

STRATEGIC STUDIES QUARTERLY
**Confidence Building Measures
for the Cyber Domain**

*Erica D. Borghard and Shawn W. Lonergan*¹

Abstract

There is a growing debate among scholars and practitioners in the cyber conflict field regarding the extent to which the cyber domain is likely to be characterized by inadvertent escalatory spirals and arms races between increasingly cyber-capable states. Historically, technological innovation or geopolitical dynamics have propelled states to form confidence building measures (CBM) or create arms control regimes to institutionalize constraints on offensive military technology and guard against inadvertent conflict and escalation. We argue that cyber CBMs could blunt some of the factors that contribute to crises and escalation. Given the absence of arms control regimes for the cyber domain, cyber CBMs could be used to mitigate the risks to stability between states and possibly change the incentives that could lead to crises. In assessing current cyber confidence building initiatives, this article creates a novel framework to better understand these efforts. It also identifies limits of cyber CBMs and provides prescriptions for new steps in cyber CBMs to enhance mutual security and guard against inadvertent conflict stemming from cyber operations.



There is a growing debate among scholars and practitioners in the cyber conflict field regarding the extent to which the cyber domain is likely to be characterized by inadvertent escalatory spirals and arms races

Erica D. Borghard is assistant professor at the Army Cyber Institute at the United States Military Academy at West Point and a Council on Foreign Relations International Affairs Fellow. She holds a PhD in political science from Columbia University.

Shawn W. Lonergan is a research affiliate of the Army Cyber Institute at the United States Military Academy at West Point and a cyber officer in the US Army currently assigned to US Cyber Command. He holds a PhD in political science from Columbia University.

between increasingly cyber-capable states.² Furthermore, policy makers find themselves grappling with competing incentives. On the one hand, actions taken to limit the use of destructive cyber weapons or the targeting of civilian infrastructure could provide some assurances for digitally dependent societies. On the other hand, policy makers are loath to support self-imposed limits on capabilities in an environment where future technological trends are uncertain and adversary capability and motivations are difficult to discern and predict. Historically, technological innovation or geopolitical dynamics have propelled states to form confidence building measures or create arms control regimes to institutionalize constraints on offensive military technology and guard against inadvertent conflict and escalation. But to what extent can cyber CBMs be used to mitigate the risks to stability between cyber powers? Is it possible to change the incentives that could lead to crises? We argue that, while there are fundamental attributes of operating in the cyber domain that impede efforts to build effective and enforceable arms control regimes, CBMs, which are distinct from arms control, could blunt some of the factors that contribute to crises and escalation. In assessing current cyber confidence building initiatives, this article creates a novel framework to better understand these efforts and to identify areas that are not being addressed and remain as potential flashpoints that could exacerbate tensions and spark conflict.

First, we conduct a brief discussion of the role of CBMs in fostering stability and reducing the risk of inadvertent escalation and situate their development in a historical context. Next, we review the hurdles to establishing arms control regimes for the cyber domain and demonstrate how, despite these hurdles, states have demonstrated a willingness to enter into CBM agreements to clarify acceptable behavior in cyberspace, avoid inadvertent conflict, and stabilize potential disruptions to international security stemming from cyber operations. We use Cold War frameworks for evaluating CBMs as a benchmark for developing realistic CBMs for the cyber domain in light of the latter's distinct characteristics.³ Specifically, cyber CBMs must take into account the multi-stakeholder nature of the cyber domain, as distinguished from other domains of warfare; the different types of information that should be shared for CBMs to be effective; the dual-pronged nature of the objectives of CBMs, which could be used not only to avoid cyber conflict, but also to bolster norm development efforts; and the administration and main-

tenance of cyber CBMs through unique mechanisms such as the United Nations Group of Governmental Experts (GGE) and the Organization for Security and Co-operation in Europe (OSCE). The article concludes by identifying the limits of cyber CBMs and provides prescriptions for next steps in cyber CBM development. Importantly, there are additional measures that could be taken to enhance mutual security and guard against inadvertent conflict stemming from cyber operations.

Confidence Building Measures as Reassurance

When arms control is perceived to be a bridge too far between adversaries that hold many points of disagreement and mistrust, yet both acknowledge the potential for inadvertent conflict, decision makers have employed confidence building measures in lieu of establishing arms control regimes.⁴ The post–Cold War literature on international law and institutions has reconsidered the value of soft law and information norms and institutions in terms of their contributions to fostering stability and reassurance between strategic rivals.⁵ Like arms control, CBMs may constitute bi- or multilateral agreements or take the form of unilateral action. As trust is built between parties, CBMs may give way to more formalized arms control agreements due to the role the former have in reassuring a potential adversary—though this is by no means determinative. According to this logic, CBMs are a form of reassurance that seeks to demonstrate intent among rivals, therefore (ideally) conveying a desire to maintain the status quo and foster a sense of security between otherwise threatened states.⁶ Indeed, they are designed to ensure crisis situations, routine tensions, or localized conflicts between states do not become inadvertent lightning rods that spark a general war.⁷ As CBMs are only intended to signal the aim of military activities, they do not change the overall balance of power between adversaries. Rather, CBMs are simply designed to preserve a fragile stability in the context of potentially intense security competition between states.

Confidence building measures provide reassurance through four mechanisms. First, they seek to demonstrate nonaggressive postures by increasing the transparency of military actions. This could occur, for instance, through inviting designated observers or the public to witness events that otherwise could be construed as threatening.⁸ Second, they place self-imposed limits on security activities, such as military exercises, that could cause another state to feel threatened. Third, CBMs often op-

erate in a time of crisis by enabling a vital communications link between adversaries. In other words, CBMs contribute to stability and détente by helping convey intent behind a state's unilateral security policies and actions that would otherwise be cloaked in uncertainty.⁹ Finally, CBMs inject predictability into a potential adversary's actions and, therefore, serve as an early warning function. Specifically, CBMs make it easier for another state to detect a deviation from an established norm of behavior and thus enable it to take measures in advance to mitigate the damage stemming from a surprise attack.¹⁰ Though CBMs do not replace the vital role of national technical means of intelligence in assessing another actor's capabilities and intent, they supplement it by enabling a fuller picture of the significance of a military policy or action than otherwise would have been available.¹¹

During the Cold War, there were concerns among scholars and policy makers that CBMs could be used to mask a surprise attack, but these were overcome due to the mutually paramount interest of avoiding inadvertent conflict that could spiral into nuclear war.¹² Specifically, governments mitigated these apprehensions through voluntarily implementing more, rather than less, transparency to reassure rivals about the intent behind a military action or policy. Though CBMs can be unilaterally implemented, they often take the form of multi- or bilateral agreements so all parties can understand the level of transparency necessary to foster mutual and reciprocal confidence in the intent behind another actor's security policy or action. The Helsinki Final Act in 1975 is a case in point.

CBMs in Historical Context: The Helsinki Final Act

CBMs need not be formalized in international law or codified in a formal agreement to be effective. While they sometimes become institutionalized over time as they evolve from state practice, CBMs can have an independent effect on stability and cooperation through informal and norms-based mechanisms.¹³ The exemplar for all subsequent CBM efforts was the Final Act of the Conference on Security and Co-operation in Europe that took place in Helsinki, Finland, in 1975.¹⁴ We use the Helsinki Final Act, therefore, as our benchmark for assessing cyber CBMs. Broadly speaking, the 1975 conference had the goal of creating stability, noting "the need to contribute to reducing the dangers of armed conflict and of misunderstanding of military activities which could give rise to apprehension, particularly in a situation where the

participating States lack clear and timely information about the nature of such activities.”¹⁵ The Helsinki Final Act, initially signed by 35 states, sought to foster stability by addressing issues that strained East-West relations on topics ranging from sovereignty to freedom of the press and cultural exchanges.¹⁶ Arguably, no part of the agreement has been as closely scrutinized as the establishment of CBMs between the signatories. The original act stipulated voluntary reporting with at least a 21-day prior notification of military maneuvers that would exceed over 25,000 troops and that would occur within 250 kilometers from a state’s border.¹⁷ The provision also enabled the exchange of observers for these maneuvers as well as the hosting of military delegations.¹⁸

The Helsinki Final Act noted that “the experience gained from the implementation of the provisions . . . together with further efforts, could lead to developing and enlarging measures aimed at strengthening confidence” and as such created a framework for follow on meetings. The first of these occurred in Belgrade in 1977, followed by Madrid in 1980, Stockholm in 1984, and Vienna in 1986.¹⁹ Each of these conferences comprised multiyear efforts that endeavored to innovate new and creative means to demonstrate intent and promote transparency in response to changing security policies and technology. By the time the 2011 Vienna Document was finalized, CBMs had expanded to include the annual exchange of military information such as organizational charts, manning and equipment numbers, unit locations, defense budgets, and information relating to the employment of new weapon systems.²⁰ Furthermore, additional CBMs included the development of more robust communication regimes that could operate in a time of a crisis as well as for routine exchanges of officers and demonstrations of new major weapon systems. The original provisions for troop notifications were also refined to require at least a 42-day warning of exercises of at least 9,000 troops or 250 battle tanks. There were also controls addressing the number of major exercises that a state could perform per year and restrictions on the number of short-notice inspections of another signatory’s military maneuvers and other troubling sites that a state could annually perform.²¹

The Helsinki Final Act illustrates how CBMs could offer a means to mitigate the risk of inadvertent conflict even under conditions when formalized arms control agreements that seek to change the incentives for military action are not feasible. CBMs do so through facilitating increased transparency and openness surrounding a state’s security policies

and operations. However, changes in security requirements, polices, and technology suggest that, for CBMs to promote lasting stability, they must be reassessed and amended on an iterative basis, as was the case throughout the Cold War and in the ensuing years.

Initial Steps Toward Cyber Confidence Building Measures

CBMs were neither the sole nor most effective means of cultivating stability between nuclear-armed rivals during the Cold War. Mutual fear of miscalculation and escalation drove the United States and the Soviet Union to form arms control regimes.²² Arms control can alter the incentives for the use of offensive military technologies, limit the damage to states in the event these technologies are used, and contribute to stable interstate relations, even between adversaries. However, there are reasons to be less sanguine about the feasibility of arms control for cyberspace.

First, several fundamental characteristics of operating in cyberspace confound the establishment of effective arms control agreements. Specifically, arms control in cyberspace is difficult due to the ambiguity surrounding the strategic balance of cyber weapons and the measurement of relative capabilities of cyber powers, the lack of transparency and issues with monitoring for compliance, the dynamic nature of the methods and means of cyber operations, uncertainty about the military implications of technological innovations, and problems of assigning and enforcing responsibility for cyber operations or capability development.²³ Put simply, this endemic uncertainty means governments do not want to find themselves at a strategic disadvantage if and when a future cyber war breaks out. Furthermore, the offensive parity that exists between many states (and even nonstate actors) in the cyber domain is likely to heighten these fears of being in a potential position of military disadvantage.²⁴ Indeed, while serving as chairman of the Joint Chiefs of Staff, Gen Martin Dempsey noted that the cyber domain is the only domain where the United States possesses peer competitors.²⁵ Second, cyber capabilities have been “weaponized” to deliver effects across two broad categories: to support traditional kinetic war fighting and for the purposes of punishment, subversion, or coercion. The more significant source of instability in cyberspace lies in the latter category rather than the former. Specifically, a key source of instability lies in exploiting national economies and critical infrastructure and manipulating the public’s perception of the integrity of essential systems via cyber means to achieve

strategic objectives. Therefore, traditional concepts of arms control that limit the “quantity” or “quality” of cyber arms, for instance, are poorly suited to address the key contributors to strategic instability between cyber rivals.

