Taming the Trolls: The Need for an International Legal Framework to Regulate State Use of Disinformation on Social Media

ASHLEY C. NICOLAS*

INTRODUCTION

Consider a hypothetical scenario in which hundreds of agents of the Russian GRU arrive in the United States months prior to a presidential election.[1] The Russian agents spend the weeks leading up to the election going door to door in vulnerable communities, spreading false stories intended to manipulate the population into electing a candidate with policies favorable to Russian positions. The agents set up television stations, use radio broadcasts, and usurp the resources of local newspapers to expand their reach and propagate falsehoods. The presence of GRU agents on U.S. soil is an incursion into territorial integrity—a clear invasion of sovereignty.[2] At every step, Russia would be required to expend tremendous resources, overcome traditional media barriers, and risk exposure, making this hypothetical grossly unrealistic.

Compare the hypothetical with the actual actions of the Russians during the 2016 U.S. presidential election. Sitting behind computers in St. Petersburg, without ever setting foot in the United States, Russian agents were able to manipulate the U.S. population in the most sacred of domestic affairs—an election. Russian "trolls" targeted vulnerable populations through social media, reaching millions of users at a minimal cost and without reliance on established media institutions.[3] Without using

---

* Georgetown Law, J.D. expected 2019; United States Military Academy, B.S. 2009; Loyola Marymount University M.Ed. 2016. The author is a former U.S. Army Intelligence Officer. The author would like to thank David and Griffey for their love and encouragement.

[1] The GRU is the Russian Military Intelligence Service and was the agency implicated in the 2016 hack of the Democratic National Committee. *See* Alina Polyakova & Spencer P. Boyer, *The Future of Political Warfare: Russia, the West, And the Coming Age of Global Digital Competition*, BROOKINGS 9 (2018), https://www.brookings.edu/wp-content/uploads/2018/03/the-future-of-political-warfare.pdf; *see also* Roland Oliphant, *Who Are Russia's Cyber-Warriors and What Should the West Do About Them?*, TELEGRAPH (May 6, 2017, 5:16 AM), https://www.telegraph.co.uk/news/2016/12/16/russias-cyber-warriors-should-west-do/ [https://perma.cc/5JJ8-YDL5].

[2] Some may argue that the efforts of these agents are also a violation of the principle of non-intervention, representing an attempt by a foreign state to "bear[] on matters" that a state is "permitted . . . to decide freely." Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S), Judgment, 1986 I.C.J. Rep. 14, 108, ¶ 205. This is a weak argument because the efforts of the Russians, while intrusive, are non-coercive and inherently limited in scope.

[3] A "troll" is "someone who joins a social media discussion on Facebook or Twitter, for example, and posts provocative comments, perhaps inflammatory or even off the topic, to

force, threatening the use of force, or invading sovereignty, the Russians were able to intervene in the domestic affairs of the United States. Under the current legal framework, this type of behavior by belligerent states escapes the reach of international law.

The "essential foundation" of the international order is state sovereignty manifested through respect for both "territorial sovereignty" and "political integrity."[4] The deceptively simple idea that states are entitled to make "the choice of a political, economic, social and cultural system" independently and free from the interference of other sovereigns has long been dependent on the sanctity of physical borders.[5] However, in the digital age, when states can project power from data centers located thousands of miles from an adversary, reliance on traditional ideas of Westphalian sovereignty, non-intervention, and use of force are insufficient to regulate state behavior.

The principle of non-intervention is a well-established norm in general and customary international law.[6] Under the pre-digital age analysis, to rise to the level of unlawful intervention, the offending state had to engage in some type of coercion to force a foreign sovereign into a choice it would not have otherwise made.[7] Although the use of force is a "particularly obvious" form of coercion, the International Court of Justice (I.C.J.) has had few opportunities to opine on the limits of prohibited behavior.[8] One reason for the dearth of I.C.J. litigation related to non-intervention is that, until the dawn of the digital age, most violations of the principle of non-intervention also required a physical invasion of sovereignty.

---

sow discord." Mike Snider, *Robert Mueller Investigation: What Is a Russian Troll Farm?*, USA TODAY (Feb. 16, 2018, 6:13 PM), https://www.usatoday.com/story/tech/news/2018/02/16/robert-mueller-investigation-what-russian-troll-farm/346159002 [https://perma.cc/Y38J-QNV9]; *see, e.g.*, *Dark Web: How Russian Trolls Spread Fake News*, WEEK (Nov. 6, 2017), http://www.theweek.co.uk/us/89497/dark-web-how-russian-trolls-spread-fake-news [https://web.archive.org/web/20180606033522/http://www.theweek.co.uk/us/89497/dark-web-how-russian-trolls-spread-fake-news], (estimating that "[a]t least 15 million Americans were exposed to content from 200 or so websites that were either operated by paid Russian trolls or by genuine conservative organisations [sic] using Russian propaganda as their source").

