options for enforcement than public broadcast of that same information. Blanket proscriptions (calling on the accused to stop doing something) and wholesale stigmatization can be nuanced in useful ways. An alternative approach might be accusations critiquing states for a failure to act to control behaviour within their territory. This leaves the accused room to both save face and moderate behaviour. Accused states may respond to such accusations claiming these operations are perpetrated by actors beyond government control. Accusations aimed at a supposed lack of territorial control might induce the accused state to exercise more control over sub-state actors' unwanted behaviour.[154]

In addition, states and other stakeholders might consider the value of generating lists or rankings of states with good and poor records of international law compliance in cyberspace in much the same way as the United States has done with its annual Human Rights Report or as the World Bank does with its 'ease of doing business' rankings.[155] Tying poor performance in such listings to resources and assistance for capacity building (which helps the accused address the problem and reduce unwanted

[153] Goldsmith and Russell, 'Strengths Become Vulnerabilities How a Digital World Disadvantages the United States in Its International Relations', Aegis Series Paper No. 1806 (2018), at 3, available at www.hoover.org/sites/default/files/research/docs/381100534-strengths-become-vulnerabilities.pdf.

[154] Healey, *supra* note 89, at 4 ('Under international pressure, most nations could likewise reduce attacks from their territory of cyberspace through several well-established steps … ').

[155] Kelley and Simmons, 'Introduction: The Power of Global Performance Indicators', 73 *Int'l Org.* (2019) 491; Kelley and Simmons, 'Politics by Number: Indicators as Social Pressure in International Relations', 59 *American Journal of Political Science* (2014) 55, at 56.

cyber behaviour emanating from its territory) might add a 'carrot' to the 'stick' of a poor ranking.

Our vision of accusations suggests, however, that states and stakeholders should not limit their expectations to international law enforcement. Properly constructed, accusations may create opportunities to clarify the international law that currently governs state cyber operations and/or to constitute new rules that do so. Simple steps could improve the credibility of existing accusations – and the legal norms they promote – within the relevant communities of states or other stakeholders. Agreeing to more standardized attribution methodologies, for example, would make it easier for audiences to weigh an accusation's credibility. Standardizing condemnations, in turn, would help build the case for a 'uniform' practice – one of the elements in identifying new rules of customary international law.

States could also make more accusations collectively. In September 2019, 27 states issued a 'Joint Statement' that contemplated a set of unspecified collective actions to advance responsible state behaviour in cyberspace.[156] One action these states might consider would be collective accusations. Some precedents already exist for this.[157] Accusations involving WannaCry, NotPetya, the OPCW hack and Georgia suggest that increasing the number of accusers might raise the credibility of the claims made.[158] Custom requires not only a uniform practice, but a general one practised by most (but not necessarily all) nation states. More accusers expressing disapproval of certain cyber operations makes it easier to argue that those expressions of disapproval comprise a sufficiently general practice that could be accepted as *opinio juris.*

Alternatively, states might advance the power of accusations by creating an impartial institution to do attribution or advocate for international law.[159] A neutral or independent third party could collect attribution data from state and non-state actors who might be reluctant to share it publicly (or even with each other privately).[160] Doing so might dispel some fog surrounding accusations and counteraccusations. It might

[156] See US Department of State, 'Joint Statement on Advancing Responsible State Behavior in Cyberspace' (23 September 2019), available at www.state.gov/joint-statement-on-advancing-responsible-state-behavior-in-cyberspace/ ('When necessary, we will work together on a voluntary basis to hold states accountable when they act contrary to this framework, including by taking measures that are transparent and consistent with international law').

[157] See *supra* notes 15, 36, 58 and accompanying text for examples.

[158] See, e.g., Roguski, *supra* note 6 (in February 2020, 20 states collectively accuse Russia of conducting cyber operations against Georgia); BBC, 'Russia Cyber-Plots: US, UK and Netherlands Allege Hacking', *BBC* (4 October 2018), available at www.bbc.com/news/world-europe-45746837 (noting organized accusations by Canadian, Dutch, US and UK officials against the GRU).

[159] A new non-profit organization, the CyberPeace Institute, was recently established in Geneva. Its mission contemplates advocacy and accountability, which may involve harmonizing approaches to attribution. As yet, however, it does not appear likely to make accusations on its own. See CyberPeace Institute, available at https://cyberpeaceinstitute.org/about-us.