Despite the significant hurdles to arms control for cyberspace, states have already taken steps to develop cyber CBMs through multi- or bilateral agreements to create mechanisms to share information about their intended uses of cyberspace and law enforcement information concerning nefarious actors, as well as to share information in a crisis. This is because governments have recognized that the secretive nature of cyber operations and the difficulties of signaling in cyberspace can be destabilizing to interstate relations, increasing tensions and the risk of inadvertent conflict. Therefore, though it is impossible to completely eliminate the incentives for actors to misrepresent or disguise their aggressive cyber actions, CBMs that facilitate a dialogue between states have become a first step toward mitigating the destabilizing effects posed by the cyber domain.²⁶ For example, in the past few years, several countries, such as the United States and Russia, entered into bilateral agreements establishing hotlines to guard against misunderstandings stemming from cyber operations in a crisis. During the fall of the 2016 US presidential election, President Obama used the hotline connection between the Nuclear Threat Reduction Centers, which was bilaterally designated to be used for cyber related events three years prior, to convey to President Putin that the laws of armed conflict applied to cyberspace.²⁷ The efficacy of President Obama’s use of the hotline remains uncertain; Jeanette Manfra, the National Protection and Programs Directorate (NPPD) Assistant Secretary for the Office of Cybersecurity and Communications (CS&C) at the US Department of Homeland Security (DHS), disclosed in February 2018 that the Russians succeeded in penetrating a small number of state election systems, though it is not known if these breaches occurred prior or subsequent to President Obama’s call.²⁸ Thus, in lieu of banning specific capabilities or seeking an agreement that depends on verification, states have sought to use informal, voluntary measures to grapple with the fundamental drivers of instability between cyber rivals by promoting clarity of the domain and enabling effective crisis management.

In the multilateral context, several international organizations have spearheaded attempts to develop cyber CBMs, with varying degrees of

success. Of particular note are the UN GGE and OSCE, which sponsored the original Helsinki Final Act. Additionally, beyond the OSCE, there have been other efforts to foster cyber information sharing and confidence building between states. Groundbreaking regional agreements such as the *African Union Convention on Cyber Security and Personal Data Protection*, the Organization of American States' *Inter-American Strategy to Combat Threats to Cybersecurity*, and the *ASEAN Regional Forum Work Plan on Security of and In the Use of Information and Communications Technologies* have all focused on addressing regional security needs stemming from cyber threats.²⁹ Similarly, there have been efforts by economic organizations, such as the Groups of 7 and 20 (G7 and G20, respectively), to promote norm creation that reflects the interests of the largest economies in the world.³⁰ Both the G7 and G20 declarations explicitly express support for the UN GGE and OSCE CBM development efforts but restrict their focus to the establishment of normative state behavior related to the use of cyber capabilities.

Despite representing the most advanced efforts by the international community to develop cyber CBMs, both the GGE and the OSCE have made only halting progress to arrive at mutually agreeable measures to promote stability and transparency between states in the cyber domain. Within the UN, the GGE on Developments in the Field of Information and Telecommunications in the Context of International Security was convened in 2004 to discuss potential areas of cooperation.³¹ A year later it failed to reach a consensus and no report was submitted. A second GGE was convened in 2009, and, after four meetings over the course of two years, it devised the first set of cyber CBMs that focused on information sharing, reducing risk to critical national infrastructure, and devising a set of commonly accepted terms; it also provided recommendations for continued dialogue.³² The CBMs were expanded by a third and fourth round of GGE panels that concluded in 2013 and 2015, respectively, with notable agreements regarding the application of international law and the concept of sovereignty to cyberspace as well as state responsibility for attributed cyber acts.³³ However, the most recent GGE round in 2016–2017 failed to build on the success of previous iterations. For instance, while the 2013 GGE promulgated that international law, especially the UN Charter, is applicable to the cyber domain, members at the 2017 GGE summit were unable to arrive at a consensus regarding *how* international law should apply. Specifically, the

breakdown of the talks centered around questions of how concepts such as sovereignty, the right to self-defense, and appropriate countermeasures apply to cyberspace, with some members taking the position that it was premature to address these issues given the dynamic nature of the domain.³⁴ It is possible that the most recent GGE round was doomed to fail when assessed against unrealistically high expectations leading up to it. Ongoing processes of interpreting and applying international law are chronically difficult.³⁵ However, the 2017 GGE summit produced a regression from previous agreements that international law itself applied in the first place, not simply a failure to push forward the agenda. Relatedly, the Permanent Council of the OSCE directed efforts in 2012 to begin drafting CBMs specific for cyberspace, noting that CBMs were necessary to “enhance interstate co-operation, transparency, predictability, and stability, and to reduce the risks of misperception, escalation, and conflict that may stem from the use of ICTs.”³⁶ These efforts led to the drafting of additional CBMs in 2013 and a more comprehensive list in 2016.³⁷

While governments have taken initial efforts to establish cyber CBMs, current academic work on the topic is at a nascent stage. Though multiple scholars have noted the need to avoid inadvertent conflict, few have postulated specific measures that states could implement to move in that direction.³⁸ Herbert Lin attributes this dearth of measures to the revolutionary nature of the domain. In Lin’s words,

Meaningful analogs to . . . [confidence building] measures in cyberspace are difficult to find. For example, there is no analog to large-scale troop movements—cyber forces can be deployed for attack with few visible indicators. Agreed conventions for behavior, such as “rules of the road,” do not cover intent, and in cyberspace, intent may be the difference between a possibly prohibited act, such as certain kinds of cyberattack, and an allowed one such as cyber espionage.³⁹

Tughral Yamin notes this dilemma but argues that, “A necessary precondition for developing cyberspace CBMs is to have good national cyber security policies and practices, particularly for the protection of critical infrastructure.”⁴⁰ Yamin does not quantify the requisite level of policy creation necessary for the effective formation of CBMs. However, he does make an important contribution by noting that institutional development of cybersecurity organizations within a state are necessary, in part, because they play an important role in knowledge generation and information sharing in a domain that is difficult to conceptualize.

Absent institutions that assist in information sharing of vulnerabilities, known threats, remediation strategies, and national policies and attitudes for approaching the cyber domain, it is unlikely that actors within and external to a state would understand the risks posed by cyber operations. Indeed, there is an *a priori* need to deliberately cultivate an epistemic community comprised of multidisciplinary and multinational academics, policy makers, the private sector, and operators/planners to arrive at a consensus on pivotal concepts and definitions that drive how actors operate in and through cyberspace similar to the epistemic community that developed during the Cold War to grapple with the implications of nuclear weapons. Additionally, in a thought piece on cyber CBMs, Jason Healey, John C. Mallery, Klara Tothova Jordan, and Nathaniel V. Youd note that, given the plethora of actors in cyberspace, a multistakeholder approach that incorporates the private sector and other nonstate actors is vital to the development and adoption of any measure.⁴¹ However, the implications of this analysis focus on the influence of domestic-level veto players on a state's international bargaining position with respect to the creation of specific CBMs. Therefore, there are opportunities for scholars to make conceptual and analytical contributions to cyber CBM development to better inform policy making.

A Framework for Cyber Confidence Building Measures

The objectives of CBMs—to foster exchanges that help states avoid conflict, rather than actually change the military balance of power—may make these mechanisms more amenable to application to the cyber domain than arms control. Indeed, continued efforts by governments and international organizations to support the development of cyber CBMs are important because they represent the first step in injecting stability and transparency into a domain characterized by secrecy and uncertainty. However, even the most “successful” efforts at developing CBMs have thus far been disappointing. Developing a framework to conceptualize and evaluate different categories of cyber CBMs, taking into account how cyber CBMs are likely to differ from previous types of CBMs, is a necessary foundation to support future CBM development efforts. Therefore, as an initial contribution, we use a model for categorizing CBMs developed during the Cold War as a benchmark for assessing the extent to which it is applicable to the cyber domain, identifying important gaps, and developing cyber-specific approaches

for evaluating CBMs. Johan Holst argued in 1983 that CBMs come in four varieties (information, notification, observation, and stabilization) and noted that some measures may encompass several of these categories.⁴² Information measures involve the sharing of defense-related information, such as budgets and organizational structures, between interested parties. Notification pertains to the advanced warning of major military activities within a geographic concentration, such as a military exercise or a major change in force distribution. Observation measures include activities such as inviting potential adversaries to physically observe military exercises, the fielding of new weapon systems, or other related military activities firsthand. However, as Holst notes, stabilization measures were multifaceted and encompass three dimensions: “crisis stability (relative absence of pressures to take early military action to forestall moves by the adversary); arms-race stability (relative absence of inducement to expand military forces); and political stability (relative absence of pressures for breakdown of the international order).”⁴³ Applying Holt’s framework to the cyber domain, we identify three different categories of information CBMs (with the exception of crisis stability), incorporating into our analysis important factors that were not considered in Holt’s framework; demonstrate why the notification, observation, and stabilization categories of CBMs are likely to be particularly difficult and complex in cyberspace; and account for the development of administrative measures that are designed to promote transparency and the role of the hosting institution. We organize all of the existing OSCE and GGE cyber CBMs into our new framework, which can be found in the appendix.

Three Categories of Information CBMs for Cyberspace

When Holt developed his framework for organizing CBMs during the Cold War, he envisioned the information category as simply an exchange of defense-related data. However, this category should be disaggregated given the diversity of threat actors and the unique complexities associated with operating in cyberspace. For instance, the multistakeholder nature of cyberspace and, in particular, the fact that the private sector owns and operates the vast majority of its infrastructure and is the primary target of cyberattacks means that including private industry as participants in CBMs is essential for their relevance and success.⁴⁴ Private actors may have better information than governments about adversary tactics, tech-

niques, and procedures (TTP) and capabilities. Relatedly, private actors already participate in information sharing independent of government actions. For example, private security firms are often quicker to publicly attribute malicious behavior than governments.

Therefore, information-based cyber CBMs should be categorized into three components: threat actor, security, and use.⁴⁵ First, the sharing of threat actor information identifies threat actors and emerging methods and means for exploitation and attack. This could include sharing information that pertains to specific online personas, country profiles, threat signatures, and TTPs as well as law enforcement information about state and nonstate actors. This type of threat actor information sharing contributes to stability by enabling states to proactively counteract malicious actors and activities in cyberspace directly, rather than defend solely within the perimeter of one's network. An example of this is the December 2013 CBM developed through the OSCE, encouraging states to establish "modern and effective national legislation to facilitate . . . time-sensitive information exchange between competent authorities, including law enforcement agencies, of the participating states, in order to counter terrorist or criminal use of ICTs."⁴⁶

Second, security information pertains to the dissemination of system vulnerability reports as well as instructions for remediation. This contrasts with threat actor information in that it is oriented around systems and networks to be defended, rather than threat actors. Security information contributes to stability by enabling defenders to take proactive measures to protect networks and systems. A common element of both the GGE and the OSCE list of measures is a reliance on computer emergency response teams (CERT) for the dissemination of both threat and security information. Since the first CERT was created at Carnegie Mellon University in 1989, the concept has expanded to include over 420 teams operating in over 80 countries that mutually promote security cooperation by sharing technical vulnerability and remediation information.⁴⁷ Parties to the 2015 GGE, for instance, agreed to share information through the CERT infrastructure about "vulnerabilities, attack patterns and best practices for mitigating attacks."⁴⁸ For example, the US National Institute of Standards and Technology (NIST) publishes publicly accessible, real-time information about ICT vulnerabilities that defenders can use to bolster security.⁴⁹

Third, use information incorporates Holst's conceptualization of the sharing of state-level defense related materials, such as doctrine and national policies. However, for the cyber domain this category should be broadened to incorporate other stakeholders, particularly the private sector as participants in CBMs. The recognized influence and role of the private sector is already evident in both the GGE and OSCE CBMs that address the sharing of information relating to "national attitudes" and views from both public and private sources.⁵⁰ An example of this is the July 2015 GGE CBM in which parties agreed to "the voluntary sharing of national views and information on various aspects of national and transnational threats to and in the use of ICTs . . . and national organizations, strategies, policies, and programmes relevant to ICT security."⁵¹ These CBMs reflect the fact that the actors in this space are not solely states and, therefore, information about the uses of cyberspace must extend beyond traditional state actors and should necessarily include information provided by—not simply about—private actors.