[4] Nicar. v. U.S., 1986 I.C.J. at 106, ¶ 202 (internal quotation marks omitted).

[5] *Id*. at 108, ¶ 205.

[6] *See, e.g.*, Conference on Security and Co-operation in Europe: Final Act, Aug. 1, 1975, 14 I.L.M. 1292, *reprinted in* 73 DEP'T ST. BULL. 323 (1975) (Providing that the participating states "will refrain from any intervention, direct or indirect, individual or collective, in the internal or external affairs falling within the domestic jurisdiction of another participating State, regardless of their mutual relations."); Treaty of Friendship, Co-operation and Mutual Assistance, May 14, 1955, 219 U.N.T.S. 3, Art. 8 (addressing "principles of respect for each other's independence and sovereignty and of non-intervention in each other's domestic affairs").

[7] *See* Nicar. v. U.S., 1986 I.C.J. at 108, ¶ 205 ("[i]ntervention is wrongful when it uses methods of coercion").

[8] *Id*.

The appropriate legal framework to analyze the online social media behavior of state actors is the same framework that applies to psychological operations. Throughout history, psychological operations have been inherently limited in scope and considered legal insofar as they did not constitute perfidy or violate the prohibition of intervention.[9] Not all interference equates to intervention. To qualify as intervention, "the interference must be forcible or dictatorial, or otherwise coercive, in effect depriving the state intervened against of control over the matter in question."[10] Without the Internet, it is difficult to imagine a scenario in which the use of information, on its own, could be considered coercive.

The Internet changes the coercion calculation. Social media as an information platform expands the reach of psychological operations so considerably that it rises to a level of prohibited intervention. The viral spread of falsehoods online deprives the victim state of control and is nearly impossible to defend against. The current international legal framework governing the use of psychological operations by state actors to shape foreign populations is insufficient to address the fundamental shift in the scale and scope of the use of disinformation in the digital age.[11] "[N]arrative manipulation through social media cyber operations"[12] represents the next great threat to the international community in cyberspace—a place of developing customary international law and few formal agreements.[13]

This Note is divided into five parts. Part I explores the history of psychological operations broadly, examining how information has been used to manipulate adversaries and foreign populations. Part II examines the development of technology and how social media has changed the way psychological operations are employed during peacetime to shape attitudes and intervene in sovereign affairs. Part III examines the current legal

---

[9] *See infra* Part III.

[10] 1 OPPENHEIM'S INTERNATIONAL LAW 432 (Sir Robert Jennings & Sir Arthur Watts eds., 9th ed. 1992).

[11] This Note uses the term "psychological operations" narrowly to refer to the use of disinformation to manipulate and coerce a civilian population. This term, along with "information operations," "influence operations," "cyberwarfare," "cyber operations," and "cyberattack" are often used interchangeably in the literature. Further, this Note will be limited in scope to the role of state actors. However, private industry has a powerful role to play in stopping the spread of deceptive information on social media. Facebook's recent decision to remove more than 100 accounts believed to be associated with "the Russia-based Internet Research Agency" highlights the critical role of industry and need for public-private collaboration. *See Facebook Removes More than 100 Accounts Linked to Russian Troll Factory*, GUARDIAN (Apr. 4, 2018), https://www.theguardian.com/technology/2018/apr/04/facebook-removes-more-than-100-accounts-linked-to-russian-troll-factory [https://perma.cc/J594-Y3LL].

[12] Jarred Prier, *Commanding the Trend: Social Media as Information Warfare*, 11 STRATEGIC STUD. Q. 50, 75 (Winter 2017), http://www.airuniversity.af.mil/Portals/10/SSQ/documents/Volume-11_Issue-4/Winter2017.pdf.

[13] This Note will not address the threat posed by information operation weapons to disrupt, jam, or misdirect signals equipment.

framework surrounding psychological operations and demonstrates how the gaps in that framework create legal grey zones for states to exploit through the use of disinformation on social media. Part IV discusses the role of international agreements in qualifying state use of "weaponized social media" as a prohibited intervention.[14] It then considers the design of a multilateral treaty that addresses the limits of acceptable deliberate state behavior on social media when the use is intended to manipulate foreign populations during peacetime. This Note concludes by addressing the threat of emerging technologies and the need to reach an international consensus regarding permissible online behavior.

