### NAMING WITHOUT SHAMING? ACCUSATIONS AND INTERNATIONAL LAW IN GLOBAL CYBERSECURITY

Martha Finnemore<sup>\*</sup> and Duncan B. Hollis<sup>\*\*</sup>

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, the fear of reputational harm often kept them from disclosing it. In either case, cyberspace's technical architecture meant those responsible for a cyber operation could often remain anonymous.<sup>1</sup> 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, the conventional wisdom held that this "attribution problem" posed serious—and perhaps insuperable—obstacles to engendering compliance by States with any rules in cyberspace.<sup>2</sup>

Times have changed.<sup>3</sup> Over the last decade, several States—including China, Iran, North Korea, Russia, the United Kingdom, and the United States—stand accused of conducting or supporting cyber operations with

<sup>\*</sup> University Professor of Political Science and International Affairs, George Washington University. Professor Finnemore would like to thank Amoz Hor for his excellent research assistance in the preparation of this paper.

<sup>&</sup>lt;sup>\*\*</sup> Professor of Law, Temple University Beasley School of Law. 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.

<sup>&</sup>lt;sup>1</sup> See Jon R. Lindsay, *Tipping the scales: the attribution problem and the feasibility of deterrence against cyberattack*, 1 J. CYBERSECURITY 53, 54 (2015).

<sup>&</sup>lt;sup>2</sup> See, e.g., Steve Ranger, US Intelligence: 30 Countries Building Cyber Attack Capabilities, ZD Net, Jan. 5, 2017; P.W. SINGER AND ALLEN FRIEDMAN, CYBERSECURITY AND CYBERWAR 73 (OUP, 2014); Larry Greenemeier, <u>Seeking Address: Why Cyber Attacks Are So Difficult</u> to Trace Back to Hackers, SCI. AM. (June 11, 2011); Martha Finnemore & Duncan B. Hollis, *Constructing Norms for Global Cybersecurity*, 110 AM. J. INT'L L. 425, 435-36 (2016).

<sup>&</sup>lt;sup>3</sup> It is not, however, obvious *why* things changed. Certainly, technology evolved to allow some State (and non-State) actors greater visibility into a cyber-attack's origins. *See, e.g.,* Benjamin Edwards et al. *Strategic aspects of cyberattack, attribution, and blame,* 114 Proc. Nat'l Acad. Sciences 2825 (March 14, 2017) ("Sources in or close to the US Government assert that its ability to trace back a cyber operation to its geographic origin (e.g., an urban neighborhood in China) is excellent"); John S. Davis II et al, *Stateless Attribution: Towards International Accountability for Cyberspace* 2 (Rand Corp. 2017). It is less clear, why victims (or victim States) began making their accusations publicly.

serious impacts on governments, peoples, and resources.<sup>4</sup> Today, there are public allegations of more than 20 State or State-sponsored cyber operations.<sup>5</sup> Accusations "naming" a State and its cyber operation(s) come from a variety of sources, including private cybersecurity firms and academic institutions. Even certain States have demonstrated an increased willingness to "name names."<sup>6</sup>

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 or decline to comment.<sup>7</sup> They also appear willing to continue engaging in the same or similar behavior.<sup>8</sup> China's agreement to forgo commercial cyber espionage activities might be an exception. China entered into a political commitment disavowing the practice after an extended naming and shaming campaign, including the U.S. indictment of five officers of the People's Liberation Army.<sup>9</sup> Yet, countervailing evidence suggests that China's changed stance actually derived from domestic politics.<sup>10</sup> Recent reports suggest in any case that China has resumed its commercial cyberespionage operations.<sup>11</sup> As such, there is

<sup>&</sup>lt;sup>4</sup> We define a "cyber operation" as the use of information and communication technologies to generate significant losses of confidentiality, integrity, and/or access in a computer system or network. Mohammad Nazmul Alam et al, *Security Engineering Towards Building a Secure Software*, 81 INT'L J. COMP. APPLICATIONS 32, 33–34 (2013). To capture the full range of accusations against States, our definition includes both cases of cyber-espionage—in which ICTs supplant more traditional spying tools—as well as more novel forms of cyber-attack that degrade, disrupt, or damage a computer system and (perhaps) the infrastructure it supports.

<sup>&</sup>lt;sup>5</sup> See Council of Foreign Relations, *Cyber Operations Tracker*, at <u>https://www.cfr.org/interactive/cyber-operations;</u> see also Center for Strategic & International Studies, *Significant Cyber Incidents*, at <u>https://www.csis.org/programs/cybersecurity-and-governance/technology-policy-program/other-projects-cybersecurity;</u> Davis II et al, *supra* note 3; Dan Efrony and Yuval Shany, *A Rule Book on the Shelf? Tallinn Manual 2.0 on Cyber Operations and Subsequent State Practice*, 112 AM. J. INT'L L. (forthcoming 2018).

<sup>&</sup>lt;sup>6</sup> See, e.g., Tim Starks, *Trump administration ratchets up 'naming and shaming' nation-state hackers*, POLITICO, June 6, 2018.

<sup>&</sup>lt;sup>7</sup> See, e.g., Davis et al, *supra* note 3, at 2; Thomas Grove, *Russian Agency at Center of U.S. Hacking Indictment Has Long Operated in the Shadows*, WALL ST. J. (July 14, 2018) ("Russia denies it had attempted to influence the U.S. elections or was behind the hacking of the DNC.").

<sup>&</sup>lt;sup>8</sup> See, e.g., Jack Goldsmith, Uncomfortable Questions in the Wake of Russia Indictment 2.0 and Trump's Press Conference With Putin, LAWFARE (July 16, 2018).

<sup>&</sup>lt;sup>9</sup> Dep't of Justice, <u>U.S. Charges Five Chinese Military Hackers with Cyber Espionage against</u> <u>U.S. Corporations and a Labor Organization for Commercial Advantage</u> (May 19, 2014).

<sup>&</sup>lt;sup>10</sup> See, e.g., FIREEYE, RED LINE DRAWN: CHINA RECALCULATES ITS USE OF CYBER ESPIONAGE (June 2016) (sourcing the changed stance to President Xi's desire to reign in freelance operations by Chinese government agencies).

<sup>&</sup>lt;sup>11</sup> See, e.g., Claude Barfield, <u>*Renewed Chinese cyberespionage: Time for the US to act*</u>, AMERICAN ENTERPRISE INSTITUTE (April 16, 2018).

widespread skepticism about the capacity of naming and shaming to change an accused's behavior in cyberspace.<sup>12</sup>

For international lawyers, the recent spate of accusations is troubling for a different reason—the absence of international law in any of these accusations. In other contexts (e.g., human rights, the environment), naming and shaming efforts were explicitly tied to violations of treaty provisions or other international law rules.<sup>13</sup> When it comes to State-sponsored cyber operations, however, the accusations have studiously avoided invoking international law, let alone assessing if these operations comport with its rules. Cyber operations are simply labeled as malicious, as irresponsible, or violations of "international norms."<sup>14</sup> Thus, 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 . . ."<sup>15</sup>

This reluctance to invoke international law might suggest that law is weak or—worse—irrelevant in holding State actors accountable for their cyber operations.<sup>16</sup> We believe, however, such concerns risk missing the forest for the trees. Focusing on law's absence in accusations risks missing larger effects accusations may have on both compliance, and on law, itself. Certainly, social science suggests accusations *can* change an accused's behavior.<sup>17</sup> International relations scholars have spent years exploring the conditions under which naming and shaming may be effective in doing so.<sup>18</sup> In the same

<sup>&</sup>lt;sup>12</sup> See, e.g., Starks, supra note 6; Jack Goldsmith and Stuart Russell, Strengths Become Vulnerabilities How a Digital World Disadvantages the United States in Its International Relations, Aegis Series Paper No. 1806 (Hoover Institution, 2018), pp.13-14.

<sup>&</sup>lt;sup>13</sup> See, e.g., Human Rights Watch, *Egypt: Al-Sisi Should End Rights Abuses* (April 10, 2018) (disclosing irregularities in 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 People's Rights").

<sup>&</sup>lt;sup>14</sup> See Press Releases, U.S. Department of the Treasury, <u>Treasury Sanctions Russian Cyber</u> <u>Actors for Interference with the 2016 US Elections and Malicious Cyber-Attacks</u> (Mar. 15, 2018); Press Release, the President at The White House, <u>Statement on Actions in Response to</u> <u>Russian Malicious Cyber Activity and Harassment</u> (Dec. 29, 2016); Press Statement, John Kerry, Secretary of State, US Dep't of State, Condemning Cyber-Attacks by North Korea (Dec. 19, 2014) (describing 2014 Sony hack as a violation of "international norms").

<sup>&</sup>lt;sup>15</sup> Efrony and Shany, *supra* note 5, at 73.

<sup>&</sup>lt;sup>16</sup> If there are references to law, they usually involve domestic legal standards like the U.S. indictments of foreign government agents for participating in various cyber operations. *See, e.g.*, Press Release, <u>U.S. 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 (March 24, 2016); Mark Mazzetti & Katie Benner, *12 Russian Agents Indicted in Mueller Investigation*, N. Y. TIMES (July 13, 2018), U.S. Charges Five Chinese Military Hackers, *supra* note 9.</u>

<sup>&</sup>lt;sup>17</sup> See Ray Pawson, Evidence and Policy in Naming and Shaming, 23 POL'Y STUD. 211 (2002).