Furthermore, in addition to enhancing transparency regarding motivations and intent, several of the information-use cyber CBMs also serve the purpose of tracking and driving norm emergence and development. There is an essential interdependence and complementarity between CBMs and norms in international politics. CBMs can contribute to norms through creating shared expectations about appropriate behavior (such as acceptable targets) or capability development (such as offensive weapons); norms, in turn, can help foster stability through facilitating identification of defection.⁵² Some cyber CBMs, for instance, are designed to share information concerning promoted norms of state and societal use of the internet within its borders, such as a desire for a free and open internet or a more closed protectionist posture, as well as other information, such as what it considers to be critical infrastructure.⁵³

Finally, there are administrative measures that have been instituted to maintain cyber CBMs and disseminate information that reflect unique needs of the cyber domain and, therefore, are beyond the scope of Holt's initial framework. This also reflects the interdependence of CBMs and norms, because the latter are also often promulgated and propagated through institutions.⁵⁴ Indeed, private sector actors, such as global financial services firms, have used the G7 and G20 as forums to advocate for norms against targeting financial institutions. Specifically, these administrative measures are designed to enable the preservation and continued rele-

vance of the cyber CBMs, as well as the conservation of the respective organizations that facilitated their creation. Two notable examples of this are the CBMs developed in December 2013 through the OSCE in which parties, including allies and competitors, agreed to “exchange views using OSCE platforms and mechanisms” and “meet at least three times each year . . . to discuss information exchanged and explore appropriate development of CBMs.”⁵⁵

The Limitations of Notification, Observation, and Stabilization Measures for Cyberspace

While there is a plethora of information sharing CBMs that have a reasonable chance of successful adoption in cyberspace, significant hurdles remain for the acceptance of other categories of CBMs due to some of the same confounding factors that thus far have impeded the development of arms control regimes in the domain. This explains why there are few stability measures—with the exception of crisis stability—and no notification and observation measures present in both the UN GGE and OSCE frameworks.

Notification and Observation

Notification and observation CBMs are designed to provide advance warning of an exercise to other states so that the exercise is not misperceived to be preparations for an offense and to generally provide reassurance regarding motivations. However, notification of a cyber event or an exercise, to include allowing potential adversaries to observe it, is counterproductive in cyberspace due to the central importance of secrecy. Exercises would likely reveal information about vulnerabilities that an observing adversary could later exploit, or about capabilities or accesses against which an adversary could preemptively develop and employ defensive measures, making them ineffective. Thus, while some scholars such as Paul Meyer have promoted cyber CBMs calling for exchanges of personnel to observe “cybersecurity exercises” (defensive exercises) between potential adversaries, meaningful exchanges of this nature are unrealistic for the cyber domain given the necessary role of secrecy surrounding cyber capabilities and operations.⁵⁶

Nevertheless, there is a role for observation of cyber exercises among allied states. Including allies as observers or even participants in defen-

sive exercises may be collectively beneficial for the purposes of building capacity. It could also demonstrate how an actor intends to respond to and remediate a cyberattack; help allies grow their own cyber defensive infrastructure, to include clarifying national authorities necessary to respond to a crisis; and identify opportunities for allies to augment and complement a state's efforts and enable a unified cyber defense. For instance, for the past 10 years NATO's Cyber Coalition cyber defensive exercise has grown to include over 700 participants from 25 allied countries.⁵⁷

However, rather than solely observing defensive exercises, the spirit behind the exchange of observers in the Helsinki Final Act was to provide reassurance among potential adversaries regarding each other's offensive forces—in other words, those that could pose a threat to stability. However, building offensive cyber operations into existing defensive exercises is fraught with difficulties. Currently, cybersecurity exercises typically have a defensive focus and are used to identify both technical and procedural vulnerabilities on internal networks.⁵⁸ For example, most exercises spearheaded by the United States typically do not showcase the units that would conduct offensive operations or their capabilities and, therefore, are not designed to signal confidence in the command and control and efficacy of their offensive cyber forces.

It is possible to incorporate offensive actions into existing defensive exercises. For instance, a state could build into a defensive scenario a counterstrike that targets an infected server commanding the attack. However, any capability for access and attack that would be used in the scenario would most likely be limited to publicly available open source tools or would be fictionalized so as not to give away to the adversary the specific vulnerability in the target system it would be exploiting. Again, this reflects the fundamental requirement of secrecy for operational success. The ephemeral nature of offensive cyber capabilities and accesses means that revealing information about them effectively renders them moot.⁵⁹ If a state used real cyber weapons from its arsenal, it is likely that any observing state (including allies) would develop hardware and software upgrades to render the demonstrated capability inert. Similar to the paradox presented by cyber arms control, this may undermine the very stability CBMs seek to create. However, public notification of the successful execution of such an exercise could increase the adversary's confidence in the actor's ability to command and control cyber capabilities, thereby serving a confidence building purpose.

The Three Forms of Cyber Stabilization CBMs

Stabilization CBMs under Holst's framework come in three varieties: crisis, political, and arms racing. Crisis stability CBMs involve the exchange of points of contact and defense-related information and are designed to eliminate misperception. Unlike the political and arms racing CBMs, the crisis stability CBMs are cornerstones of the UN GGE and OSCE CBM agreements and are also prominent in several seminal bilateral agreements, as will be discussed in greater detail in the subsequent section.

Political stability CBMs. Achieving mutual consensus around political stability in cyberspace is one of the most significant hurdles for cyber CBMs and accounts for their absence from the current frameworks. The internet has created a relatively cheap and plausibly deniable avenue to undermine the political stability of other states—observed in spades in recent elections in Western democracies. Both authoritarian and democratic regimes view the internet as a medium to influence not only their own but also each other's citizenry. However, while there is some consensus on the utility of cyber capabilities to intervene in the political affairs of other states, there are sharp divisions between states—often reflected in differences in regime type—in terms of how they perceive the role and use of the internet internal to their physical sovereign borders. This tension has implications for stability.⁶⁰ Table 1 highlights the divergent view of the internet internal and external to the state according to regime type, although the latter is an imperfect but useful proxy for this distinction. These differences, we argue, are likely to confound the meaningful development of political stabilization CBMs across dyads of varying regime types.

Political stability CBMs are likely to be confounded by the varying perceptions of the internet internal to state borders on the one hand, and the profligate activities across cyber powers of all regime types to infringe on the sovereignty of their adversaries (or even allies) on the other hand. External to state borders, all major cyber powers perceive a strategic value in using cyber capabilities to conduct shaping operations in support of conventional war fighting and as a tool of coercion, influencing operations, and undermining political stability. The 2016 US presidential election, for instance, exposed how the internet could be used as a vehicle for a state (in this case, Russia) to intervene in the sovereign affairs of another through digital means to achieve strategic

Table 1. Contrasting Approaches to the Internet, by Regime Type

	View of the internet internal to their physical borders	View of the internet external to their physical borders
Authoritarian regimes	<ul style="list-style-type: none"> • Internet censorship and monitoring necessary for state security • States link allowing access to open internet as undermining regime stability 	<ul style="list-style-type: none"> • Need for rigidly defined concept of cyber sovereignty • Internet affords a means to achieve strategic objectives through infringing on sovereignty of others
Democratic regimes	<ul style="list-style-type: none"> • Limited censorship across most democratic regimes; most restrictions deal directly with illicit activities^a • Monitoring of online activity limited by civil liberty protections • Free and open access to the internet is in keeping with democratic ideals 	<ul style="list-style-type: none"> • Access to a free and open internet may be a human right • Internet affords a means to achieve strategic objectives through infringing on sovereignty of others

^aVariations exist among democracies as to the extent and means by which they block fake news and some forms of political speech

objectives.⁶¹ Democratic governments, of course, also conduct information operations.⁶² Democracies perceive a strategic benefit in the spread of democratic principles enabled by the internet.⁶³ For instance, the United States government has invested in the development of anonymity technology through the US State Department’s Bureau of Democracy, Human Rights, and Labor, which historically has sought annual grants for the development of software that contributes to internet freedom.⁶⁴ It is also consistent with the US government spending “approximately \$2 million annually during the past decade to help enable Internet users in China and other Internet restricting countries to access its websites, such as Voice of America and Radio Free Asia.”⁶⁵

Internal to state borders, most democratic states have viewed access to a free and open internet as consistent with broader democratic principles, with some going so far as to define such access as a human right and, therefore, a moral imperative for states to safeguard.⁶⁶ However, there are limits and nuances in these cases, as some democratic governments have taken steps to block or prevent access to illicit content or even limit some forms of political speech. For instance, following Russian interference in the 2016 US presidential election and pervasive information warfare campaigns in Europe, French President Emmanuel Macron advocated for new laws to ban “fake news” during elections, while Germany has enacted new hate speech laws (known as NetzDG) that levy fines on social media companies that fail to remove offensive content.⁶⁷

While some democratic states have enacted measures to limit information on the internet, this stands in stark contrast to how authoritarian

governments view the internet within their borders. The latter perceive an open internet with fundamental suspicion, finding that it encroaches on their sovereign rights and threatens regime survival by undermining state efforts to control the population and by providing a forum for potential dissidents to coordinate and organize against the government. The most notable example of this is China's "Great Firewall," which is integral to the Chinese Communist Party's monitoring and control not only of its citizenry but also of anyone accessing the internet within Chinese borders.⁶⁸ However, other governments, such as those of Russia, Iran, and Turkey, employ similar mechanisms to surveil and control the domestic population. For instance, Russia—particularly in the wake of antigovernment protests in March 2017 that were enabled, in part, by online organizing and activism—attempted to institute limits on domestic access to the internet. The prior year, Russia invited Chinese experts on the Great Firewall to share information and expertise about internet control.⁶⁹

An important wrinkle in the distinction between democratic and authoritarian governments is the role of private Western firms in enabling or collaborating with authoritarian governments to provide capabilities or enforce regulations that support internet control or sharing user information about citizens.⁷⁰ This again reflects the complexities of the multistakeholder nature of the internet. Facebook, for example, has shared user information with China through several data-sharing partnerships with parastatal Chinese electronics firms.⁷¹

Thus, the fact that the internet affords a means to directly reach the citizenry of another state in a way that was not previously possible has complicated the development of political stabilization CBMs. Many authoritarian regimes have moved to block this access through censorship, and many democratic governments struggle with finding policy solutions to thwart external or nefarious interference without sacrificing their democratic ideals. At the same time, all cyber powers benefit from the current ambiguities surrounding violating sovereignty via cyber means. Together, these factors prevent consensus regarding a set of political stabilization CBMs.

Arms racing stability CBMs. Arms racing stability CBMs are similar to more formal arms control agreements in that they typically limit the proliferation of certain technologies, but they are distinct in being entirely voluntary. In cyberspace, the viability of these types of CBMs is

tenuous. Arms racing stability CBMs appeared in the OSCE framework, but they were limited to periodic information exchanges intended to prevent misperceptions that could lead to arms racing behavior—specifically, pressures that encourage increasing forces or capability. Other types of self-imposed limits are unlikely due to the near-universal proliferation of cyber tools. For instance, many offensive tools are publicly available via online forums or for sale on the Dark Web, a section of the internet that is accessible through most web browsers and is known to facilitate illicit transactions.⁷² The source code for Stuxnet as well as US National Security Agency capabilities for surpassing firewalls and other exploit technologies have been compromised and made publicly available by actors such as Shadow Brokers, among others; a tech-savvy actor could learn how to morph these into something even more advanced.⁷³

Additionally, efforts have been made to control the export of information and communications technology that could support offensive operations through amending the Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies to apply to cyberspace. However, the 2013 amendment was quickly met with industry opposition because the technology that supports offensive operations is also necessary to discover vulnerabilities that need to be patched, thus highlighting the offensive and defensive dual-use nature of many cyber security tools.⁷⁴ Indeed, this provision triggered significant resistance from the private sector, which felt it would inevitably and counterproductively lead to greater insecurity by placing restrictions on cybersecurity-related technology and activities, such as penetration testing technology, the sharing of threat information, and the use of multinational computer bug bounty programs.⁷⁵ This represents another example of the challenges of multistakeholder governance. To date, the provisions of ICT technology on cyber security capabilities of the Wassenaar Arrangement are still being refined both collectively by the Wassenaar Plenary and by member countries as they nest domestic regulation with their obligations under the Arrangement. For instance, in response to public feedback, the specific 2013 Wassenaar amendments that covered the training and employment of vulnerability detection systems were never implemented in the United States.⁷⁶ However, the 2016 Plenary relaxed or removed several of the contentious export controls given continued integration of these tools into consumer products.⁷⁷