I. THE HISTORY AND EVOLUTION OF PSYCHOLOGICAL OPERATIONS

A. PURPOSE AND STRUCTURE OF PSYCHOLOGICAL OPERATIONS

Throughout history, psychological operations (PSYOPS) haven take many forms at varying levels of complexity and effectiveness. The United States Department of Defense defines PSYOPS as "planned operations to convey selected *truthful* information and indicators to foreign audiences to influence their emotions [and] motives . . . . to induce or reinforce foreign attitudes and behavior favorable to the originator's objectives."[15] In Iraq and Afghanistan, the United States carried out PSYOPS through the use of Military Information Support Teams that employed a wide range of traditional tactics and tools including print media, broadcasts and leaflet campaigns.[16] In some cases, these efforts were intended to alert populations to danger to avoid unnecessary civilian suffering.[17] In other

---

[14] The weaponization of social media refers to "the adaptation (something existent or developed for other purposes) in such a way that it can be used as a weapon (platform / system) in order to achieve 'military' effect(s)." THOMAS ELKJER NISSEN, #THEWEAPONIZATIONOFSOCIALMEDIA: @CHARACTERISTICS_OF_CONTEMPORARY_CONFLICTS 96 (2015), https://www.stratcomcoe.org/thomas-nissen-weaponization-social-media [https://web.archive.org/web/20180412144439/https://www.stratcomcoe.org/thomas-nissen-weaponization-social-media].
"One of the potentially fertile areas for weaponizing social network media is the psychological warfare (PsyWar) area." *Id.* at 67.

[15] ROBERT J. KODOSKY, PSYCHOLOGICAL OPERATIONS AMERICAN STYLE: THE JOINT UNITED STATES PUBLIC AFFAIRS OFFICE, VIETNAM AND BEYOND xiv (2007) (emphasis added).

[16] *See* Meghann Myers, *The Army's Psychological Operations Community Is Getting its Name Back*, ARMYTIMES (Nov. 6 2017), https://www.armytimes.com/news/your-army/2017/11/06/the-armys-psychological-operations-community-is-getting-its-name-back/ [https://perma.cc/2XNX-5FNT] (noting that in 2017 the United States Army changed the name of the program from "Military Information Support Operations" back to Psychological Operations).

[17] *See* U.S. DEP'T OF ARMY, FIELD MANUAL 3-05.302, TACTICAL PSYCHOLOGICAL OPERATIONS TACTICS, TECHNIQUES, AND PROCEDURES 7-5 ¶ 7-19 (Oct. 2005); *see also* ARTURO MUNOZ, U.S. MILITARY INFORMATION OPERATIONS IN AFGHANISTAN: EFFECTIVENESS OF PSYCHOLOGICAL OPERATIONS 2001–2010 64–70 (2012), https://www.rand.org/content/dam/rand/pubs/monographs/2012/RAND_MG1060.pdf.

instances, PSYOPS were intended to hasten the end of hostilities or shape the population in a manner consistent with strategic objectives.[18]

PSYOPS can be divided into white, grey, and black operations. White operations are those which are overt and for which the state openly takes responsibility.[19] Grey information may be true or false but is presented without a source.[20] Black propaganda is presented with a false source.[21] That is, the information is attributed to a single source when it emanates from another.[22] As a matter of policy, the United States does not engage in disinformation through PSYOPS.[23] In other words, while the U.S. engages in white and grey operations, it does not admit to engaging in black operations during peacetime. This policy is an acknowledgment of the threat to legitimacy and credibility that exists when disinformation is attributed to a state actor. In the words of Colonel James Treadwell, former commander of the 4th Psychological Operations Group, "'the truth is the best propaganda' . . . . [o]therwise 'you lose credibility,' . . . and the audience tunes out."[24]

B. EXAMPLES OF PSYCHOLOGICAL OPERATIONS: DECEPTION TO DISINFORMATION

Efforts to deceive an enemy state have been underway since the earliest days of warfare. Even as the fledgling Army of the United States struggled to match the capabilities of a superior force, General George Washington used deception and covert operations to mislead British officials and gain a strategic advantage.[25] During World War II, a roughly 1,110-man unit, the 23rd Headquarters Special Troop, known as the "Ghost Army," manufactured alternate realities to confuse and mislead

---

[18] *See generally* MUNOZ, *supra* note 17, at 64–70 (providing examples of PSYOPS used in Afghanistan).

[19] Col. Frank L. Goldstein & Col. Daniel W. Jacobowitz, *Psychological Operations: An Introduction*, *in* PSYCHOLOGICAL OPERATIONS: PRINCIPLES AND CASE STUDIES 6 (Frank L. Goldstein & Benjamin F. Findley, Jr. eds., 1996).

[20] *Id*.

[21] *Id*. The use of disinformation on social media should be considered a black operation. This is because disinformation on social media typically uses false narratives, manufactured sources, and fake identities. *See* discussion of disinformation on social media, *infra* Part II.

[22] *See* Goldstein & Jacobowitz, *supra* note 19.