<sup>&</sup>lt;sup>18</sup> See, e.g., Mathrew Krain, J'Accuse! Does Naming and Shaming Perpetrators Reduce the Severity of Genocides or Politicides? 56 INT'L STUD. Q. 574 (2012); James Franklin, Shame on You: The Impact of Human Rights Criticism on Political Repression in Latin

vein, naming and shaming can improve international law compliance; accusations of international law violations—most often in the human rights context—have led certain accused States to conform, or at least reduce, the extent of their deviance from international law.<sup>19</sup> Such successes have led States and scholars to perceive naming and shaming as a single concept with a unitary function—"shaming" a "named" State into changing its unwanted behavior.<sup>20</sup>

It is a mistake, however, to lump all accusations under this "naming and shaming" rubric. In both functions and contents, accusations involve a more varied and dynamic set of processes than contemplated by the existing naming and shaming literature. Of course, accusations can engender *compliance* by leading the accused to cease unwanted acts. But accusations may also serve *defensive*, *deterrent* or *punitive* purposes. Most importantly, accusations may play a *constitutive* role, constructing new norms, including customary international law.

This variation in the possible functions of an accusation is matched by variation in the contents of accusations themselves. For us, accusations comprise not one or two, but three discrete processes:

- (i) *attribution* (the process of associating what happened with a particular author or territory);
- (ii) *exposure* (the process of disclosing what happened to third parties);
- (iii) *condemnation* (the process of signaling disapproval of what happened).<sup>21</sup>

*America*, 52 INT'L STUD. Q. 187, -204-07 (2008); Emilie M. Hafner-Burton, *Sticks and Stones: Naming and Shaming the Human Rights Enforcement Problem*, 62 INT'L ORG. 689 (2008).

<sup>&</sup>lt;sup>19</sup> See, e.g., Krain, *supra* note 18; Hafner-Burton, *supra* note 18 (noting that naming and shaming works for certain types of human rights violations, but not others); *see generally* BETH SIMMONS, MOBILIZING HUMAN RIGHTS: INTERNATIONAL LAW IN DOMESTIC POLITICS (2009).

<sup>&</sup>lt;sup>20</sup> Despite the conjunctive terminology, most scholarship in international relations conceives of naming and shaming as a unitary mechanism. *See, e.g.,* H. Richard Friman, *Introduction: Unpacking the Mobilization of Shame,* in THE POLITICS OF LEVERAGE 3 (H. Friman, ed., 2015) ("unpacking naming and shaming" by examining what "exactly *the* concept means") (emphasis added). It is, moreover, often defined in terms of its capacity to alter the accused's behavior. *See* Molly Beutz Land, *Networked Activism,* 22 HARV. HUM. RTS. J. 205, 208 (2009) (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.").

<sup>&</sup>lt;sup>21</sup> We are not the first to identify naming and shaming's discrete processes. *See, e.g.*, Faradj Koliev, *Book Review, The politics of leverage in international relations: name, shame, and sanction, Edited by H. Richard Friman*, 91 INT'L AFF. 1168, 1169 (2015) (praising the volume for "its conceptual distinction between public exposure (naming) and public condemnation (shaming)"). For his part, Friman describes the phenomenon in terms of "[p]ublic exposure and condemnation." Friman, *supra* note 20, at 5, 203. As discussed in Part II *infra*, however, we do not view accusations to require exposure and also believe attribution is a separate potential component of accusations.

Accusations can encompass all three processes, as when the United States accused the Russian Federation of interference in its 2016 presidential election.<sup>22</sup> Other accusations may feature only two elements. Accusers can choose to attribute and condemn what happened without exposing it—i.e., making their accusation via private or diplomatic channels. Or, accusers can expose and condemn what happened without disclosing (or even knowing) to whom it may be attributed. Accusations can even expose an attribution without explicitly condemning it; in other words, there can be naming without shaming.

In this essay, we identify and explore the concept of accusations in cybersecurity, with particular attention to their role in international law. We do so in four parts. We begin by examining the different functions accusations may serve based on the cyber accusations made to date. Second, we identify different components of an accusation and how they may be constructed. Third, we look beyond the accusation's contents to identify external conditions that may impact its efficacy. We hypothesize, for example, that the conditions for constructing a norm from an accusation need not fully align with those needed to change the accused's behavior. Fourth, and finally, we examine the implications of accusation dynamics in cybersecurity for international law. We offer some hypotheses about why accusations regarding cyber operations have yet to include international legal condemnations and suggest several concrete steps for improving their utility.

Taken together, our essay offers a broader and more nuanced assessment of the utility of accusations for global cybersecurity than those who have examined naming and shaming to date. For international relations scholars, we hope to inspire further research on how the various functions and components of accusations created varied political effects in different contexts. For international lawyers, the cybersecurity context provides a valuable case-study of how international law may be constituted in the shadows. Finally, we aim to provide States and other stakeholders a more accurate and detailed map for when and how to employ accusations to various ends, including the construction of customary international law.

#### I. WHAT CAN ACCUSATIONS ACHIEVE?

Accusations are a regular feature of all social interactions. A parent may accuse her child of causing a sibling to cry; an NGO may accuse a company of using child labor; shareholders may accuse CEOs of mismanagement. In the context of global cybersecurity, we define an accusation as the process by which one or more actors claim that a State bears

<sup>&</sup>lt;sup>22</sup> See Press Release, Joint Statement, Department of Homeland Security & Office of the Director of National Intelligence (ODNI), <u>Election Security</u> (Oct. 7, 2016); Intelligence Community Assessment (ICA), <u>Assessing Russian Activities and Intentions in Recent US Elections</u> (Jan. 6, 2017).

responsibility for a cyber incident or operation.<sup>23</sup> What purpose do such accusations serve? We believe there are at least five different reasons an accuser may deploy an accusation: (i) compliance; (ii) defense; (iii) deterrence; (iv) punishment; or (v) constitution.<sup>24</sup> 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.

*Compliance* is the function most often associated with "naming and shaming" in the extant literature. Accusations often seek to have an accused "comply" with the accuser's behavioral expectations, whether by altering its ongoing behavior or avoiding undesired behavior in the future. 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.<sup>25</sup> Companies engaged in questionable financial practices may not welcome public scrutiny.<sup>26</sup> Human rights violating governments usually prefer to torture and "disappear" their opponents in secret.<sup>27</sup> Public exposure or revelation of the bad behavior ("naming") will create reputational damage and/or moral discomfort ("shaming") in the bad actor thereby inducing a change in that behavior.

The compliance 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 accusing North Korea of responsibility for the Sony Pictures hack and of subsequent U.S. charges and sanctions

<sup>&</sup>lt;sup>23</sup> See supra note 4 (defining "cyber operation"); Herbert Lin, Attribution of Malicious Cyber Incidents: From Soup to Nuts, Aegis Paper Series, No. 1607 (Hoover Institution, 2016), p. 5. Accusations in international relations may, of course, have a broader ambit. They can include other subject-matter beyond cyber-space. Accusations may also target other categories of actors, including international organizations, insurgent groups, multi-national enterprises, or transnational civil society organizations. Given our focus on international law, we have limited our attention to accusations where the accused is a State or a non-State actor for which a State may have responsibility.

<sup>&</sup>lt;sup>24</sup> This is not an exhaustive list. Private cyber-security companies, for example, may find that making accusations offers a tangible reward. Credible accusations by cybersecurity companies may boost client sales or profitability, as Mandiant's financial success after accusing China in its APT1 report shows. *See* Jim Finkel, *Mandiant goes viral after China Hacking report*, REUTERS, Feb. 22, 2013.

<sup>&</sup>lt;sup>25</sup> JAMES T. HAMILTON, REGULATION THROUGH REVELATION: THE ORIGIN, POLITICS, AND IMPACTS OF THE TOXICS RELEASE INVENTORY PROGRAM (CUP, 2005); James T. Hamilton, *Pollution as news: Media and stock market reactions to the toxics release inventory data*, 28 J. ENV. ECON. & MGMT. 98 (1995).

<sup>&</sup>lt;sup>26</sup> Judith van Erp, *Naming without Shaming: the publication of sanctions in the Dutch financial market*, 5 REG. & GOVERNANCE 287 (2011).

<sup>&</sup>lt;sup>27</sup> Amanda Murdie and Dursun Peksen, *Women's rights INGO shaming and the government respect for women's rights*, 10 REV. INT'L ORG.1 (2015); Amanda M. Murdie and David R. Davis, *Shaming and blaming: Using events data to assess the impact of human rights INGOs*, 56 INT'L STUD.Q. 1 (2012); Hafner-Burton, *supra* note 18.

against a named Pyongyang operative.<sup>28</sup> Indictments of specific Chinese and Iranian individuals affiliated with their respective governments had a similar purpose, especially in the absence of mechanisms to bring them before U.S. courts.<sup>29</sup> The Trump Administration has recently touted its "naming and shaming" strategy. In describing the increasing number of U.S. accusations against State-sponsored cyber operations, Jeanette Manfra, 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.<sup>30</sup>

Secretary Manfra, however, also articulated a second function that accusations can serve: *defense*.<sup>31</sup> Simply put, the accused may not be the only audience for an accusation. Accusations provide information on what happened that can have great utility to third parties. This is especially true for cybersecurity where an accusation "may encourage victims or other vulnerable populations to bolster network defenses."<sup>32</sup> 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.<sup>33</sup> Similar defensively-orientated 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.<sup>34</sup>

<sup>&</sup>lt;sup>28</sup> Press Release, The White House, <u>Statement by the Press Secretary on the Executive Order</u>, <u>Imposing Additional Sanctions with Respect to North Korea</u> (Jan. 2, 2015); Sean Sullivan, Obama: North Korea hack 'cyber-vandalism,' not 'act of war, WASH. POST (Dec. 21, 2014); Ellen Nakashima and Devlin Barrett, Justice Department announces charges against North Korean operative for Sony Pictures hack, WASH. POST (Sept. 6, 2018).