Since curbing the proliferation of technology is impractical, a potential alternative avenue for consideration would be for states to voluntarily curb the nonstate actors that take part in cyber operations by instituting domestic laws that make such activities illegal. Understandably, CBMs addressing criminal behavior were not part of Holt's framework because crime was perceived to be distinct from national security considerations. However, criminal activity and national security are profoundly interwoven in the cyber domain. States have used and provided safe haven to criminal actors as proxies to conduct plausibly deniable cyber operations at the behest of the state.⁷⁸ For example, in the spring of 2017 the US Department of Justice indicted members of the FSB, one of Russia's intelligence agencies, as well as two hackers who were alleged to have worked with the FSB to steal information from what is now reported to be 3 billion Yahoo user accounts in 2014. The hack was a joint endeavor by an intelligence agency and criminal actors and was carried out for both intelligence and criminal purposes, illustrating the nexus between these two forces.⁷⁹ Additionally, governments have directly engaged in crime via the cyber domain to circumvent economic sanctions or build military and industrial capability through intellectual property theft. North Korea has allegedly netted millions of dollars from cybercrime to evade the crippling effects of economic sanctions including, recently, the WannaCry ransomware attack in the spring of 2017 and financial theft operations targeting banks in the SWIFT network, including the Bank of Bangladesh in 2016 and Taiwan's Far Eastern Bank in 2017.⁸⁰

While there have been ad hoc agreements between states to grapple with certain aspects of criminal activity in cyberspace (notably, the 2015 agreement between the US and China to refrain from economic espionage and intellectual property theft, discussed in greater detail below), there are no CBMs either in the GGE or OSCE lists that directly address cooperation on cybercrime. Most Western states have already institutionalized domestic laws criminalizing illicit cyber activity and have agreed to cooperate on the prevention of cybercrime by becoming signatories to the Budapest Convention on Cybercrime.⁸¹ In contrast, the Russian Federation is the only member of the Council of Europe that has not signed the Budapest Convention.⁸²

Opportunities and Recommendations

Extending the above analysis, we explore potential avenues for cooperation between rivals in cyberspace. Given recent disappointments at multilateral forums for cyber CBMs, we evaluate opportunities for bilateral CBMs when the conditions for effective multilateral CBMs are not met. Finally, we provide specific recommendations for new cyber CBMs.

Bilateral Cyber CBMs

Consistent with the CBM literature, the above discussion has focused primarily on assessing multilateral efforts to develop measures for cyberspace. However, there have been some notable examples of bilateral cyber CBMs outside of the GGE and OSCE, specifically between the US and Russia and the US and China.⁸³ These cases are consistent with the thrust of the analysis above: both of these dyads are cases in which there is a mutually recognized, non-negligible risk of escalation and inadvertent conflict in the domain and, therefore, would benefit from CBMs even as multilateral efforts involving the same countries have failed.

With respect to China and the US, some progress has been made in developing mechanisms that promote transparency and cooperation during peacetime as well as in a crisis. In 2015 Presidents Obama and Xi signed an agreement to abstain from cyber-enabled intellectual property theft for gaining a commercial competitive advantage, to exchange vulnerability and law enforcement information, and to create a working group to further discuss the UN GGE 2015 Report.⁸⁴ While advancing the agenda of the latter was clearly unsuccessful, as evidenced by the failed 2017 GGE summit, the 2015 agreement between the US and China did provide some clarity regarding how each state intends to use the domain (if only within the confines of economics). Furthermore, by mutually agreeing to refrain from economic espionage, the 2015 agreement enabled states to identify potential defections from a pattern of compliance. Most recently, following a meeting between Presidents Trump and Xi at Mar-a-Lago in April 2017, the US and China initiated another round of bilateral talks in October 2017 that reaffirmed the CBMs agreed to in 2015.⁸⁵ However, bilateral agreements have been limited to economic issues rather than political or national security ones. This likely reflects the enduring strategic value both governments perceive in developing cyber capabilities at a relatively low cost/risk for national security purposes. Moreover, the evidence is mixed with respect

to the extent of China's compliance with the 2015 agreement. In the March 2018 *Worldwide Threat Assessment*, the US Director of National Intelligence assessed that Chinese cyberespionage has decreased since the 2015 agreement but noted that "most Chinese cyber operations against US private industry are focused on cleared defense contractors or IT and communications firms whose products and services support government and private sector networks worldwide."⁸⁶ The findings of the March 2018 US Trade Representative report on China are similarly ambiguous about Chinese behavior post-2015.⁸⁷

In a separate landmark agreement, in 2014 the US Department of Defense and the People's Liberation Army allowed the exchange of observers for major military activities and created a military crisis notification system utilizing the Defense Telephone Link between the two countries that was established in 2008.⁸⁸ Though neither of these agreements contained the term "cyber" or "ICT," it was understood at the signing that the catalyst was uncertainty stemming from the potential for inadvertent escalation during a crisis.⁸⁹ Again, however, the fact that this agreement did not directly address notification and observation for cyberspace highlights some of the major hurdles to effective cyber CBMs in these categories as much as it does the opportunities for cooperation. This reflects the delicate balance cyber powers such as the US and China must strike between preventing an inadvertent spiral into an unwanted conflict and protecting cyber assets and capabilities in the event they are needed if the former occurs.

In June 2013, the US and Russia created a working group within the context of the Bilateral Presidential Commission that sought to "promote transparency and reduce the possibility that an incident related to the use of ICTs could unintentionally cause instability or escalation."⁹⁰ Though the United States suspended its participation in the Bilateral Commission following Russia's invasion of Ukraine in 2014, the agreement mentioned three measures of note.⁹¹ First was the continuous sharing of cyber threat information between the US CERT located at the Department of Homeland Security (DHS) and its Russian equivalent. Second was an agreement to utilize the Nuclear Risk Reduction Center (NRRC), first established in 1987, to facilitate inquiries about cybersecurity incidents. In the closing days of the 2016 presidential election, it was reported that the United States used the NRRC to deter Russia from directly interfering with US voting systems.⁹² What is unique about this case is

not that the hotline was used but, rather, that it was used for deterrence rather than for détente. Finally, the commission also created a direct line between the White House's Cybersecurity Coordinator and the Kremlin's Deputy Secretary of the Security Council integrated into the Direct Secure Communications System that, like the NRRC, was first developed to manage nuclear crises during the Cold War. That said, while these are examples of confidence building measures developed with the intent to promote both peacetime and crisis stability, their efficacy remains to be seen. As noted earlier, the Russians succeeded in penetrating the voting systems of several states, although it is not known whether this occurred prior or subsequent to the use of the hotline.⁹³

Specific Recommendations for New Cyber Confidence Building Measures

Based on the framework articulated in this article we identify several potential CBMs that could be adopted. Broadly speaking, these recommendations focus (not exclusively) on promoting stability. While there are non-negligible obstacles to CBM formation, particularly in reference to crisis and arms racing stability, the imperative to prevent unintended conflict escalation and promote crisis stability should compel policy makers to devote energy to this effort. Furthermore, crisis and arms racing stability CBMs are more practical to conceptualize and implement than notification, observation, or political stability measures. We submit the following five areas for CBM creation.

First, as an *a priori* CBM, stakeholders across adversaries and allies should work to build an epistemic community to work toward consensus on key concepts and definitions for cyberspace.

Second, the private sector should be systematically included as an actor in—not simply the subject of—information CBMs. This is particularly relevant for threat actor information CBMs because private actors play a central role in attribution, understanding adversary TTPs and capabilities, and information about their own vulnerabilities.


Third, states could make a commitment to state control of offensive cyber operations. Specifically, a CBM could articulate a concept of command and control (C2) for offensive cyber operations in which offensive operational capabilities remain in the hands of the military, while oversight and launch authorities reside with policy makers. This is similar to what many states have already done with respect to nuclear weapons. Cur-

rently, in cyberspace, many military organizations lack complete control of offensive cyber capabilities. This is due to several factors. First, often states rely upon proxy actors and maintain ambiguous C2 to buttress a government's plausible deniability of offensive cyber operations. Second, due to the often-superior capabilities of private actors, states may rely on civilian industry for expertise and development, or states operating parastatals may depend heavily on cyber espionage for economic growth. Relatedly, some states lack robust indigenous cyber capabilities, personnel, and the resources to produce them and are thus forced to employ cyber proxies to fulfill national security objectives.⁹⁴ This arms racing stability CBM could be built on existing efforts, such as the Budapest Convention, to standardize laws between states for prosecution of cyber crime and other types of nefarious cyber related activity. However, limiting nonstate actors that engage in cyber espionage and offensive operations may only be possible when the perceived risk of escalation outweighs the economic or plausible deniability benefits.

Fourth, and related to above, effort could be dedicated to a measure that addresses the delegation of authorities that each state mandates for the approval of various types of cyber operations. This would assist in understanding what organizations and individuals are behind specific operations, thus adding clarity to attribution efforts. Furthermore, such a measure would assist in building confidence between states that these operations are maintained through a rigid C2 structure.

Finally, states could achieve consensus on an arms racing stability CBM that limits the indiscriminate and mass compromise of a supply chain. States largely agree that espionage is acceptable under customary international law and, therefore, would be reluctant to ascribe to a CBM that limits cyber espionage. However, the mass targeting of a supply chain can be particularly destabilizing, especially if there is a concern that intrusions represent preparations for a cyber attack, rather than simply espionage. An example of this would be if a state maintained a backdoor into every computer that happens to employ a certain brand of antivirus software, or every cell phone manufactured by a specific developer (as allegedly occurred with both Russia and China, respectively).⁹⁵ Beyond the national security concerns, there are implications for international trade if states perceive the need to resist market forces and only purchase software and hardware manufactured domestically or by a trusted ally. While capable cyber powers will likely continue to

seek to disrupt the supply chain to gain access to an adversary, limiting mass (versus tailored) operations through a CBM could enhance stability among cyber rivals.

Creating new cyber CBMs and the continued maintenance of those already in existence is a necessary step toward mitigating the risk of inadvertent conflict in cyberspace. While traditional arms control regimes are unrealistic and ill-suited for managing the risks associated with cyber operations, CBMs that take into account the unique attributes and dynamics of operating in the cyber domain could help to share information, mitigate uncertainty, and facilitate crisis management, thereby promoting much-needed stability between states. 

Notes

1. We are grateful to Robert Jervis and Richard Betts for their extensive and insightful feedback on earlier versions of this article. We are also thankful to individuals at the Army Cyber Institute at the United States Military Academy at West Point; the US Departments of Commerce, State, and Homeland Security; and the French Mission to the 2017 United Nations Group of Governmental Experts for sharing their candid thoughts with us on this topic.

2. For recent work on the cyber security dilemma, see Ben Buchanan, *The Cybersecurity Dilemma: Hacking, Trust, and Fear Between Nations* (Oxford, UK: Oxford University Press, 2017). For a discussion of the debate regarding escalation dynamics in cyberspace, see Erica D. Borghard and Shawn W. Loneragan, “Escalation Dynamics in Cyberspace” (paper presented at the American Political Science Association annual conference, San Francisco, 31 August 2017). Also see Martin Libicki, *Crisis and Escalation in Cyberspace*, RAND Monograph MG-1215-AF (Santa Monica, CA: RAND, 2012); William A. Owens, Kenneth W. Dam, and Herbert S. Lin, eds., “Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities” (Washington, DC: National Academy of Sciences, 2009); Brandon Valeriano, Benjamin Jensen, and Ryan C. Maness, *Cyber Strategy: The Evolving Character of Power and Coercion* (Oxford, UK: Oxford University Press, 2018); Austin Long, “A Cyber SIOP? Operational Considerations for Strategic Offensive Cyber Planning,” *Journal of Cybersecurity* 3, no. 1 (March 2017): 19–28, <https://doi.org/10.1093/cybsec/tyw016>; and Sarah E. Kreps and Jacquelyn Schneider, “Escalation Firebreaks in the Cyber, Conventional, and Nuclear Domains: Moving Beyond Effects-Based Logics” (working paper, SSRN, Rochester, NY, January 2018), <https://ssrn.com/abstract=3104014>.

3. For a discussion of how the failure to establish norms for interstate relations in cyberspace underscores the imperative of constructing CBMs for the domain, see Alex Grigsby, “The End of Cyber Norms,” *Survival: Global Politics and Strategy* 59, no. 6 (December 2017–January 2018): 109–22, <https://doi.org/10.1080/00396338.2017.1399730>.