[160] See, e.g., Davis II et al., *supra* note 4, at 3; Healey et al., 'Confidence-Building Measures in Cyberspace', *Atlantic Council* (November 2014), available at www.atlanticcouncil.org/wp-content/uploads/2014/11/Confidence-Building_Measures_in_Cyberspace.pdf; Smith, 'The Need for a Digital Geneva Convention', *Microsoft On the Issues* (24 February 2017), available at https://blogs.microsoft.com/on-the-issues/2017/02/14/need-digital-geneva-convention/.

shift arguments from 'what happened?' to 'is what happened permissible and legal?'. An attribution organization might thus supplement and even strengthen currently disaggregated accusation and attribution efforts.[161] Similarly, an attribution organization could build and concentrate technical expertise. This would particularly benefit states that lack the capacity to adequately attribute, and thus broaden participation in the creation of new cyber norms. If neutral actors issue specific and credible reports of unwanted behaviour, they may lead states (either globally or in more like-minded groups) to coalesce around new international legal rules proscribing such behaviour.

Finally, policymakers might consider whether and how additional legal instruments could be helpful in making these accusations. If they want to clarify what the rules and norms *are* for cyber operations, there are clear steps they can take to do this, beyond simply waiting for practice and *opinio juris* to emerge organically over time. Having some agreed upon treaty framework for international law's application in cyberspace (whether bilateral, regional or global) would open up a range of new possibilities for using accusations about those treaty commitments to further construct rules of the road in cyberspace.

7 Conclusions

Generating compliance with international norms and legal rules is an ongoing struggle in world politics. States have a variety of tools for this purpose ranging from discrete – often private – criticism to public, even forceful, coercion. In this paper, we have investigated one such tool – the accusation – and the many ways it might be used to steer states towards more pro-social and norm-compliant behaviour in cyberspace.

For international relations (IR) scholars, our investigation builds on the well-understood dynamics of 'naming and shaming' but opens up that concept to reveal a much richer array of political possibilities. One such possibility is that naming and shaming do not always go together. IR literatures tend to assume they do – that once a state is 'named' in an accusation, shame and shaming behaviour will follow. Our paper starts from the puzzle that, in cybersecurity, this link between naming and shaming is weak and, of particular interest for readers of this journal, that international law is largely absent from naming, shaming and accusations in cybersecurity. Unpacking the structure of accusations helps us understand why this is so. Accusations are flexible tools. They can be constructed in diverse ways to accomplish diverse goals. They can also have effects beyond those expected.

Of particular interest to IR constructivists will be the role accusations can play in the constitution of new social norms, rules and law. This is a classic case of social construction in action. Accusers have to decide which accusations to make and how to frame (and justify) those charges. Accused parties, and third parties, have to decide how to respond – whether to accept or deny the accusation and, importantly, they must articulate their reasons for doing so. These repeated social interactions will, over

[161] See, e.g., Eichensehr, 'Decentralized Cyberattack Attribution', 113 *AJIL Unbound* (2019) 213. For Eichensehr's argument that decentralized attribution should continue, see Eichensehr, *supra* note 26.

time, determine the social contours of the cybersecurity issue space – what its rules are, who has authority there and how those rules and authorities came to be so.

For international lawyers, these political possibilities can play a critical role in identifying the existing legal rules and building new ones, *even when* states avoid the rhetoric of international law. Accusations – and the responses they generate – could advance international law enforcement. But the failure of accusations to enforce international law does not make law irrelevant. The interchanges following an accusation help reveal what behaviour is – and is not – accepted by the international community of states. The resulting delineation of wrongful behaviour in cyberspace can constitute the practice from which *opinio juris* may emerge over time.

For policymakers, our investigation offers a menu and a toolkit for thinking about whether and how accusations can be used to further their cybersecurity goals. When framing an accusation, accusers have choices. Do they want to name a perpetrator, or just announce that a cyber operation has happened and alert others to the threat? If they want to name perpetrators, do they want to name a government or specific individuals, or simply say the operation emanated from a named territory? How much evidence do they want or need to divulge to elicit the desired reaction either from the accused or from third parties? Do they want to make their accusation public immediately, or can they begin with a private conversation with the accused, and then escalate the accusation to a larger audience as needed? Different answers to each of these questions will lead to a different framing of an accusation, and different political consequences down the road.

States and other stakeholders thus have strategic choices to make as they survey global cybersecurity today. As visibility improves on who is doing what, there will be more opportunities to use accusations to set the rules of the road. In doing so, accusations may help, if not enforce, at least construct relevant constraints (and permissions) derived from international law.