<sup>&</sup>lt;sup>29</sup> See, e.g., Seven Iranians Working for Islamic Revolutionary Guard Corps-Affiliated Entities Charged, supra note 16; U.S. Charges Five Chinese Military Hackers, supra note 9.

<sup>&</sup>lt;sup>30</sup> Starks, *supra* note 6.

<sup>&</sup>lt;sup>31</sup> *Id.* (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").

<sup>&</sup>lt;sup>32</sup> Davis et al., *supra* note 3, at 17.

<sup>&</sup>lt;sup>33</sup> Blake Johnson et al, <u>Attackers Deploy New ICS Attack Framework "TRITON" and cause</u> <u>Operational Disruption to Critical Infrastructure</u>, FIREEYE BLOG, Dec. 14, 2017; see also Chris Bing, <u>Trisis has the security world spooked</u>, <u>stumped and searching for answers</u>, CYBERSCOOP, Jan. 16, 2018.

<sup>&</sup>lt;sup>34</sup> See, e.g., ICS-CERT, <u>Cyber-Attack Against Ukrainian Critical Infrastructure</u> (Feb. 26, 2016); Joint Analysis Report, National Cybersecurity and Communications Integration Center's (NCCIC) and Federal Bureau of Investigation, *Grizzly Steppe – Russian Malicious Cyber Activity*, Ref. No. JAR-16-20296A (Dec. 29, 2016).

Accusations may not only seek to assist third parties in defenses, they may also seek to deter potential perpetrators as well. Accusers may expose a State's cyber operations to signal to other States that they cannot engage in similar behavior without public attention. Cyber operations are often attractive to States precisely because States think that they can be deployed anonymously-i.e., either the operation proceeds undetected, or the State can keep its own role unclear, or even have another State or non-State party take the blame (a "false flag operation"). Accusers who identify the State(s) responsible for a cyber operation may disrupt such expectations, signaling to States that they cannot automatically expect to operate unobserved. That fact may, in turn, effect the cost-benefit calculus of States contemplating cyber operations; in some cases, it could deter them from acting at all. Deterrence was 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 being responsible for launching the NotPetya ransomware.<sup>35</sup> Deterrence is also often a motive in accusations discrediting false flag cyber-operations, including reports that Russia-not ISISconducted a cyber-attack knocking TV5Monde off the air in France and that Russia-not North Korea-disrupted the information infrastructure associated with the 2018 Winter Olympic Games.<sup>36</sup>

Separate from any deterrent function, some accusations may have a *punitive* purpose. Instead of imposing social pressure for others to conform or comply with the accuser's expectations, accusations may serve as building blocks in a strategy to punish. Accusers may, for example, issue accusations to "persuade a set of third-party actors to generate support for sanctions."<sup>37</sup> Accusations are also required to deploy domestic criminal penalties. Since first indicting the five PLA officers, the United States has pursued indictments with increased frequency. And, although most of the accused have escaped punishment, the United States did arrest a Chinese national in 2017 on charges of participating in the OPM hack.<sup>38</sup>

<sup>&</sup>lt;sup>35</sup> See, e.g., The White House, <u>Statement from the Press Secretary</u> (Feb.15, 2018); see also Stilgherrian, Blaming Russia for NotPetya was coordinated diplomatic action, ZDNet, April 12, 2018. Ukraine also blamed Russia. <u>SBU establishes involvement of the RF special</u> <u>services into Petya. A virus-extorter attack</u>, SBU PRESS-CENTER (July 1, 2017). NotPetya was a ransomware attack that experts suggest was 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 Conner Forrest, NotPetya ransomware outbreak cost Merck more than \$300M per quarter, TECHREPUBLIC, Oct. 30, 2017.

<sup>&</sup>lt;sup>36</sup> Sheera Frenkel, *Experts Say Russians May Have Posed As ISIS To Hack French TV Channel*, BUZZFEED, June 9, 2015 (discussing FireEye report accusing Russia of responsibility for the TV5Monde hack); Ellen Nakishima, *Russian spies hacked the Olympics and tried to make it look like North Korea did it, U.S. officials say*, WASH. POST, Feb. 24, 2018.

<sup>&</sup>lt;sup>37</sup> Davis et al., *supra* note 3, at 17.

<sup>&</sup>lt;sup>38</sup> Joseph Menn, *Chinese National Arrested in Los Angeles on U.S. hacking charge*, REUTERS, Aug. 24, 2017.

Accusations may open the door to punishments for their own sake or to facilitate restitution. But accusations may also be central to strategies for improved compliance by the accused. International law provides two vehicles for obtaining reparations and the cessation of wrongful behavior. Acts of retorsion are unfriendly—but lawful—acts (e.g., the expulsion of diplomats) designed to respond to an unlawful act.<sup>39</sup> Counter-measures are non-forceful acts—which are themselves illegal—but which international law permits when conducted by a State in response to another State's prior wrongful act(s).<sup>40</sup> For a State to engage in either retorsion or counter-measures, however, requires some accusation articulating the wrongful acts that trigger the accuser's right of response.<sup>41</sup>

Finally, accusations may be *constitutive*. In many cases, an accusation "sends a public message about correct and appropriate behavior."<sup>42</sup> In the human rights context, accusations often involve well-established legal norms of behavior (e.g., the prohibitions on torture or genocide; freedom of expression; religious freedom) against which the accused's behavior is measured.<sup>43</sup> 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 where an 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 behavior.

Particularly important in the cybersecurity context, accusations may play a key role in constructing new norms. The most prominent cyber operations (Estonia, Stuxnet, WannaCry) are defined by their novelty; they do things never seen before or on a scale not previously thought possible.<sup>44</sup> Thus,

<sup>&</sup>lt;sup>39</sup> Rebecca Crootof, *International Cybertorts: Expanding State Accountability in Cyberspace*, 103 CORNELL L. REV. 565, 579 (2018).

<sup>&</sup>lt;sup>40</sup> See ILC, "Draft Articles on the Responsibility of States for Internationally Wrongful Acts" in *Report on the Work of its Fifty-first Session* (3 May-23 July, 1999), UN Doc A/56/10 55 Art. 22 ("ASR")

<sup>&</sup>lt;sup>41</sup> *Id.*, Art. 43 ("An injured State which invokes the responsibility of another State shall give notice of its claim to that State."); Art. 52 ("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 also measured in part by its proportionality to the originally wrongful act. *Id.* at Art. 51.

<sup>&</sup>lt;sup>42</sup> van Erp, *supra* note 26.

 <sup>&</sup>lt;sup>43</sup> See, e.g., International Covenant on Civil and Political Rights, Dec. 16, 1966, 999 U.N.T.S.
171; Convention on the Prevention and Punishment of the Crime of Genocide, Dec. 9 1948, 78 U.N.T.S. 277.

<sup>&</sup>lt;sup>44</sup> See, e.g., KIM ZETTER, COUNTDOWN TO ZERO DAY: STUXNET AND THE LAUNCH OF THE WORLD'S FIRST DIGITAL WEAPON (2014); Chris Graham, *NHS Cyber attack: Everything you need to know about 'biggest ransomware' offensive in history*, THE TELEGRAPH, May 20, 2017.

it is often unclear if *any* norm exists to govern States engaging in these operations.<sup>45</sup> In such cases, an accusation serves as an opening bid 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. Accusations may thus lay out the contours of the "bad behavior" along with an argument about why, exactly, the behavior is undesirable. Other actors may then respond to the accusation. They may accept some of it; they may accept all of it; or they may accept it in some situations but not others. Or, they may reject it. It is these interactions between the accuser, the accused, and third parties in the larger community that, over time, may result in the creation of a new norm (or its failure).<sup>46</sup>

The United States has employed such a constructive strategy in suggesting that certain cyber operations (e.g., the Sony Hack, 2016 election interference) violated "established international norms."<sup>47</sup> Ambiguity in the U.S. statements leaves open which norms it believes were violated, however, and the accused have denied the U.S. charges.<sup>48</sup> Nonetheless, the U.S. accusations also served as an invitation to other like-minded States to express similar views on the appropriate norms of behavior. In the case of U.S. accusations about election interference, for example, Foreign and Security Ministers from the G7 subsequently issued a joint statement denouncing foreign attempts to interfere in democratic processes, including "through cyber-enabled activities."<sup>49</sup>

There is no reason that cyber accusations could not feature international law and build out legal norms in similar ways. Although States have not done so, several scholars have made the effort to examine accusations of cyber operations such as WannaCry and the 2016 election interference in terms of the existing rules of international law (e.g., the prohibition on the use of force, the duty of non-intervention, sovereignty, self-determination, human rights).<sup>50</sup> Moreover, it would be a mistake to assume that a State's silence on

<sup>&</sup>lt;sup>45</sup> See Peter J. Katzenstein, *Introduction: Alternative Perspectives on National Security*, in THE CULTURE OF NATIONAL SECURITY: NORMS AND IDENTITY IN WORLD POLITICS 1, 5 (Peter J. Katzenstein ed., 1996) (defining norms as "collective expectations for the proper behavior of actors with a given identity").