4. This is not to suggest that arms control regimes and CBMs are always mutually exclusive. In many instances, states pursue formal, institutionalized arms control agreements in tandem with informal, voluntary CBMs. In the context of Cold War strategic rivalries, CBMs often preceded and sometimes evolved into more formal arms control regimes.

5. For further reference on soft law and information institutions, see Kenneth W. Abbott and Duncan Snidal, "Hard and Soft Law in International Governance," *International Organization* 54, no. 3 (Summer 2000): 421–56, <http://www.jstor.org/stable/2601340>; Martha Finnemore and Stephen J. Toope, "Alternatives to 'Legalization': Richer Views of Law and Politics," *International Organization* 55, no. 3 (Summer 2001): 743–58, https://home.gwu.edu/~finnemor/articles/2001_legalization_io.pdf; and Jutta Bruneo and Stephen J. Toope, *Legitimacy and Legality in International Law* (Cambridge, UK: Cambridge University Press, 2010).

6. Jonathan Alford, "Confidence-Building Measures in Europe: The Military Aspects," *Adelphi Papers* 19, no. 149 (1979): 4–13, <https://doi.org/10.1080/05679327908448540>.

7. Kevin N. Lewis and Mark A. Lorell, *The Utility of Confidence-Building Measures in Crisis Situations: Some Case Studies*, RAND Paper P-6947 (Santa Monica, CA: RAND, January 1984), 2–7, <https://www.rand.org/pubs/papers/P6947.html>.

8. In this sense, states act according to Thomas C. Schelling and Morton H. Halperin's "positive-evidence principle," which notes that states are motivated to provide evidence that they are not violating the understanding; Schelling and Halperin, *Strategy and Arms Control*, (Washington, DC: Pergamon-Brassey, 1985), 97–98. However, providing the level of transparency that could completely mitigate the fears of defection is unlikely in cyberspace due to the necessary secrecy that surrounds these operations. CBMs that rely on inspection for compliance are unrealistic.

9. Johan Jørgen Holst and Karen Alette Melander, "European Security and Confidence-Building Measures," *Survival* 19, no. 4 (July/August 1977): 146–54, <https://doi.org/10.1080/00396337708441688>.

10. Johan Jørgen Holst, "Confidence-Building Measures: A Conceptual Framework," *Survival* 25, no. 1 (January/February 1983): 2–15, <https://doi.org/10.1080/00396338308442072>; and Rolf Berg, "Military Confidence-Building in Europe," in *Building Security in Europe: Confidence-Building Measures and the CSCE*, ed. Allen Lynch (New York: Institute for East-West Security Studies, 1986), 13–68.

11. Holst, "Confidence-Building Measures," 3. However, CBMs do not decrease mistrust of an adversary or limit its capabilities; that would require an arms control regime that addresses specific security concerns from two or more parties. Also see Richard E. Darilek, "Reducing the Risks of Miscalculation: The Promise of the Helsinki CBMs," in *Confidence-Building Measures in Europe*, ed. F. Stephen Larrabee and Dietrich Stobbe (New York: Institute for East-West Security Studies, 1983), 59–90.

12. See Richard K. Betts, "Hedging Against Surprise Attack," *Survival* 23, no. 4 (1981): 146–56, <https://doi.org/10.1080/00396338108441973>; and Thomas C. Schelling, "Confidence in Crisis," *International Security* 8, no. 4 (Spring 1984): 55–66, <https://www.jstor.org/stable/2538562>.

13. Holst and Melander, "European Security and Confidence-Building Measures," 148.

14. In addition to Holst, "Confidence-Building Measures," see James Macintosh, "Confidence-Building Measures: A Conceptual Exploration," in *Confidence Building Measures and International Security*, ed. R. B. Byers, F. Stephen Larrabee, and Allen Lynch (New York: Institute for East-West Security Studies, 1987), 9–29, for a conceptual overview of confidence building measures that a state may choose to enact.

15. US State Department, "Conference on Security and Co-operation in Europe Final Act," Department of State Publication 8829, August 1975, 84, <https://www.osce.org/helsinki-final-act?download=true>.

16. The Helsinki Final Act solely dealt with cooperation between the East and West during the Cold War. However, it has been used as the benchmark for other regional security competi-

tions. Similar agreements can be seen in Ariel E. Levite and Emily B. Landau, "Confidence and Security Building Measures in the Middle East," *Journal of Strategic Studies* 20, no. 1 (1997): 143–71, <https://doi.org/10.1080/01402399708437667>; Laurie Nathan, "With Open Arms: Confidence-and Security-Building Measures in Southern Africa," *South African Journal of International Affairs* 1, no. 2 (1994): 110–26, <https://doi.org/10.1080/10220469409545106>; Ralph A. Cossa, *Asia Pacific Confidence and Security Building Measures* (Washington, DC: Center for Strategic & International Studies, 1995); Michael Krepon, Dominique M. McCoy, and Matthew C. J. Rudolph, *A Handbook of Confidence-Building Measures for Regional Security* (Washington, DC: The Henry L. Stimson Center, 1993); and United Nations, Department for Disarmament Affairs, *Confidence and Security-Building Measures—From Europe to Other Regions* (New York: United Nations, 1991).

17. With only one exception between 1975 and 1982, the Soviet Zapad-81 military exercise, all signatories observed the reporting requirements of the original Helsinki Final Act. For a more detailed analysis of the exercises conducted during this period, see Holst, "Confidence-Building Measures," 7–11. However, in more recent times, the Russian Zapad military exercises in the summer of 2017 were said to have violated Cold War agreements. See Michael R. Gordon and Eric Schmitt, "Russia's Military Drills Near NATO Border Raise Fears of Aggression," *New York Times*, 31 July 2017, <https://www.nytimes.com/2017/07/31/world/europe/russia-military-exercise-zapad-west.html>.

18. For the exact language for the requirements of this provision, see State Department, "Conference on Security and Co-operation in Europe Final Act," 85–86.

19. For a concise history of CBMs up to the 1992 Vienna document, see James Macintosh, "Confidence Building Measures in Europe: 1975 to the Present," in *Encyclopedia of Arms Control and Disarmament*, ed. Richard Dean Burns (New York: Charles Scribner's Sons, 1993).

20. The 2011 Vienna document builds on previous agreements, specifically the 1975 Helsinki Final Act, the Document of the Stockholm Conference of 1986, the 1992 Helsinki Document, and the Vienna Documents of 1990, 1992, 1994, and 1999.

21. For a more detailed overview of the 2011 Vienna Document, see Organization for Security and Co-operation in Europe (OSCE), *Vienna Document 2011 on Confidence- and Security-building Measures* (Vienna: OSCE, 30 November 2011), <http://www.osce.org/fsc/86597>.

22. For instance, see Schelling and Halperin, *Strategy and Arms Control*, ix–6; and David W. Kern Jr., *Great Power Security Cooperation: Arms Control and the Challenge of Technological Change* (Lexington, MA: Lexington Books, 2014), 15–48. Examples of arms control agreements designed to minimize instability include the 1963 "Hot Line" Agreement, the 1971 "Accidents Measures" Agreement, and the 1972 Antiballistic Missile (ABM) Treaty. Hedley Bull describes the objectives of arms control across economic, moral, and the international security domains in chapter 1 of his seminal work, *The Control of the Arms Race: Disarmament and Arms Control in the Missile Age*, vol. 2 (London: Praeger for the Institute for Strategic Studies, 1961). Bull's objectives of arms control that relate to inadvertent conflict escalation are of particular relevance to this article. Schelling and Halperin, *Strategy and Arms Control*, 141–42.

23. See Erica D. Borghard and Shawn W. Lonergan, "The Logic of Coercion in Cyberspace," *Security Studies* 26, no. 4 (2017): 452–81, <https://doi.org/10.1080/09636412.2017.1306396>; Adam P. Liff, "Cyberwar: A New 'Absolute Weapon'? The Proliferation of Cyberwarfare Capabilities and Interstate War," *Journal of Strategic Studies* 35, no. 3 (June 2012): 401–28, <http://dx.doi.org/10.1080/01402390.2012.663252>; Ilai Saltzman, "Cyber Posturing and the Offense-Defense Balance," *Contemporary Security Policy* 34, no. 1 (April 2013): 40–63, <https://doi.org/10.1080/13523260.2013.771031>; Rebecca Slayton, "What Is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment," *International Security* 41,

no. 3 (Winter 2016/17): 72–109, https://doi.org/10.1162/ISEC_a_00267; Jon R. Lindsay, “Stuxnet and The Limits of Cyber Warfare,” *Security Studies* 22, no. 3 (2013): 365–404, <https://doi.org/10.1080/09636412.2013.816122>. Additionally, it is difficult to differentiate between offensive and defense expenditures given that, at least from a military perspective, the forces that conduct offensive cyber operations may very well be the same as those that conduct defensive cyber operations.

24. As noted, there is no effective measure of relative offensive cyberpower between states. However, on a subjective scale, certain states (e.g., the United States, China, Russia, France, Great Britain, and Israel) are capable of wreaking widespread havoc against another actor via the offensive use of cyberpower.

25. Chris Wallace, “Gen. Dempsey Reacts to Paris Attacks; Sens. Hoeven, Coons Talk Keystone Showdown,” Fox News Sunday, 11 January 2015, <http://www.foxnews.com/transcript/2015/01/11/gen-dempsey-reacts-paris-attacks-sens-hoeven-coons-talk-keystone-showdown.html>.

26. The existing CERT infrastructure, which fosters a common threat picture for state and nonstate actors alike, reduces some of the uncertainty and vulnerability that exists between actors in this volatile space. However, the CERTs do not help manage crisis escalation nor do they promote transparency beyond the technical threat and security information that participants freely share.

27. William M. Arkin, Ken Dilanian, and Cynthia McFadden, “What Obama Said to Putin on the Red Phone About the Election Hack,” NBC News, 19 December 2016, <http://www.nbcnews.com/news/us-news/what-obama-said-putin-red-phone-about-election-hack-n697116>.

28. Cynthia McFadden, William R. Arkin, and Kevin Monahan, “Russians Penetrated U.S. Voter Systems, Top U.S. Official Says,” NBC News, 8 February 2018, <https://www.nbcnews.com/politics/elections/russians-penetrated-u-s-voter-systems-says-top-u-s-n845721>.

29. For further information on these efforts, see Barbara Rosen Jacobson, Roxana Radu, and Vladimir Radunovic, “Towards a Secure Cyberspace via Regional Cooperation,” *Diplo*, 9 December 2016, <https://www.diplomacy.edu/blog/towards-secure-cyberspace-regional-cooperation>. While these initiatives may eventually contribute to fostering stability in cyberspace within their respective regions, we focus on the UN and OSCE efforts because they are the most mature on promoting cyber confidence building to date.

30. Group of 7 (G7), *G7 Declaration on Responsible States Behavior in Cyberspace* (Taormina, Italy: G7, 11 April 2017), <https://www.mofa.go.jp/files/000246367.pdf>; Group of 20 (G20), *G20 Leaders’ Communique* (Antalya, Turkey: G20, 15–16 November 2015), <http://www.consilium.europa.eu/en/press/press-releases/2015/11/16/g20-summit-antalya-communicue/>.

31. UN General Assembly (UNGA), Resolution 58/32, A/RES/58/32, “Developments in the Field of Information and Telecommunications in the Context of International Security,” 8 December 2003, 47, <http://undocs.org/A/RES/58/32>.

32. UNGA, Report A/65/201, “Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security,” 30 July 2010, <http://undocs.org/A/65/201>.

33. UNGA, Report A/68/98, “Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security,” 24 June 2013, reissued for technical reasons on 30 July 2013, <http://undocs.org/A/68/98>; and UNGA, Report A/70/174, “Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security,” 22 July 2015, <http://undocs.org/A/70/174>.

34. See the appendix for a list of these CBMs. Also see Michele G. Markoff, “Explanation of Position at the Conclusion of the 2016–2017 UN Group of Governmental Experts (GGE) on Developments in the Field of Information and Telecommunications in the Context of International Security” (Washington, DC: US State Department, 23 June 2017), <https://www.state.gov/s/cyberissues/releasesandremarks/272175.htm>. For further commentary, see Adam Segal, “The Development of Cyber Norms at the United Nations Ends in Deadlock. Now What?,” *Council of Foreign Relations-Net Politics*, 29 June 2017, <https://www.cfr.org/blog/development-cyber-norms-united-nations-ends-deadlock-now-what>; and Owen Bowcott, “Dispute Along Cold War Lines Led to Collapse of UN Cyberwarfare Talks,” *The Guardian*, 23 August 2017, <https://www.theguardian.com/world/2017/aug/23/un-cyberwarfare-negotiations-collapsed-in-june-it-emerges>; and Arun M. Sukumar, “The UN GGE Failed. Is International Law in Cyberspace Doomed as Well?,” *Lawfare* (blog), 4 July 2017, <https://www.lawfareblog.com/un-gge-failed-international-law-cyberspace-doomed-well>.

35. See, for example, Finnemore and Toope, “Alternatives to ‘Legalization’; Bruneo and Toope, *Legitimacy and Legality*; and Wayne Sandholtz, “Dynamics of International Norm Change: Rules against Wartime Plunder,” *European Journal of International Relations* 14, no. 1 (March 2008): 101–31, <http://journals.sagepub.com/doi/10.1177/1354066107087766>.

36. Organization for Security and Co-operation in Europe, *Permanent Council Decision no. 1039*, 26 April 2012, <http://www.osce.org/pc/90169>.

37. Organization for Security and Co-operation in Europe (OSCE), *Permanent Council Decision*, no. 1106, 3 December 2013, <http://www.osce.org/pc/109168>; OSCE, *Permanent Council Decision*, no. 1202, 10 March 2016, <http://www.osce.org/pc/227281>. However, also see Transnational Threat Department, “Cyber/ICT Security,” OSCE Secretariat, <http://www.osce.org/secretariat/256071?download=true>; and Lamberto Zannier, “Cyber/ICT Security: Building Confidence,” *Security Community* 2 (June 2014): 4–5, <http://www.osce.org/magazine/112525>.

38. For instance, the need for cyber CBMs has been noted by James Andrew Lewis, “Confidence-Building and International Agreement in Cybersecurity,” *Disarmament Forum: Confronting Cyberconflict* 4 (Geneva: UN Institute for Disarmament Research, 2011), 51–60, <http://www.unidir.org/files/publications/pdfs/confronting-cyberconflict-en-317.pdf>.

39. Herbert Lin, “Arms Control in Cyberspace: Challenges and Opportunities,” *World Politics Review* (website), 6 March 2012, <https://www.worldpoliticsreview.com/articles/11683/arms-control-in-cyberspace-challenges-and-opportunities>.

40. Tughral Yamin, *Cyberspace CBMs between Pakistan and India* (Islamabad: National University of Sciences and Technology, 2014), 102.

41. Jason Healey, John C. Mallery, Klara Tothova Jordan, and Nathaniel V. Youd, *Confidence-Building Measures in Cyberspace: A Multistakeholder Approach for Stability and Security* (Washington, DC: Atlantic Council, November 2014), http://www.atlanticcouncil.org/images/publications/Confidence-Building_Measures_in_Cyberspace.pdf.

42. Holst, “Confidence-Building Measures,” 4–5.

43. Holst, “Confidence-Building Measures,” 4.

44. For literature on multistakeholder governance, see Milton L. Mueller, *Networks and States: The Global Politics of Internet Governance* (Cambridge, MA: MIT Press, 2010); Tana Johnson, *Organizational Progeny: Why Governments Are Losing Control over Proliferating Structures of Global Governance* (Oxford, UK: Oxford University Press, 2014); Kenneth W. Abbott, Jessica F. Green, and Robert O. Keohane, “Organizational Ecology and Institutional Change in Global Governance,” *International Organization* 70, no. 2 (Spring 2016): 247–77, <https://doi.org/10.1017/S0020818315000338>.

45. The CBMs in the appendix are classified by an updated version of John Holst's CBM classification guidelines.

46. See appendix, table 3: December 2013 OSCE—CBM 6.

47. Forum of Incident Response and Security Teams (FIRST), "FIRST Teams," accessed 25 June 2018, <https://www.first.org/members/teams>.

48. See appendix, table 2: July 2015 GGE—CBM D.

49. See the US National Institute of Standards and Technology (NIST) National Vulnerability Database, <https://nvd.nist.gov/>.

50. For example, see GGE Measure 3 and OSCE Measure 7, in appendix, table 1 and table 3, respectively.

51. See appendix, table 1: July 2015 GGE—CBM 3.

52. See Martha Finnemore and Duncan B. Hollis, "Constructing Norms for Global Cybersecurity," *American Journal of International Law* 110, no. 3 (July 2016): 425–79, <https://doi.org/10.1017/S0002930000016894>; Emanuel Adler and Vincent Pouliot, "International Practices," *International Theory* 3, no. 1 (February 2011): 1–36, <https://doi.org/10.1017/S175297191000031X>; and Emanuel Adler and Michael Barnett, eds., *Security Communities* (Cambridge, UK: Cambridge University Press, 1998).

53. Though beyond the scope of this article, there are signs of states developing dramatically different strategies for operating militarily in the cyber domain that may reflect variation in strategic culture within those states. For instance, Russian interference in the 2016 US presidential election is consistent with the Soviet Union's approach to information warfare employed throughout the Cold War (Dov H. Levin, "Partisan Electoral Interventions by the Great Powers: Introducing the PEIG Dataset," *Conflict Management and Peace Science* 33, no. 4 [September 2016], <https://doi.org/10.1177/0738894216661190>). North Korea has reportedly benefited from financially motivated cybercrime to generate capital and circumvent economic sanctions (Paul Mozur and Choe Sang-Hun, "North Korea's Rising Ambition Seen in Bid to Breach Global Banks," *New York Times*, 25 March 2017, <https://www.nytimes.com/2017/03/25/technology/north-korea-hackers-global-banks.html>). China has used cyber espionage to steal industrial and intellectual property to grow its economy (Office of the National Counterintelligence Executive, *Foreign Spies Stealing US Economic Secrets in Cyberspace: Report to Congress on Foreign Economic Collection and Industrial Espionage, 2009–2011* [Washington, DC: Director of National Intelligence, October 2011]). Finally, the United States has historically pressed for the "rules based order" that exists in the physical domains of strategic interaction to apply to cyberspace (for instance, see Dan Seifert, "President Obama Wants to Prevent a Cyber Weapon 'Arms Race,'" *The Verge*, 5 September 2016, <https://www.theverge.com/2016/9/5/12798836/president-obama-prevent-cyber-weapon-arms-race>; Harold Hongju Koh, "International Law in Cyberspace," *Faculty Scholarship Series*, Paper 4854 [2012], http://digitalcommons.law.yale.edu/fss_papers/4854/; and International Cyberspace Policy Strategy, Public Law 114-113, Division N, Title IV, Section 402, US Department of State, March 2016, <https://www.state.gov/documents/organization/255732.pdf>).

54. See Martha Finnemore and Kathryn Sikkink, "International Norm Dynamics and Political Change," *International Organization* 52, no. 4 (Autumn 1998): 887–917, <https://www.jstor.org/stable/2601361>.

55. See appendix, table 3: December 2013 OSCE—CBMs 10, 11.

56. Paul Meyer, "Cyber-Security through Arms Control: An Approach to International Cooperation," *RUSI Journal* 156, no. 2 (2011): 22–27, <https://doi.org/10.1080/03071847.2011.576471>. Note that Yamin, *Cyberspace CBMs between Pakistan and India*, 108–12, also takes a rudimentary step toward identifying cyber CBMs that could be established in a bilateral context;

however, many of the measures fall under the convention of norm creation, such as “avoiding hostile propaganda,” and have limited applicability beyond the India-Pakistan relationship. For more on the role of secrecy surrounding cyber operations, see Shawn W. Loneragan, “Cyber Power and the International System” (PhD diss., Columbia University, 2017), 4–7, <https://doi.org/10.7916/D88D07PH>.

57. “NATO’s Flagship Cyber Exercise Begins in Estonia,” NATO, 4 December 2017, https://www.nato.int/cps/ic/natohq/news_149233.htm.

58. For example, Cyber Guard is an annual defensive exercise cohosted by the US Cyber Command, the Federal Bureau of Investigation, and the Department of Homeland Security that incorporates the private sector and numerous governmental organizations to respond to a cyberattack against US public and private critical infrastructure. The exercise is focused on identifying existing technical vulnerabilities, developing and testing response actions, and improving defense of DOD information systems. For further information, see Mark Pomerleau, “Cyber Forces Prepare for Attack on a Grand Scale,” *Defense Systems*, 20 June 2016, <https://defensesystems.com/articles/2016/06/20/cyber-guard-dod-civilian-industry-exercise.aspx>. However, US Cyber Command does have exercises, such as its annual Cyber Flag, which incorporate offensive capabilities and involve multinational coalition partners. During this exercise the offensive response actions may be fictionalized or involve real capabilities. However, despite the multinational representation at the exercise, the cyber weaponry used is employed on isolated test networks, and only those with appropriate security clearances are aware of the specific capabilities involved. In addition to promoting a common understanding of response actions between partners, the exercise, as is the case with Cyber Guard, is widely used by the military to validate the proficiency of cyber teams. Chief Warrant Officer Judy Esquibel (Army Cyber Institute), interview by the author, 30 April 2017.

59. Max Smeets, “A Matter of Time: On the Transitory Nature of Cyberweapons,” *Journal of Strategic Studies* 41, nos. 1–2 (2018): 6–32, <https://doi.org/10.1080/01402390.2017.1288107>.

60. Though beyond the scope of this article, there is a growing body of literature addressing the differences in internet policy across regime type. Most authoritarian states have instituted hierarchies in their intrastate internet infrastructure to prevent their citizens from accessing prohibited material. However, Western democracies (for the most part) have pursued a free and open internet that is largely devoid of state censorship. These conflicting visions for the internet were evident during the 2012 breakdown of the United Nations International Telecommunications Union’s World Conference on International Communication (WCIT), when China and Russia used the Arab Spring to get support from many Middle Eastern countries to push for a treaty that limited the openness of the internet and created restrictions on free speech. In response, most Western democracies refused to ratify the treaty. This divide has given rise to extensive debates about internet governance, state sovereignty in cyberspace, and the “Balkanization” of the internet. See James D. Fielder, “The Internet and Dissent in Authoritarian States,” in *Conflict and Cooperation in Cyberspace: The Challenge to National Security*, ed. Panayotis A. Yannakogeorgos and Adam B. Lowther (New York: Taylor & Francis, 2013), 161–94; Daniel W. Drezner, “The Global Governance of the Internet: Bringing the State Back In,” *Political Science Quarterly* 119, no. 3 (Fall 2004): 477–98, <https://doi.org/10.2307/20202392>; Stephen K. Gourley, “Cyber Sovereignty,” in *Conflict and Cooperation in Cyberspace*, 277–90; Michael N. Schmitt, ed., *Tallinn Manual on the International Law Applicable to Cyber Warfare* (Cambridge: UK: Cambridge University Press, 2013); Dana Polatin-Reuben and Joss Wright, “An Internet with BRICS Characteristics: Data Sovereignty and the Balkanisation of the Internet” (paper presented at the 4th USENIX Workshop on Free and Open Communications on the Internet, San Diego, CA, 18 August 2014). For a comprehensive report on state censorship

by country, see Sanja Kelly, Madeline Earp, Laura Reed, Adrian Shahbaz, and Mai Truong, *Privatizing Censorship, Eroding Privacy: Freedom on the Net 2015* (New York: Freedom House, October 2015), <https://freedomhouse.org/sites/default/files/FOTN%202015%20Full%20Report.pdf>.

61. A comprehensive report on Russian interference into the 2016 US Presidential election has yet to be written. However, there are multiple ongoing US government investigations across several agencies, as well as investigative journalists doing the same. Moreover, there is consensus within the US intelligence community regarding Russian interference of the election. The most comprehensive assessment to date is the Office of the Director of National Intelligence, National Intelligence Council, “Assessing Russian Activities and Intentions in Recent US Elections,” ICA 2017-01D, 6 January 2017, https://www.dni.gov/files/documents/ICA_2017_01.pdf.

62. For a good discussion on the Western view of information operations and public diplomacy, see John M. Wilson, “The Hunting of the Snark: Organizing and Synchronizing of Informational Elements for Homeland Defense and Civil Support” (master’s thesis, Naval Postgraduate School, Monterey, CA, June 2009), <http://www.dtic.mil/dtic/tr/fulltext/u2/a501553.pdf>; and Leigh Armistead, ed. *Information Operations: Warfare and the Hard Reality of Soft Power* (Lincoln, NE: Potomac Books, 2004).