<sup>&</sup>lt;sup>46</sup> Finnemore and Hollis, *supra* note 2, at 475-477.

<sup>&</sup>lt;sup>47</sup> See White House, Statement on Actions in Response to Russian Malicious Cyber Activity and Harassment, supra note 14 (opposing "Russia's efforts to undermine established international norms of behavior and interfere with democratic governance"); Condemning Cyber-Attacks by North Korea, supra note 14 (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").

<sup>&</sup>lt;sup>48</sup> See, e.g., Davis et al, *supra* note 3, at 2; Grove, *supra* note 7.

<sup>&</sup>lt;sup>49</sup> G7, Joint Statement of Foreign and Security Ministers, <u>Defending Democracy: Addressing</u> <u>Foreign Threats</u> (June 2018).

<sup>&</sup>lt;sup>50</sup> See, e.g., Michael Schmitt and Sean Fahey, WannaCry and the International Law of Cyberspace, JUST SECURITY, Dec. 22, 2017; Jens David Ohlin, Did Russian Cyber Interference in the 2016 Election Violate International Law?, 95 TEXAS L. REV. 1579 (2017).

the international legal implications of its accusation means that the accusation has none. Customary international law rarely, if ever, emerges immediately and fully formed. Rather, it is the product of interactions and iterations over time that eventually reach a point where a sufficiently uniform practice is generally (although not universally) accepted as *opinio juris* (i.e., recognized as being legally obligatory). <sup>51</sup> 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 can also serve as evidence of *opinio juris*, delineating which acts are wrongful as a matter of international law.

Whether or not accusations construct or elaborate specific international law prohibitions, they may play an important role in defining what behavior international law permits.<sup>52</sup> By objecting and making accusations of wrong-doing, States and other actors can limit the potential for the accused's behavior to become legally accepted. The International Law Commission emphasized this point in its most recent *Draft Conclusions on Identifying Customary International Law*, noting how a failure to react can constitute evidence that such behavior is lawful.<sup>53</sup> In other words, "toleration of a certain practice may indeed serve as evidence of acceptance as law (*opinio juris*) when it represents concurrence in that practice."<sup>54</sup> Thus, whether or not States currently characterize their cyber accusations in explicitly legal terms, they are signaling that they disapprove of certain cyber acts, and these accusations may counter-act claims that the accused States' operations are (or are becoming) permitted by international law.<sup>55</sup>

<sup>&</sup>lt;sup>51</sup> 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.,* George Norman & Joel P. Trachtman, *The Customary International Law Game,* 99 AJIL 541, 542 (2017); Andrew T. Guzman, *Saving Customary Int'l Law,* 27 MICH. J. INT'L L. 115, 122 (2005); Anthea Roberts, *Traditional and Modern Approaches to Customary International Law: A Reconciliation,* 95 AJIL 757, 757-758 (2002).

<sup>&</sup>lt;sup>52</sup> 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" (Fr. v. Turk.), 1927 P.C.I.J. (ser. A) No. 10 (Sept. 17), 18-19 (Given the "very nature and existing conditions of international law . . . [r]estrictions upon the independence of States cannot therefore be presumed" and finding "all that can be required of a State is that it should not overstep the limits which international law places upon its jurisdiction.").

<sup>&</sup>lt;sup>53</sup> See ILC, Draft Conclusions on Identification of Customary International Law, U.N. Doc. A/CN.4/L.908 (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"); ILC, Draft Conclusions on Identification of Customary International Law U.N. Doc. A/71/10 (2016), pp. 100-101 (Commentary on Conclusion 10).

<sup>&</sup>lt;sup>54</sup> ILC, Commentary on Draft Conclusions, *supra* note 53, at 100-101.

<sup>&</sup>lt;sup>55</sup> Alternatively, if other States do accept or acquiesce in the legality of certain State or Statesponsored cyber operations, the accusing State may be able to employ its accusation to claim the status of a persistent objector. *See* JAMES CRAWFORD, BROWNLIE'S PUBLIC INTERNATIONAL LAW (8<sup>th</sup> ed. 2012), p. 28.

### II. DISAGGREGATING ACCUSATIONS: ATTRIBUTION, EXPOSURE, CONDEMNATION

Successful accusations require knowledge of the facts or events that prompted them. Such information is not always cheap or easy to obtain, but in the contemporary information environment, assembling the corroborating details of malicious activity can be a widely available, non-violent, and, at least in democracies, legal tool for an array of savvy actors seeking to curb bad behavior 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: (a) attribution; (b) exposure; (c) condemnation. We explain how each of these operate in the cyber context below before exploring how they may be constructed into an accusation.

#### A. Attribution

Attribution is the process of answering the age-old question of "who did what exactly."<sup>56</sup> 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.<sup>57</sup>

For our purposes, attribution is the assignment of responsibility for a cyber operation. Unlike physical and static identifiers used in other contexts (e.g., DNA, fingerprints), digital attribution involves very different technical indicators and patterns that may complicate the process.<sup>58</sup> Much of the cybersecurity literature focuses extensively on these technical aspects of attributing responsibility for cyber incidents.<sup>59</sup> Yet, as Herb Lin emphasizes, cyber attributions may require more than a technical process depending on the goal. 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) ultimately responsible for the incident.<sup>60</sup> The latter two efforts will usually require other (or "all") sources of intelligence beyond technical indicators pointing to a particular IP address or network.

<sup>&</sup>lt;sup>56</sup> See Thomas Rid and Ben Buchanan, Attributing Cyber-Attacks, 38 J. STRAT. STUD. 4, 4 (2014).

<sup>&</sup>lt;sup>57</sup> See, e.g., Nicaragua v. United States, 1986 I.C.J. 14, 18 (1986) (holding US responsible for mining Nicaraguan harbors); Shaun Walker, *MH17 Downed by Russian Military Missile System, say Investigators*, THE GUARDIAN (May 24, 2018); Department of Justice and Constitutional Development of the Republic of South Africa, *The Truth and Reconciliation Official Website, at* www.justice.gov.za (accessed Sept. 6, 2018).

<sup>&</sup>lt;sup>58</sup> See Davis et al, supra note 3 at 9-10.

<sup>&</sup>lt;sup>59</sup> See, e.g., David Wheeler and Gregory Larsen, *Techniques for Cyber Attack Attribution*, (Institute for Defense Analyses, Oct. 2003).

<sup>&</sup>lt;sup>60</sup> Lin, *supra* note 23, at 8-19; Davis et al, *supra* note 3, at 9.

Whatever the goal, attributions can vary in terms of certainty and precision. As Rid and Buchanan explain, cyber attributions are not binary—where attribution is possible/impossible—but are situated along a spectrum.<sup>61</sup> Thus, when the University of Toronto's Citizen Lab uncovered the "Ghostnet" cyber espionage network targeting Tibetan institutions, its analysis "circumstantially point[ed] to China as the culprit" but never formally named "the identity of the attacker(s)."<sup>62</sup> In contrast, the U.K. Foreign Ministry indicated that it was "highly likely" that "North Korean actors known as the Lazarus Group were behind the WannaCry ransomware campaign."<sup>63</sup> The U.S. cybersecurity company Mandiant concluded that Unit 61398 of China's People's Liberation Army was the source of a long-standing commercial cyber espionage campaign, barring

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.<sup>64</sup>

It is possible, moreover, to have attribution at one level (e.g., to a machine, to a person, to a State) but not others. Jason Healey, for example, highlights how it is possible to attribute responsibility for a cyber operation to a particular State even without evidence permitting attribution to particular individuals.<sup>65</sup>

Attributions may also vary in their specificity. An attribution can be highly precise, identifying specific individual(s) associated with perpetuating a cyber operation. The United States has, for example, issued indictments that attribute responsibility for U.S. election interference to more than a dozen named Russian intelligence operatives.<sup>66</sup> Many cybersecurity firms attribute responsibility to known "groups" bearing diverse monikers (e.g., Strontium, Fancy Bear, Cozy Bear, Lazarus) that may have some affiliation with a State.<sup>67</sup> In other cases, an attribution may only indicate the territorial origin of a cyber incident without actually identifying a responsible individual, group, or State.<sup>68</sup>

<sup>&</sup>lt;sup>61</sup> Rid and Buchanan, *supra* note 56, at 7.

<sup>&</sup>lt;sup>62</sup> INFORMATION WARFARE MONITOR, TRACKING GHOSTNET: INVESTIGATING A CYBER ESPIONAGE NETWORK 12–13 (2009).

<sup>&</sup>lt;sup>63</sup> Press Release, Foreign and Commonwealth Office and Lord Ahmad of Wimbledon, <u>Foreign</u> <u>Office Minister condemns North Korean actor for WannaCry attacks</u> (Dec.19, 2017).

<sup>&</sup>lt;sup>64</sup> Mandiant, *APT1: Exposing One of China's Cyber Espionage Units* (2013), available at <u>https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf</u>.

<sup>&</sup>lt;sup>65</sup> See Jason Healey, Beyond Attribution: Seeking National Responsibility for Cyber Attacks, Issue Brief, Atlantic Council (Jan. 2012).