63. Though beyond the scope of this article, this view is in line with the literature on the Democratic Peace Theory that holds that democracies are more likely to be peaceful with one another because of shared culture, norms, and structural mechanisms that promote peaceful conflict resolution. Additionally, scholarship has shown that common ideologies encourage alliance formation whereas divergent ideologies can have a positive effect on threat perception and domestic stability. Indeed, authoritarian regimes view access to the unfettered internet as a venue that can encourage civil unrest because it exposes their citizenry to differing ideological thought and serves as a venue for likeminded individuals to assemble in relative safety. This view is rational given the role social media played during the Arab Spring, which resulted in multiple regime changes. Some Western policy makers and strategists may hold that by keeping a free and open internet, democracies can spread democratic values in hopes that it will encourage democratic revolutions in authoritarian regimes—thus resulting in a strategic victory and an international environment where they face fewer threats and potentially new allies.

64. US State Department Bureau of Democracy, Human Rights, and Labor, “Internet Freedom Annual Program Statement,” 2 June 2014, <https://2009-2017.state.gov/j/drl/p/previous/calls/227048.htm>. As an interesting aside, the development of the most commonly used anonymity software available, The Onion Router (commonly known as “Tor”), was sponsored by the Defense Advanced Research Projects Agency (DARPA) and the US Navy’s Office of Naval Research (ONR).

65. Thomas Lum, Patricia Moloney Figliola, and Matthew C. Weed, “China, Internet Freedom, and U.S. Foreign Policy,” *Congressional Research Service*, 13 July 2012.

66. UNGA, Report A/HRC/17/27, “Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, Frank La Rue,” 16 May 2011, <http://undocs.org/A/HRC/17/27>; UNGA Human Rights Council, A/HRC/32/L.20, “The Promotion, Protection and Enjoyment of Human Rights on the Internet,” 27 June 2016, <http://undocs.org/A/HRC/32/L.20>. However, the view that access to the internet is a human right remains hotly contested. For further reference, see Daniel Joyce, “Internet Freedom and Human Rights,” *European Journal of International Law* 26, no. 2 (May 2015): 493–514, <https://doi.org/10.1093/ejil/chv021>; Vinton G. Cerf, “Internet Access Is Not a Human Right,” *New York Times*, 4 January 2012, <https://www.nytimes.com/2012/01/05/opinion/internet-access-is-not-a-human-right.html>; and Michael L. Best, “Can the Internet Be a Human Right?,” *Human Rights*

Human Welfare 4, no. 1 (2004): 23–31, <https://www.du.edu/korbel/hrhw/volumes/2004/best-2004.pdf>.

67. Angélique Chrisafis, “Emmanuel Macron Promises Ban on Fake News During Elections,” *The Guardian*, 3 January 2018, <https://www.theguardian.com/world/2018/jan/03/emmanuel-macron-ban-fake-news-french-president>; Mark Scott and Janosch Delcker, “Free Speech vs. Censorship in Germany,” *Politico*, 4 January 2018, <https://www.politico.eu/article/germany-hate-speech-netzdg-facebook-youtube-google-twitter-free-speech/>.

68. For instance, see Gary King, Jennifer Pan, and Margaret E. Roberts, “How Censorship in China Allows Government Criticism but Silences Collective Expression,” *American Political Science Review* 107, no. 2 (2013): 326–43, <https://doi.org/10.1017/S0003055413000014>; Chris C. Demchak and Peter Dombrowski, “Rise of a Cybered Westphalian Age,” *Strategic Studies Quarterly* (Spring 2011): 32–61, http://www.airuniversity.af.mil/Portals/10/SSQ/documents/Volume-05_Issue-1/Demchak-Dombrowski.pdf; and Fielder, “Internet and Dissent in Authoritarian State.”

69. Emily Parker, “Russia Is Trying to Copy China’s Approach to Internet Censorship,” *Slate*, 4 April 2017, http://www.slate.com/articles/technology/future_tense/2017/04/russia_is_trying_to_copy_china_s_internet_censorship.html.

70. Ronald J. Deibert, *Black Code: Surveillance, Privacy, and the Dark Side of the Internet* (Oxford, UK: Signal, 2013).

71. Michael LaForgia and Gabriel J. X. Dance, “Facebook Gave Data Access to Chinese Firms Flagged by U.S. Intelligence,” *New York Times*, 5 June 2018, <https://www.nytimes.com/2018/06/05/technology/facebook-device-partnerships-china.html>.

72. For further information on the Dark Web marketplace, see Paul Stockton and Michele Golabek-Goldman, “Curbing the Market for Cyber Weapons,” *Yale Law & Policy Review* 32, no. 1 (2013): 239–66, <http://digitalcommons.law.yale.edu/ylpr/vol32/iss1/11/>.

73. For instance, see “Stuxnet Source Code Released Online, Download Now,” *The Hacker News*, 2 July 2011, <http://thehackernews.com/2011/07/stuxnet-source-code-released-online.html>; and Ellen Nakashima, “Powerful NSA Hacking Tools Have Been Revealed Online,” *Washington Post*, 16 August 2016, https://www.washingtonpost.com/world/national-security/powerful-nsa-hacking-tools-have-been-revealed-online/2016/08/16/bce4f974-63c7-11e6-96c0-37533479f3f5_story.html?noredirect=on&utm_term=.9e477b899afb.

74. “Wassenaar: Cybersecurity and Export Control,” testimony before the US House of Representatives Subcommittee on Information Technology (Washington, DC: 12 January 2016), <https://oversight.house.gov/hearing/wassenaar-cybersecurity-and-export-control/>; and Trey Herr, *Malware Counter-Proliferation and the Wassenaar Arrangement, Proceedings of the 8th International Conference on Cyber Conflict* (Tallinn, Estonia, 4 January 2016), <https://dx.doi.org/10.2139/ssrn.2711070>.

75. Bug bounty programs offer financial rewards for “white hat” hackers who find vulnerabilities. Russell Brandom, “Google Says Controversial Exports Proposal Would Make the World ‘Less Secure,’” *The Verge*, 20 July 2015, <http://www.theverge.com/2015/7/20/9005351/google-wassenaar-arrangement-proposal-comments>; and Chris Bream, Facebook U.S. Public Policy, “Wassenaar Rules Are Not the Right Direction,” Facebook, 28 July 2015, <https://www.facebook.com/uspublicpolicy/posts/1047027321981746>. For a more comprehensive list of the public feedback the US Department of Commerce received regarding this regulation, see “Public Comments for Wassenaar Arrangement 2013 Plenary Agreements Implementation: Intrusion and Surveillance Items,” *Regulations.gov*, comment period closed 20 July 2015, <https://www.regulations.gov/docketBrowser?rpp=50&so=DESC&sb=postedDate&po=0&dct=PS&D=BIS-2015-0011>.

76. Senior official (US Department of Commerce, Bureau of Industry and Security), interview with the author, 28 November 2016. Specifically, Wassenaar Arrangement Category 4 rules 4.A.5, 4.D.4, and 4.E.1.C were never added to the Commerce Control List as elucidated in the department's Export Administration Regulations in either 2014 or 2015 following the 2013 amendments. The concern here was that technology that is used to detect vulnerabilities within a network could also be used to exploit the network as an essential part of gaining access. This highlights the difficulties of dual-use technologies in this space.

77. "Summary of Changes: List of Dual-Use Goods & Technologies and Munitions List," Wassenaar Arrangement (website), 17 February 2017, <https://www.wassenaar.org/app/uploads/2017/03/Summary-of-Changes-to-2016-Lists.pdf>; also see Rainer Himmelfreund-pointner, "Le Monde Wassenaar Arrangement," *Cercle Diplomatique*, Issue 1/2017, 62–66, <https://www.yumpu.com/en/document/view/57005074/cercle-diplomatique-issue-01-2017>.

78. For an in-depth discussion on the motivation of authoritarian states to employ cyber proxies, see Erica D. Borghard and Shawn W. Lonergan, "Can States Calculate the Risks of Using Cyber Proxies?" *Orbis* 60, no. 3 (Summer 2016): 395–416, <https://doi.org/10.1016/j.orbis.2016.05.009>.

79. Ellen Nakashima, "Justice Department Charges Russian Spies and Criminal Hackers in Yahoo Intrusion," *Washington Post*, 15 March 2017, https://www.washingtonpost.com/world/national-security/justice-department-charging-russian-spies-and-criminal-hackers-for-yahoo-intrusion/2017/03/15/64b98e32-0911-11e7-93dc-00f9bdd74ed1_story.html?utm_term=.e0993e6a74ea; and Alina Selyukh, "Every Yahoo Account that Existed in Mid-2013 Was Likely Hacked," NPR, 3 October 2017, <https://www.npr.org/sections/thetwo-way/2017/10/03/555016024/every-yahoo-account-that-existed-in-mid-2013-was-likely-hacked>.

80. Mozur and Sang-Hun, "North Korea's Rising Ambition"; Ellen Nakashima, "The NSA Has Linked the WannaCry Computer Worm to North Korea," *Washington Post*, 14 June 2017, https://www.washingtonpost.com/world/national-security/the-nsa-has-linked-the-wannacry-computer-worm-to-north-korea/2017/06/14/101395a2-508e-11e7-be25-3a519335381c_story.html?utm_term=.5eb85debb0f8; and Jim Finkle, "North Korea likely behind Taiwan SWIFT cyber heist—BAE," Reuters, 16 October 2017, <https://www.reuters.com/article/cyber-heist-north-korea-taiwan/north-korea-likely-behind-taiwan-swift-cyber-heist-bae-idUSL2N1MRIQC>.

81. Council of Europe, "European Treaty Series no. 185, Convention on Cybercrime" (Budapest: 23 November 2001), <https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185>.

82. Chart of Signatures and Ratifications of Treaty 185, Council of Europe, accessed 6 June 2017, <https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185/signatures>. However, it is important to note that in April 2017 Russia put forward its own convention on cybercrime to the United Nations as a direct replacement to the Budapest Convention. For further information, see "Russia Prepares New UN Anti-Cybercrime Convention—Report," RT, 14 April 2017, <https://www.rt.com/politics/384728-russia-has-prepared-new-international/>.

83. Though these bilateral agreements are examples between major state powers, there are other agreements between regional powers, particularly on cybersecurity coordination efforts, that are beyond the scope of this article.

84. The White House, "Fact Sheet: President Xi Jinping's State Visit to the United States," Office of the Press Secretary, 25 September 2015, <https://www.whitehouse.gov/the-press-office/2015/09/25/fact-sheet-president-xi-jinpings-state-visit-united-states>. Note that intellectual property theft

that supports national security objectives is still permissible for espionage, for it is considered a necessary state practice by Customary International Law.

85. “First U.S.-China Law Enforcement and Cybersecurity Dialogue,” Department of Homeland Security, <https://www.dhs.gov/news/2017/10/06/first-us-china-law-enforcement-and-cybersecurity-dialogue>.

86. Daniel R. Coats, Director of National Intelligence, *Statement for the Record: Worldwide Threat Assessment of the US Intelligence Community*, 115th Cong., 2nd sess., 6 March 2018, 6, <https://www.dni.gov/index.php/newsroom/congressional-testimonies/item/1851-statement-for-the-record-worldwide-threat-assessment-of-the-us-intelligence-community>.

87. Office of the United States Trade Representative, “Findings of the Investigation into China’s Act, Policies, and Practices Related to Technology Transfer, Intellectual Property, and Innovation under Section 301 of the Trade Act of 1974” (Washington, DC: Executive of the President, 22 March 2018).

88. Secretary of Defense, “Memorandum of Understanding Between the United States of America Department of Defense and the People’s Republic of China Ministry of National Defense on Notification of Major Military Activities Confidence-Building Measures Mechanism,” 4 November 2014, http://archive.defense.gov/pubs/141112_MemorandumOfUnderstandingOnNotification.pdf; “Military Crisis Notification Mechanism for Use of the Defense Telephone Link,” in US-China Crisis Communications, September 2015, http://www.defense.gov/portals/1/documents/pubs/us-china_crisis_communications_annex_sep_2015.pdf.

89. Senior official (United States Department of Homeland Security), interview by the author, 14 July 2016.

90. US State Department, *U.S.-Russia Bilateral Presidential Commission: 2013 Joint Annual Report* (2013), <http://www.state.gov/p/eur/ci/rs/usrussiabilat/219086.htm#8>.

91. The White House, “Fact Sheet: U.S.-Russian Cooperation on Information and Communications Technology Security,” Office of the Press Secretary, 17 June 2013, <https://www.whitehouse.gov/the-press-office/2013/06/17/fact-sheet-us-russian-cooperation-information-and-communications-technol>.

92. David Ignatius, “In Our New Cold War, Deterrence Should Come before Détente,” *Washington Post*, 15 November 2016, https://www.washingtonpost.com/opinions/global-opinions/in-our-new-cold-war-deterrence-should-come-before-detente/2016/11/15/051f4a84-ab79-11e6-8b45-f8e493f06fcd_story.html?utm_term=.0f14161b193c.

93. Office of the Press Secretary, “DHS Statement on NBC News Coverage of Election Hacking,” Department of Homeland Security, 12 February 2018, <https://www.dhs.gov/news/2018/02/12/dhs-statement-nbc-news-coverage-election-hacking>.

94. In addition to hiring cyber proxies due to a dearth of localized talent, states may also employ them for plausible deniability. For further reference to the state-proxy exchange, see Borghard and Loneragan, “Can States Calculate the Risks of Using Cyber Proxies?” Note that the authors discuss the risks states face when they chose to employ a cyber proxy.

95. See Nicole Perlroth and Scott Shane, “How Israel Caught Russian Hackers Scouring the World for U.S. Secrets,” *New York Times*, 10 October 2017, <https://www.nytimes.com/2017/10/10/technology/kaspersky-lab-israel-russia-hacking.html>; and Matt Apuzzo and Michael S. Schmidt, “Secret Back Door in Some U.S. Phones Sent Data to China, Analysts Say,” *New York Times*, 15 November 2016, <https://www.nytimes.com/2016/11/16/us/politics/china-phones-software-security.html>.

Appendix

Table A.1. United Nations General Assembly (UNGA) recommended confidence building measures (CBM) on 22 July 2015

Recommended CBM	CBM classification
1. The identification of appropriate points of contact at the policy and technical levels to address serious information and communications technology (ICT) incidents and the creation of a directory of such contacts;	Stability-crisis
2. The development of and support for mechanisms and processes for bilateral, regional, subregional, and multilateral consultations, as appropriate, to enhance inter-state confidence-building and to reduce the risk of misperception, escalation and conflict that may stem from ICT incidents;	Stability-arms race
3. Encouraging, on a voluntary basis, transparency at the bilateral, subregional, regional and multilateral levels, as appropriate, to increase confidence and inform future work. This could include the voluntary sharing of national views and information on various aspects of national and transnational threats to and in the use of ICTs; vulnerabilities and identified harmful hidden functions in ICT products; best practices for ICT security; confidence-building measures developed in regional and multilateral forums; and national organizations, strategies, policies and programmes relevant to ICT security;	Information-use, threat actor, and security
4. The voluntary provision by states of their national views of categories of infrastructure that they consider critical and national efforts to protect them, including information on national laws and policies for the protection of data and ICT-enabled infrastructure. States should seek to facilitate cross-border cooperation to address critical infrastructure vulnerabilities that transcend national borders. These measures could include:	Information-use
– a repository of national laws and policies for the protection of data and ICT-enabled infrastructure and the publication of materials deemed appropriate for distribution on these national laws and policies;	
– the development of mechanisms and processes for bilateral, subregional, regional and multilateral consultations on the protection of ICT-enabled critical infrastructure;	
– the development on a bilateral, subregional, regional and multilateral basis of technical, legal and diplomatic mechanisms to address ICT-related requests;	
– the adoption of voluntary national arrangements to classify ICT incidents in terms of the scale and seriousness of the incident, for the purpose of facilitating the exchange of information on incidents.	

Table A.2. UNGA-recommended additional CBMs on a bilateral, subregional, regional, and multilateral basis

Recommended CBM	CBM classification
A. Strengthen cooperative mechanisms between relevant agencies to address ICT security incidents and develop additional technical, legal and diplomatic mechanisms to address ICT infrastructure-related requests, including the consideration of exchanges of personnel in areas such as incident response and law enforcement, as appropriate, and encouraging exchanges between research and academic institutions;	Information-security
B. Enhance cooperation, including the development of focal points for the exchange of information on malicious ICT use and the provision of assistance in investigations;	Information-security
C. Establish a national computer emergency response team and/or cybersecurity incident response team or officially designate an organization to fulfill this role. States may wish to consider such bodies within their definition of critical infrastructure. States should support and facilitate the functioning of and cooperation among such national response teams and other authorized bodies;	Information-threat actor and security Stability-crisis
D. Expand and support practices in computer emergency response team and cybersecurity incident response team cooperation, as appropriate, such as information exchange about vulnerabilities, attack patterns and best practices for mitigating attacks, including coordinating responses, organizing exercises, supporting the handling of ICT-related incidents and enhancing regional and sector-based cooperation;	Information-threat actor and security Stability-crisis
E. Cooperate, in a manner consistent with national and international law, with requests from other states in investigating ICT-related crime or the use of ICTs for terrorist purposes or to mitigate malicious ICT activity emanating from their territory.	Information-threat actor and security

Table A.3. CBMs adopted through OSCE Permanent Council Decision no. 1106 on 3 December 2013

Recommended CBM	CBM classification
1. Participating states will voluntarily provide their national views on various aspects of national and transnational threats to and in the use of ICTs. The extent of such information will be determined by the providing parties.	Information-security
2. Participating states will voluntarily facilitate co-operation among the competent national bodies and exchange of information in relation with security of and in the use of ICTs.	Information- use and security
3. Participating states will on a voluntary basis and at the appropriate level hold consultations in order to reduce the risks of misperception, and of possible emergence of political or military tension or conflict that may stem from the use of ICTs, and to protect critical national and international ICT infrastructures including their integrity.	Stability-arms race
4. Participating states will voluntarily share information on measures that they have taken to ensure an open, interoperable, secure, and reliable internet.	Information-use

Table A.3. CBMs adopted through OSCE Permanent Council Decision no. 1106 on 3 December 2013 *(continued)*

Recommended CBM	CBM classification
5. The participating states will use the OSCE as a platform for dialogue, exchange of best practices, awareness-raising and information on capacity-building regarding security of and in the use of ICTs, including effective responses to related threats. The participating states will explore further developing the OSCE role in this regard.	Administrative
6. Participating states are encouraged to have in place modern and effective national legislation to facilitate on a voluntary basis bilateral co-operation and effective, time-sensitive information exchange between competent authorities, including law enforcement agencies, of the participating states in order to counter terrorist or criminal use of ICTs. The OSCE participating states agree that the OSCE shall not duplicate the efforts of existing law enforcement channels.	Information-threat actor
7. Participating states will voluntarily share information on their national organization; strategies; policies and programmes – including on co-operation between the public and the private sector; relevant to the security of and in the use of ICTs; the extent to be determined by the providing parties.	Information-use
8. Participating states will nominate a contact point to facilitate pertinent communications and dialogue on security of and in the use of ICTs. Participating states will voluntarily provide contact data of existing official national structures that manage ICT-related incidents and co-ordinate responses to enable a direct dialogue and to facilitate interaction among responsible national bodies and experts. Participating states will update contact information annually and notify changes no later than thirty days after a change has occurred. Participating states will voluntarily establish measures to ensure rapid communication at policy levels of authority, to permit concerns to be raised at the national security level.	Stability-crisis, Information-use
9. In order to reduce the risk of misunderstandings in the absence of agreed terminology and to further a continuing dialogue, participating states will, as a first step, voluntarily provide a list of national terminology related to security of and in the use of ICTs accompanied by an explanation or definition of each term. Each participating state will voluntarily select those terms it deems most relevant for sharing. In the longer term, participating states will endeavor to produce a consensus glossary.	Information-use
10. Participating states will voluntarily exchange views using OSCE platforms and mechanisms inter alia, the OSCE communications network, maintained by the OSCE secretariat's Conflict Prevention Centre, subject to the relevant OSCE decision, to facilitate communications regarding the CBMs.	Administrative
11. Participating states will, at the level of designated national experts, meet at least three times each year, within the framework of the security committee and its informal working group established by permanent council decision no. 1039 to discuss information exchanged and explore appropriate development of CBMs. Candidates for future consideration by the IWG may include inter alia proposals from the consolidated list circulated by the chairmanship of the IWG under PC.DEL/682/12 on 9 July 2012, subject to discussion and consensus agreement prior to adoption.	Administrative

Table A.4. CBMs adopted through OSCE Permanent Council Decision no. 1202 on 10 March 2016

Recommended CBM	CBM classification
<p>12. Participating states will, on a voluntary basis, share information and facilitate inter-state exchanges in different formats, including workshops, seminars, and roundtables, including on the regional and/or subregional level; this is to investigate the spectrum of co-operative measures as well as other processes and mechanisms that could enable participating states to reduce the risk of conflict stemming from the use of ICTs. Such activities should be aimed at preventing conflicts stemming from the use of ICTs and at maintaining peaceful use of ICTs.</p>	Information-security
<p>With respect to such activities participating states are encouraged, inter alia, to:</p> <ul style="list-style-type: none"> – conduct such activities in the spirit of enhancing inter-state cooperation, transparency, predictability and stability; – complement, through such activities, un efforts and avoid duplicating work done by other fora; and – take into account the needs and requirements of participating states taking part in such activities. <p>Participating states are encouraged to invite and engage representatives of the private sector, academia, centres of excellence and civil society in such activities.</p>	
<p>13. Participating states will, on a voluntary basis, conduct activities for officials and experts to support the facilitation of authorized and protected communication channels to prevent and reduce the risks of misperception, escalation, and conflict; and to clarify technical, legal and diplomatic mechanisms to address ICT-related requests. This does not exclude the use of the channels of communication mentioned in Permanent Council Decision no. 1106.</p>	Stability-arms race, crisis
<p>14. Participating states will, on a voluntary basis and consistent with national legislation, promote public-private partnerships and develop mechanisms to exchange best practices of responses to common security challenges stemming from the use of ICTs.</p>	Information-security
<p>15. Participating states, on a voluntary basis, will encourage, facilitate and/or participate in regional and subregional collaboration between legally-authorized authorities responsible for securing critical infrastructures to discuss opportunities and address challenges to national as well as trans-border ICT networks, upon which such critical infrastructure relies.</p>	Information-security, threat actor
<p>Collaboration may, inter alia, include:</p> <ul style="list-style-type: none"> – sharing information on ICT threats; – exchanging best practices; – developing, where appropriate, shared responses to common challenges including crisis management procedures in case of widespread or transnational disruption of ICT-enabled critical infrastructure; – adopting voluntary national arrangements to classify ICT incidents in terms of the scale and seriousness of the incident; – sharing national views of categories of ICT-enabled infrastructure states consider critical; 	

Table A.4. CBMs adopted through OSCE Permanent Council Decision no. 1202 on 10 March 2016 <i>(continued)</i>	
Recommended CBM	CBM classification
– improving the security of national and transnational ICT-enabled critical infrastructure including their integrity at the regional and subregional levels; and	
– raising awareness about the importance of protecting industrial control systems and about issues related to their ICT-related security, and the necessity of developing processes and mechanisms to respond to those issues.	
16. Participating states will, on a voluntary basis, encourage responsible reporting of vulnerabilities affecting the security of and in the use of ICTs and share associated information on available remedies to such vulnerabilities, including with relevant segments of the ICT business and industry, with the goal of increasing cooperation and transparency within the OSCE region. OSCE participating states agree that such information exchange, when occurring between states, should use appropriately authorized and protected communication channels, including the contact points designated in line with CBM 8 of Permanent Council Decision no. 1106, with a view to avoiding duplication.	Information-security
<p>Disclaimer</p> <p>The views and opinions expressed or implied in SSQ are those of the authors and are not officially sanctioned by any agency or department of the US government. We encourage you to send comments to: strategicstudiesquarterly@us.af.mil</p>	