<sup>&</sup>lt;sup>66</sup> See Mazzetti & Benner, supra note 16.

<sup>&</sup>lt;sup>67</sup> See, e.g., Who is FANCY BEAR, CROWDSTRIKE BLOG, Sept. 12, 2016, Microsoft Security Intelligence Report: Strontium, MICROSOFT SECURE, Nov. 16, 2015.

<sup>&</sup>lt;sup>68</sup> The scope of attributions may also vary from those focused on a single incident to larger patterns of overall conduct. As an example of the latter, 2012 media reports suggested that

2018

Actors making cyber accusations that contain attributions must, therefore, assess the level of certainty and specificity they will convey. Of course, that level may be a function of the investigation itself—the accuser may only have limited certainty and/or specificity about a cyber operation's author(s). But attributions may not align with the accuser's actual knowledge. Accusers can always attribute with less certainty or specificity than their actual knowledge, especially if protecting sources and methods is useful. Alternatively, they can make inferences (or mistakes) that result in attributions beyond what their own evidence suggests.

#### **B.** Exposure

Exposure refers to the publicity an accusation receives. Some accusations never see the light of day, but are communicated privately between the accuser and the accused.<sup>69</sup> 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 claim of Russian responsibility for the 2007 directed denial of service attacks against its systems to the U.S, U.K. and Australian accusations that North Korea launched WannaCry.<sup>70</sup>

The existence of private or semi-private accusations is harder to discern. Still, we believe both types are not only possible, but likely, vehicles for accusations in cyberspace. Transnational technical communities (e.g., FIRST) or industry collectives (e.g., the Cybersecurity Tech Accord) certainly have information that could underpin accusations but have good reasons not to share these publicly.<sup>71</sup> Similarly, where one State believes another bears responsibility for a cyber incident, the first may prefer to convey the accusation privately through diplomatic channels or other means. Such private accusations may be the only move or a first step in an escalatory ladder. After formally accusing Russia of complicity in using cyber means to interfere in the 2016 U.S. presidential election, for example, President Obama revealed that he had first privately conveyed the accusation to President Putin directly.<sup>72</sup>

U.S. officials attributed responsibility to Iran for a series of cyber incidents against U.S. financial institutions from 2011-13. Nicole Perlroth and Quentin Hardy, *Bank Hacking was the Work of Iranians, Officials Say*, N.Y. TIMES, Jan. 8, 2013.

<sup>&</sup>lt;sup>69</sup> Much of the work of the International Committee of the Red Cross operates this way. *See, e.g.*, Laura MacInnis, *International Red Cross issues rare Myanmar censure*, REUTERS (June 29, 2007) (noting that the organization "normally deals under a cloak of confidentiality.").

<sup>&</sup>lt;sup>70</sup> See, e.g., Claire Bickers, UK and US blame 'WannaCry' cyber-attack on North Korea, REUTERS, Dec. 20, 2017; Newly Nasty, THE ECONOMIST, May 24, 2007; Estonian Links Moscow to Internet Attack, NEW YORK TIMES, May 18, 2007.

<sup>&</sup>lt;sup>71</sup> See, e.g., The Forum of Incident Response and Security Teams (FIRST), at <u>https://first.org/;</u> The CyberSecurity Tech Accord, at <u>https://cybertechaccord.org/</u>.

<sup>&</sup>lt;sup>72</sup> See Mark Landler and David E. Sanger, Obama Says He Told Putin: 'Cut It Out' on Hacking, N.Y. TIMES, Dec. 16, 2016.

Accusers interested in exposure of a cyber operation must choose what vehicle they will use and what evidence to share. Accusers may proffer accusations directly. Thus, States have used press releases and speeches to make accusations, while private cyber security companies issue reports detailing their claims.<sup>73</sup> Alternatively, accusers may use proxies to expose information about a cyber operation. One might read Mandiant's APT1 report linking China to commercial cyber espionage as part of a larger U.S. effort to accuse China of acts of commercial cyber-espionage.<sup>74</sup> CrowdStrike was authorized by its client—the Democratic National Committee—to make public its accusation that Russia had hacked the DNC's systems.<sup>75</sup> 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, U.S. officials used media outlets in 2012 to publicize their views that Iran had launched a series of cyber-attacks against U.S. banks.<sup>76</sup>

In addition to deciding whether to make accusations directly or indirectly, accusers must also determine how much documentation to employ. Detailing, and documenting, what happened bolsters an accusation's credibility.<sup>77</sup> Part of what has made accusations from the likes of Mandiant or the University of Toronto's Citizen Lab so powerful is the technical details employed to support their claims.<sup>78</sup> 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, and 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 U.S. National Security Agency surveillance tools demonstrates just how much harm can follow the disclosure

<sup>&</sup>lt;sup>73</sup> See, e.g., U.S. Charges Five Chinese Military Hackers, supra note 9; Condemning Cyber-Attacks by North Korea, supra note 14; Mandiant, APT1, supra note 64; Symantec Security Response, Attackers target dozens of global banks with new malware, SYMANTEC OFFICIAL BLOG, Feb. 12, 2017 (accusing the Lazarus group affiliated with North Korea as responsible for hacking the Bangladesh Central Bank).

<sup>&</sup>lt;sup>74</sup> See Mandiant, APT1, supra note 64.

<sup>&</sup>lt;sup>75</sup> Dmitri Alperovitch, *Bears in the Midst: Intrusion into the Democratic National Committee*, CROWDSTRIKE BLOG, June 15, 2016.

<sup>&</sup>lt;sup>76</sup> See, e.g., Mike Mount, U.S. Officials believe Iran behind recent cyber attacks, CNN, Oct. 16, 2012. Several Iranians were later indicted for their participation in these operations. Seven Iranians Working for Islamic Revolutionary Guard Corps-Affiliated Entities Charged, supra note 16.

<sup>&</sup>lt;sup>77</sup> Land, *supra* note 20, at 208 (discussing how the quality of the "naming evidence" matters).

<sup>&</sup>lt;sup>78</sup> See Mandiant, APT1, supra note 64; Tracking GhostNet, supra note 62.

of means and methods: the NSA's tools provided the foundation for both the WannaCry and NotPetya ransomware attacks.<sup>79</sup>

Consequently, some cyber accusations are unsupported. 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.<sup>80</sup> This led some cyber-security experts to question its accuracy, although others confirmed the U.S. charges.<sup>81</sup> In contrast, other cyber accusations are followed by details that allow the accused and third parties to evaluate the claim as the United States attempted to do in accusing Russia of hacking the Democratic National Committee.<sup>82</sup> 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 in past claims 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 normalized, we expect demands for documentation to rise, along with efforts to harmonize the standards by which third parties can review the accuracy of an accusation's claims.

#### C. Condemnation

Condemnation refers to an expression of disapproval.<sup>83</sup> Accusations will generally involve behavior that the accuser deems wrongful in some way. Sometimes, an accuser simply expresses distaste for what occurred for reasons that may be idiosyncratic. In most cases, however, condemnations have a reference point—a normative standard from which the accused's behavior supposedly diverged.

Condemnations can vary in the specificity with which they reference the normative standard. In some case, the standard is left unstated, or the accused's behavior is simply labeled as "bad." At other times, the normative standard may be referenced explicitly. Condemnations may, moreover, invoke norms that have different bases of propriety. Norms can delineate appropriate

<sup>&</sup>lt;sup>79</sup> See, e.g., Lily Hay Newman, *The Leaked NSA Spy Tool that Hacked the World*, WIRED, March 7, 2018.

<sup>&</sup>lt;sup>80</sup> Imposing Additional Sanctions with Respect to North Korea, supra note 28.

<sup>&</sup>lt;sup>81</sup> Compare Gary Leupp, <u>A Chronology of the Sony Hacking Incident</u>, COUNTERPUNCH (Dec. 29, 2014); <u>New Clues in Sony Hack Point to Insiders, Away from DPRK</u>, SECURITY LEDGER (Dec. 28, 2014), with Novetta, <u>Operation Blockbuster: Unraveling the Long Thread of the Sony Attack</u>, BLOCKBUSTER REPORT (Feb. 2016).

<sup>&</sup>lt;sup>82</sup> Joint Analysis Report, *supra* note 34.

<sup>&</sup>lt;sup>83</sup> 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 20, at 18 (noting that "although shame discourse dominates conventional arguments and the popular human rights lexicon, 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.

behavior by reference to religion, politics, culture, and law (whether domestic or international).<sup>84</sup>

In cyberspace, accusations to date have condemned the accused's behavior in general terms (e.g., as "malicious").<sup>85</sup> 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.<sup>86</sup> President Obama referred to the Sony Pictures hack as an act of "cyber vandalism," but that was a novel phrase without any clear normative antecedents.<sup>87</sup>

Such limited condemnation is not, however, due to an absence of normative candidates. In 2015, a U.N. Group of Governmental Experts (GGE) reached consensus on a list of "voluntary" norms of responsible State behavior in peacetime.<sup>88</sup> Moreover, as the two *Tallinn Manuals* demonstrate, international law offers a range of norms that may both constrain and facilitate State cyber operations.<sup>89</sup> Yet, States have not used the language of the UN GGE (e.g., its prohibition on targeting critical infrastructure in peacetime) to condemn other States' cyber operations even as Russia purportedly targeted Ukrainian power grids.<sup>90</sup> Moreover, as Efrony and Shany's survey reveals, States have, to date, not condemned cyber operations with reference to the *Tallinn Manuals* or the international law they purport to codify.<sup>91</sup>

#### **D.** Constructing Accusations

How are accusations constructed? We believe that States should—and in many cases do—shape accusations according to the function(s) they want the accusation to serve. This may require employing all three processes in the accusation– attribution, exposure, and condemnation—but in other cases, two of the three may suffice.

<sup>&</sup>lt;sup>84</sup> See Finnemore and Hollis, *supra* note 2, at 441-42.

<sup>&</sup>lt;sup>85</sup> See supra note 14 and accompanying text.

<sup>&</sup>lt;sup>86</sup> Id.

<sup>&</sup>lt;sup>87</sup> See Imposing Additional Sanctions with Respect to North Korea, supra note 28; Sean Sullivan, Obama: North Korea hack 'cyber-vandalism,' not 'act of war, WASH. POST (Dec. 21, 2014).

<sup>&</sup>lt;sup>88</sup> Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security, U.N. Doc. A/70/174 (July 22, 2015) ¶13 ["2015 GGE Report"].

<sup>&</sup>lt;sup>89</sup> See Michael Schmitt (Ed.), Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations (NATO CCD COE, 2017) ("*Tallinn 2.0*"); Michael Schmitt (Ed.), Tallinn Manual on the International Law Applicable to Cyber Warfare (NATO CCD COE, 2013).

<sup>&</sup>lt;sup>90</sup> This may be because not all States believe the 2015 GGE Report—the product of consultations among twenty 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.

<sup>&</sup>lt;sup>91</sup> Efrony and Shany, *supra* note 5, at 73.

Attributions, for example, may not be required in cases where shoring up defenses is a priority. In other words, there can be useful and consequential accusations of bad behavior without identifying an accused.<sup>92</sup> In such cases, it may be sufficient to share the vulnerability and technical indicators of the malware. Accusations surrounding the Triton/Trisis malware, for example, have yet to identify its authors (beyond suggesting it was likely the work of a State); that has not, however, stopped cybersecurity firms from accusing someone of planting it in Saudi systems and alerting relevant communities to defend against the threat posed.<sup>93</sup> Similarly, we can envision scenarios where an accusation may catalyze norm construction even if the author of the cyber operation is unknown. In such cases, the accuser may call on other members of the relevant community to join it in condemning the behavior and settling on a norm that prohibits it.

On the other hand, at least some attribution appears necessary if an accusation involves compliance, punishment or deterrence. Pressure for compliance, for example, requires identifying which actor(s) must change their behavior.<sup>94</sup> Similarly, the deterrent value of an accusation lies in showing third parties they too could be identified and accused if they engage in the cited behavior; accusations that fail to identify the culprit are unlikely to have much, if any, deterrent effects.

For accusations that necessitate attribution, accusers must also weigh how much specificity and certainty to convey. If the point is punishment, the accusation may require as much detail as the accuser can muster – indictments, after all, must name an accused. In other cases, however, an accuser's purpose may be achieved with less certainty or specificity. States sensitive to being stigmatized by an accusation may respond to more obliquely framed accusations that source the cyber incident to the named State without directly accusing the government. Accusations of this kind give the accused opportunities and incentives to comply without losing face. For example, attributions that specify only a territorial origin of a piece of malware or only identify non-State actor authors without attributing their behavior to the State,

<sup>&</sup>lt;sup>92</sup> *Cf.* Lin, *supra* note 23, at 6 ("identification of the specific actor is not necessarily required to infer bad intention").

<sup>&</sup>lt;sup>93</sup> See Elias Groll, Cyberattack Targets Safety System at Saudi Arabia, FOREIGN POLICY (December 21, 2017). 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.

<sup>&</sup>lt;sup>94</sup> The naming and shaming literature has already recognized a version of this problem. While it 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 behavior. See M. JURKOVICH, FEEDING THE HUNGRY (forthcoming, 2019).

leave States room to respond in a variety of ways (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.<sup>95</sup> This was also the approach that China apparently took in response to media reports that it bore responsibility for hacking the U.S. Office of Personnel Management—i.e., rather than admitting it was complicit, it arrested several Chinese hackers and identified them as the real culprits (charges many U.S. officials regard as suspect).<sup>96</sup>

What about exposure? Although exposure may, in certain circumstances, improve the chances for compliance, we do not believe this will always hold true. We suspect private accusations may work just as well, and sometimes better, in at least some cases. Punishment may also 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.<sup>97</sup> Both issues, however, could use further research to confirm the attractiveness of non-public accusations in various contexts.

In contrast, exposure is clearly a pre-requisite for accusations designed to shore up defenses and deterrence; if third parties do not know of a cyber operation and of the disapproval of it they can neither defend against it nor are they likely to be deferred from engaging in similar behavior. 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 which over time may acquire the requisite *opinio juris*.

As for condemnation, we envision it will play a key role in accusations that pursue punishment. Without a condemnation, it is unclear what the accuser seeks to punish. Similarly, if the goal of an accusation involves deterring third parties, it should convey the accuser's disapproval as otherwise the accusation might be read as an invitation for others to pursue the newly exposed conduct. Condemnations may have less purchase in accusations that emphasize deterrence. Reports like Mandiant's on APT1 focused on exposing and attributing "malicious" acts but with little by way of condemnation.<sup>98</sup>

The relationship between compliance and condemnations is more complicated. Certainly, strongly condemning behavior by an accused may cause the accused to change its behavior—that is the logic that underlies a "naming and shaming" strategy of any kind. But condemnations—particularly

<sup>&</sup>lt;sup>95</sup> Chris McGreal and Bobbie Johnson, *Hillary Clinton criticises Beijing over internet censorship*, THE GUARDIAN, Jan. 21, 2010.

<sup>&</sup>lt;sup>96</sup> Ellen Nakashima, *China: Hackers' Arrested*, WASH. POST, Dec. 3, 2015, at A3.

<sup>&</sup>lt;sup>97</sup> See ASR, supra note 40, Art. 52.

<sup>&</sup>lt;sup>98</sup> See Mandiant, APT1, supra note 65.

public condemnations—risk stigmatization that may lead an accused to retrench or repeat the condemned behavior.<sup>99</sup> Those involved in truth and reconciliation commissions are often at some pains to highlight this point. Exposure may be necessary to uncover truth and to promote larger goals of legal reform and social change, but political fallout from active condemnation may alienate crucial parties in the peace process.<sup>100</sup> Similarly, when regulatory authorities try to move companies toward better behaviors, condemnations via fines or public sanctions may be a useful deterrent, but stigmatization may also make crime worse and can create adversarial relationships between regulator and companies that are counterproductive.

Where accusers fear a back-lash, accusations may substitute technical assistance for condemnation in a process known as "reintegrative shaming."<sup>101</sup> Social science research suggests this approach can produce better results, especially in situations where there is ambiguity about the relevant rules of behavior. Such engagement can provide useful guidance about what compliance actually means and contribute to a process of creating consensus about right action.<sup>102</sup>

Looking at the cyber context, this approach may have the most utility where an accused failed to act or acted negligently (say, by failing to be diligent in ensuring an otherwise lawful cyber operation stayed within its expected parameters). It is less likely to be useful where an accused actively adopted behavior that is malicious or unwanted. We would also suspect it may have little effect when the accused operates outside—or at some distance from—the relevant community. Thus, we don't envision much utility in reintegrative shaming when dealing with rogue States like North Korea or Iran.

When it comes to constructing norms, condemnations may play a key role. Condemnation of a practice may serve as the basis for articulating publicly what "good" (or lawful) behavior looks like. Such an articulation could then form the basis for a new norm or legal rule. Yet, norm construction may occur in some circumstances without shaming. Consider Stuxnet. On June 1, 2012, New York Times reporter David Sanger published a story that assigned responsibility for the virus (which destroyed up to 1000 centrifuges in Iran's nuclear program) to the United States and Israel.<sup>103</sup> Far from condemning the U.S. and Israeli actions, however, the operation was presented quite positively. Stuxnet gave the accused States a new mechanism for opposing nuclear proliferation without causing the death and destruction that

<sup>&</sup>lt;sup>99</sup> See Rebecca Alder-Nissen, Stigma Management in International Relations: Transgressive Identities, Norms, and Order in International Society, 68 INT'L ORG.143 (2014).

<sup>&</sup>lt;sup>100</sup> Eric Wiebelhaus-Brahm, *Promoting Accountability, Undermining Peace? Naming and Shaming in Transitional Justice Processes,* in FRIMAN, *supra* note 20, at 86.

<sup>&</sup>lt;sup>101</sup> van Erp, *supra* note 26, at 288.

<sup>&</sup>lt;sup>102</sup> *Id.* at 288, 290-91.

<sup>&</sup>lt;sup>103</sup> David Sanger, *Obama Order Sped Up Wave of Cyberattacks Against Iran*, THE NEW YORK TIMES (June 1, 2012).

accompany the use of conventional weapons. Thus, one could interpret the exposure of Stuxnet as an effort, in this case by media actors, to establish the propriety of using this new capacity over more traditional kinetic means (with their attendant death and destruction).<sup>104</sup> The international community has not, however, embraced that idea. When and where such operations are appropriate remains unclear, and in some cases contested, while the reverse-engineering of Stuxnet into the Shamoon and BlackEnergy malware suggests that its benefits may not so clearly outweigh its costs.<sup>105</sup>

The fact that Stuxnet was celebrated in some circles and condemned in others reveals, moreover, that accusations may work differently with different audiences. An accusation may contain a condemnation that resonates with one audience but not another. Even as the OPM hack was condemned within a U.S. domestic law framework as a breach of national security, the U.S. Director of National Intelligence, James Clapper, indicated that such behavior was acceptable among States: "You have to kind of salute the Chinese for what they did,' adding the U.S. would have done the same thing if it could."<sup>106</sup>

Thus, the efficacy of a condemnation will not depend solely on how well it pairs with its anticipated function, but also on various features of the surrounding circumstances. Those interested in pursuing accusations must attend to these circumstances in deciding whether and how to pursue an accusation. And when they do, they should also consider these same circumstances in constructing their accusation to achieve the desired outcome(s).

#### **III.** UNDER WHAT CONDITIONS DO ACCUSATIONS WORK?

Accusations will not work at all times or in all conditions. The surrounding circumstances will often dictate whether an accusation can work at all or whether it will work for certain purposes but not others. Precisely which conditions allow what actions deserves more research, but we highlight four conditions that might bear on the success of accusations: (i) the existence of a norm for measuring what happened; (ii) the relationship between the accuser and the accused; (iii) the relationship between the accuser and the community that serves as the audience for the accusation; and (iv) the relationship between the accuser and that same community. Where accusations seek compliance, we believe the first three conditions will be most

<sup>&</sup>lt;sup>104</sup> Stuxnet infected similar systems world-wide, but was designed apparently to execute only on Iran's Natanz facility, leaving other systems unharmed (although understandably requiring the owners of such systems to clean them of the virus once it became known). D Albright et al, *Did Stuxnet Take out 1000 Centrifuges at the Natanz Enrichment Plant?* (ISIS, 2010).

<sup>&</sup>lt;sup>105</sup> See Nicole Perlroth, In Cyberattack on Saudi Firm, U.S. Sees Iran Firing Back, N.Y. TIMES (Oct. 23, 2012).

<sup>&</sup>lt;sup>106</sup> Jim Sciutto, *Director of Nat'l Intelligence blames China for OPM hack*, CNN, June 25, 2015.

relevant. In contrast, where the accusation serves as the basis of norm construction, the last condition deserves priority.

#### A. No Norm? No Compliance

As a tool of compliance, accusations require a norm against which the accused's behavior can be measured. The existing naming and shaming literature has not, however, emphasized this condition to date. This may be because when it comes to areas featured in that literature—e.g., human rights, the environment—there is little debate over the existence of norms. States—including the accused—do not contest the norm prohibiting torture (or genocide, or significant transboundary pollution, etc.).<sup>107</sup> Rather, accused States focus on denying what the accused says happened or offer a different interpretation or application of the norm than that proffered by the accused.<sup>108</sup>

By contrast, norms governing online behavior are not always as clear and well-entrenched. This is problematic from a compliance perspective. If there is no norm, there can be no compliance. Even in cases where the accused does reference a norm, the more its existence is contested (or its meaning open to dispute), the more likely such circumstances will undermine the accusation's efficacy in generating compliance by the accused.

Consider, for example, recent debates over whether a State's cyber operation effecting another State's territory violates the latter State's sovereignty. *Tallinn Manual 2.0* answers the question in the affirmative.<sup>109</sup> Others, however, have questioned if sovereignty is even a rule governing State behavior as opposed to a background principle that informs the content of other rules (such as the duty of nonintervention).<sup>110</sup> Most recently, the U.K. Attorney General firmly placed the United Kingdom in the sovereignty-as-background-principle camp.<sup>111</sup> As such, accusations that one State has violated another's sovereignty is even a rule of behavior than the more focused question of whether the accused will comply with a norm whose existence it accepts.<sup>112</sup>

# **B.** What is the Relationship between the Accuser and the Accused?

The relationship between the accuser and the accused will regularly be a key factor in assessing the likely efficacy of an accusation. The more an

<sup>&</sup>lt;sup>107</sup> See supra note 43 and accompanying text.

<sup>&</sup>lt;sup>108</sup> See supra note 7 and accompanying text.

<sup>&</sup>lt;sup>109</sup> *Tallinn 2.0, supra* note 89, at 17 (Rule 4).

<sup>&</sup>lt;sup>110</sup> See, e.g., Gary Corn, Tallinn Manual 2.0—Advancing the Conversation, JUST SECURITY, Feb. 15, 2017.

<sup>&</sup>lt;sup>111</sup> Jeremy Wright, QC, MP, *Cyber and International Law in the 21st Century*, May 23, 2018, www.gov.uk/government/speeches/cyber-and-international-law-in-the-21st-century.

<sup>&</sup>lt;sup>112</sup> On the different implications of existential arguments in international law, see Duncan B Hollis, *The Existential Function of Interpretation in International Law* in A BIANCHI ET AL (EDS), INTERPRETATION IN INTERNATIONAL LAW (OUP 2015), 78-79.

accused values its relationship (whether politically, economically, or socially) with the accuser, the greater the likelihood the accusation may prove effective. Where the accusation relies on a norm, the accused may make efforts to comply. Or, if the accusation contains a call for a new norm, the accused may be more likely to accept it if it wants to retain its status vis-à-vis the accuser. We would expect, for example, that accusations of U.K. hacking a European ally's telecommunications carrier (e.g., Belgacom) are more likely to constrain its future behavior than accusations that it targeted a similarly situated Russian company.<sup>113</sup>

At the same time, the accuser-accused relationship is not unidirectional. Accusers also need to consider how an accused's response may impact their own position. On the one hand, the accused may escalate in response to the accusation, creating new problems for the accused beyond those it originally faced. On the other hand, the accused may reject or ignore the accusation, leaving the accused with the dilemma of escalating themselves, or risk looking weak for failing to do more to engender compliance. Where the accuser has material leverage over the accused this may not be an insurmountable problem. Still, this is a problem likely to be exacerbated in cyberspace. Many cyber operations fit uncomfortably below the threshold of armed conflict but above other coercive measures. This means that States have few readily available coercive measures to redress unwanted cyber operations.<sup>114</sup> Instead, accusers may be incentivized to pursue accusations with less exposure or condemnations that avoid stigmatization if possible.

# C. The relationship between the accused and the community within which the behavioral norm is situated

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.<sup>115</sup> Accusations seeking behavioral changes by the accused assume that perpetrators have prosocial reputations they want to protect and/or a moral compass of some kind. 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 having the capacity to engage in destructive cyber operations.<sup>116</sup> Indeed, they may seek to profit from it on the Dark Web or in other nefarious corners of the Internet.

There is in any case, substantial existing research on how to measure the likelihood that social ties may generate norm compliance. According to

<sup>&</sup>lt;sup>113</sup> See, e.g., Ryan Gallagher, *How U.K. Spies Hacked a European Ally and Got Away with It*, THE INTERCEPT, Feb. 17, 2018.

<sup>&</sup>lt;sup>114</sup> Goldsmith and Russel, *supra* note 12.

<sup>&</sup>lt;sup>115</sup> MARGARET E. KECK & KATHRYN SIKKINK, ACTIVISTS BEYOND BORDERS: ADVOCACY NETWORKS IN INTERNATIONAL POLITICS 29, 208 (1998).

<sup>&</sup>lt;sup>116</sup> See Adler-Nissan, supra note 99, at 170.

Goodman and Jinks, the likelihood of a positive response to an accusation from a State depends on the strength, immediacy, and size of the group in which the accused shares an identity.<sup>117</sup> Interestingly, the social science research on which they rely suggests that the most effective groups have 3-8 members, with the efficacy of compliance for larger groups dropping off rapidly. That fact does not bode well for international law and the nearly two hundred nation States subject to it.<sup>118</sup> Still, the accused's sensitivity to its reputation and the moral leverage of a group will be key factors in evaluating the potential for compliance.

# **D.** The relationship between the accuser and the community within which the behavioral norm is situated

Given its focus on "shaming" the accused and obtaining compliance, existing research has examined the accuser's identity in terms of its capacity to move the accused to a different course of action. In doing so, the literature has undertheorized an equally important connection—the capacity of the accuser to influence the larger community into constructing a new norm around the undesired behavior (or to applying an existing norm in some way). After all, it is the community—not the accused—that will be the decider on whether norm development bears fruit. The community's view of the accuser may therefore matter more to a proposed norm's reception than its view of the accused. This is especially true in the global cybersecurity context where some of the most significant operations are conducted by States (or their proxies) who already have reputations as rogue actors or marginal members of the international community. In other words, States are more likely to accommodate normative views on the impropriety of WannaCry coming from the United Kingdom (the accuser) than North Korea (the accused).<sup>119</sup>

The accuser's identity may prove relevant to norm construction in two respects. First, where the accuser has power (or material leverage) within the group, others are more likely to give its views serious weight. The power and position of accusers can also influence the willingness of third parties to dismiss or to press their accusation and recruit more accusers to their cause.

This is not to suggest that an accuser must be a powerful State; on the contrary, accusations are regularly deployed by non-governmental organizations who lack such authority and must rely on their reputation and credibility within the community. Thus, a second condition for evaluating the potential of accusations to generate norm construction lies in the accuser's reputation. Is the accuser a trusted actor? Have its previous accusations been corroborated and accepted? Or, is the accuser perceived to have a personal

<sup>&</sup>lt;sup>117</sup> RYAN GOODMAN & DEREK JINKS, SOCIALIZING STATES: PROMOTING HUMAN RIGHTS THROUGH INTERNATIONAL LAW 28 (2013). "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 id.* 

<sup>&</sup>lt;sup>118</sup> Id.

<sup>&</sup>lt;sup>119</sup> See supra note 63, and accompanying text.

agenda or motives apart from those of the system as a whole. In short, the power and reputation of accusers can have important consequences for the efficacy of accusations generally, and norm construction specifically.

### IV. THE RELATIONSHIP BETWEEN ACCUSATIONS AND INTERNATIONAL LAW

How do accusations interact with international law? Most obviously, they can be a source of compliance. If conditions are favorable, an accused may become more compliant in response to the accuser's condemnation of a legally wrongful act (or a failure to act). As noted, however, cyber accusations have yet to take advantage of this possibility. What explains this reluctance to invoke international law?

For starters, at least some of the accusations to date involve behavior currently regarded as legally appropriate. The OPM hack, for example, may have severely undermined U.S. national security at a scale not seen previously. Yet, from the perspective of international law, this was an act of espionage, that international either fails to regulate or affirmatively permits.<sup>120</sup> As such, it is not surprising to see accusations of China's responsibility for the OPM hack avoid condemnations in international legal terms.

The same rationale may explain the reluctance to invoke other international legal rules that have divided States at the GGE and elsewhere. The 2017 U.N. Group of Governmental Experts failed to achieve consensus reportedly because States divided over whether (and how) various international legal rules, including self-defense, international humanitarian law, the duty of non-intervention, sovereignty, and due diligence, applied in cyberspace.<sup>121</sup> Consequently, some States may opt to 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. In such cases, silence may actually do more to extend the norm's availability for future cases than near-term contestation.

Alternatively, States may decline to invoke international law rules out of reciprocity concerns. Iran, for example, never challenged U.S. and Israel's role in Stuxnet as a use of force or even an armed attack (triggering a right of self-defense), preferring instead to deploy its own cyber operations against U.S. financial targets without any legal framing at all.<sup>122</sup>

<sup>&</sup>lt;sup>120</sup> See Ashley Deeks, An International Legal Framework for Surveillance, 55 VIRG. J. INT'L L. 291, 300 (2015).

<sup>&</sup>lt;sup>121</sup> See Arun Mohan Sukumar, *The UN GGE Failed. Is International Law in Cyberspace Doomed as Well?*, LAWFARE, July 14, 2017; Michael Schmitt and Liis Vihul, *International Cyber Law Politicized: The UN GGE's Failure to Advance Cyber Norms*, JUST SECURITY, June 30, 2017.

<sup>&</sup>lt;sup>122</sup> See supra note 105 and accompanying text.

Even if an accuser believes that sufficient consensus exists around the existence of an international legal norm, documentation issues may serve as another barrier to referencing it. International legal accusations pose particular evidentiary challenges. Accusers must tie the accused State to the actual hackers, whether by 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 are later adopted by the State.<sup>123</sup> 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 do so). Indeed, among the norms agreed to by the 2015 U.N. GGE was that "the accusations of organizing and implementing wrongful acts brought against States should be substantiated."124

Additional challenges may have little to do with norm creation and violation. In many—but not all—cases, the accuser and the accused faced strained relations *before* an accusation about a cyber operation. Invoking international law in an already tense relationship might risk escalating the situation even further. Alternatively, 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."<sup>125</sup> 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 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 flaunt in other contexts.

As significant as these challenges are, a more nuanced understanding of how accusations work suggests some potential measures States could make to improve the utility of their international legal accusations. First, accusers could do more to reduce the risk of escalation or retrenchment by the accused. Accusations that attribute cyber operations to the territorial origins of the

<sup>&</sup>lt;sup>123</sup> See, e.g., Nicholas Tsagourias and Michael D Farrell, *Cyber attribution: technical and legal approaches and challenges*, EJIL (forthcoming 2018). Of course, the standards of control required to establish a State's responsibility for acts of non-State actors are disputed, with the International Court of Justice favoring a rule of "effective control," in contrast to the rule of "overall control" advocated by the International Criminal Tribunal for the Former Yugoslavia. *Compare Case concerning Military and Paramilitary Activities in and against Nicaragua (Nicaragua v USA)* (Merits, Judgment) [1986] ICJ Rep 14, 64–5 [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) [1997] ICJ Rep 43, 208–9 [399]–[401]; *with Prosecutor v Dusko Tadic aka 'Dule'* (Judgment) ICTY-94-1-A (15 July 1999) [131], [145].

<sup>&</sup>lt;sup>124</sup> 2015 UN GGE, *supra* note 88, at ¶28(f).

<sup>&</sup>lt;sup>125</sup> Goldsmith and Russel, *supra* note 12, at 13.

operation rather than to the responsible State itself might leave the accused State more face-saving ways to respond. Alternatively, accusers insistent on highly specific and certain charges might employ private communication channels in lieu of public exposure.

States and other stakeholders might also consider reframing accusations that, to date, have centered on attributing—and stigmatizing— State actors by proscription (calling on the accused to stop doing something). A more effective approach might be accusations critiquing States for a failure to act to control behavior within its territory. Whether or not the State was in control of the non-state actors in question at the time of the accusation, accusations of this type might induce the accused State to realign that relationship and exercise more control over sub-state actors' unwanted behavior.<sup>126</sup> In addition, States and other stakeholders might consider the value of generating lists or rankings of State cyber practices in much the same way as the United States has done with its annual Human Rights Report or the World Bank does with its "ease of doing business" rankings.<sup>127</sup> In some cases, it might even make sense to tie poor performance in such listings to capacity building opportunities that assist the accused in reducing instances where its territory originates unwanted cyber behavior.

What our broader and more nuanced vision of accusations suggests, moreover, is that States and stakeholders should not limit their expectations entirely to compliance. Properly constructed, accusations may create opportunities to clarify the international legal norms that currently govern State cyber operations and/or to build new rules that do so. Simple steps could improve the credibility of existing accusations—and the norms they promote—within the relevant communities of States or other stakeholders. Agreeing to more standardized attribution methodologies 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.

Recent accusations involving WannaCry and NotPetya suggest, moreover, that increasing the number of accusers might raise the credibility of the claims made. Since custom requires not only a uniform practice, but a general one, the more accusers expressing disapproval of certain cyber operations, the easier it becomes to have the type of practice that can acquire *opinio juris*. Eventually, accusations might find their most force when generated in multilateral or multistakeholder fora.

Finally, States and other stakeholder might improve the credibility of accusations by employing more trusted intermediaries to assist in cyber governance. This approach, for example, lends credence to efforts favoring

<sup>&</sup>lt;sup>126</sup> Healey, *supra* note 65, at 4 ("Under international pressure, most nations could likewise reduce attacks from their territory of cyberspace through several well-established steps . . .").

<sup>&</sup>lt;sup>127</sup> Judith G. Kelley and Beth A Simmons, *Politics by Number: Indicators as Social Pressure in International Relations*, 59 AM. J. POL. SCI. 55, 56 (2015).

the creation of an independent and impartial institution to engage in attribution of cyber operations (either directly or via peer review).<sup>128</sup> If neutral actors issue specific and credible claims of unwanted behavior, they may lead States (either globally or in more like-minded groups) to coalesce around new international legal rules proscribing such behavior.

### 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 toward more pro-social and norm compliant behavior.

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. International relations literatures tend to assume they do—that once a state is "named" in an accusation, shame and shaming behavior will automatically follow. Our paper starts from the puzzle that, in cybersecurity, this link between naming and shaming is weak and, of particular interest for this volume, 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 have different political effects depending on: the identity and reputation of accused and the accuser(s); the way the accusation is framed; and the types of evidence offered in support of the charge. The role of third party audiences is also crucial, and there are good reasons to expect private accusations among states to generate different effects than public ones. 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 time, determine the social contours of the cyber security 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, in turn, 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

<sup>&</sup>lt;sup>128</sup> See, e.g., Davis et al, *supra* note 3, at 3; Jay Healey et al, *Confidence Building Measures in Cyberspace: A Multistakeholder Approach for Stability and Security* (Atlantic Council, 2014); Brad Smith, "*The need for a Digital Geneva Convention*", MICROSOFT (Feb. 24, 2017) https://blogs.microsoft.com/on-the-issues/2017/02/14/need-digital-geneva-convention/.

the responses they generate—could advance international law compliance. But the failure of accusations to achieve compliance by an accused does not relegate the law to irrelevancy. The interchanges following an accusation do much to tell us what behavior is—and is not—accepted by States. The resulting delineation of wrongful behavior in cyberspace can constitute the practice from which *opinio juris* may emerge over time.

For policy makers, 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, 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 reaction they want from either 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 party 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.

Policy makers might also consider whether and how additional legal instruments might be helpful in making these accusations. If more clarity would be useful about what the rules and norms *are* for cyber operations, there are clear strategies they can employ to achieve this, beyond simply waiting for practice and *opinio juris* to emerge organically over time. States, for example, could make accusations more collectively, or stand up a neutral institution to do attribution or to advocate for international law. 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 increasing opportunities to use accusations to set the rules of the road, including constraints (and permissions) derived from international law.