[23] Fred W. Walker, *Strategic Concepts for Military Operations*, *in* PSYCHOLOGICAL OPERATIONS: PRINCIPLES AND CASE STUDIES 17 (Frank L. Goldstein & Benjamin F. Findley, Jr. eds., 1996).

[24] JOINT CHIEFS OF STAFF, JOINT PUBLICATION 3-53: DOCTRINE FOR JOINT PSYCHOLOGICAL OPERATIONS I-3 (2003).

[25] Stephen F. Knott, *America Was Founded on Secrets and Lies*, FOREIGN POLICY (Feb. 15, 2016, 9:14 AM), https://foreignpolicy.com/2016/02/15/george-washington-spies-lies-executive-power/ [https://perma.cc/FSF5-TB5V].

German officials.[26] The Ghost Army used inflatable versions of U.S. Army heavy weapons, including replicas of Sherman tanks, combined with fake radio broadcasts known as "[s]poof [r]adio" to construct elaborate ruses that allowed the Allies to mask true troop movements.[27]

In the modern era, the United States has used information to affect the mindset and motivations of the enemy. During the Gulf War, the United States employed a massive leaflet campaign that effectively persuaded Iraqi troops to surrender.[28] The Coalition radio network, known as the "Voice of the Gulf," also encouraged Iraqi surrender by broadcasting messages intended to counter adversarial disinformation.[29] In Afghanistan, Army PSYOPS teams used multimedia products, including leaflets, radio broadcasts, social networking, and billboards, to push the message that al-Qaeda and the Taliban were "un-Islamic."[30] These campaigns pointed out religious inconsistencies and highlighted atrocities, undermining the legitimacy of al-Qaeda and the Taliban.[31]

The above examples from the United States all involve the use of legal ruses or the strategic use of valid information to inform or shape the population. However, adversaries of the United States have not restricted themselves to the use of verifiable information. During the Cold War, the use of disinformation was at the "forefront of the Soviet Union's strategy to discredit and undermine the United States."[32] When employing disinformation, the goal is not for the state version of the truth to be accepted as such; rather, the goal is to undermine the credibility of established institutions by sowing confusion and creating doubt.

Modern disinformation campaigns are in many ways rooted in the Cold War Soviet disinformation campaigns[33] that sought to convince the international community that the United States was a bad actor, unworthy of the mantle of global leadership. Operation Infektion was a Soviet operation, intended to undermine American credibility abroad, by convincing the international community that the United States had

---

[26] Megan Garber, *Ghost Army: The Inflatable Tanks That Fooled Hitler*, ATLANTIC (May 22, 2013), https://www.theatlantic.com/technology/archive/2013/05/ghost-army-the-inflatable-tanks-that-fooled-hitler/276137/ [https://perma.cc/S668-7QBS].

[27] *Id*.

[28] FINAL REPORT TO CONGRESS: CONDUCT OF THE PERSIAN GULF WAR 34 (1992), https://www.globalsecurity.org/military/library/report/1992/cpgw.pdf.

[29] *Id*. at 623.

[30] *See* MUNOZ, *supra* note 17, at 64–70.

[31] *See id.*

[32] Ashley Deeks, Sabrina McCubbin & Cody M. Poplin, *Addressing Russian Influence: What Can We Learn from U.S. Cold War Counter-Propaganda Efforts?*, LAWFARE (Oct. 25, 2017, 7:00 AM), https://www.lawfareblog.com/addressing-russian-influence-what-can-we-learn-us-cold-war-counter-propaganda-efforts [https://perma.cc/9933-9Q78]; *see also* Adam Taylor, *Before 'Fake News,' There Was Soviet 'Disinformation,'* WASH. POST (Nov. 26, 2016), https://www.washingtonpost.com/news/worldviews/wp/2016/11/26/before-fake-news-there-was-soviet-disinformation/?utm_term=.a9d882c93129 [https://perma.cc/7D6D-H7N4].

[33] The term "disinformation" is "an Anglicization of the Russian term '*dezinformatsiya*.'" Deeks et al., *supra* note 32.

invented AIDS in a laboratory.[34] The campaign was carried out through the use of fake news stories, cited to manufactured sources and planted in state-controlled print media.[35] The story gained widespread notoriety and the rumor—that AIDS is a U.S.-created pathogen—persists today. Other Soviet campaigns attempted to undermine alliances, manufacture distrust, and grow anti-American sentiment.[36] The Russian use of social media to spread false information is simply the next generation of *dezinformatsiya*.

Prior to the digital age, each category of PSYOPS was inherently limited in scope. A leaflet campaign, although effective, will only reach the individuals within the drop zone. Even television and radio broadcasts are limited by the strength of the available signal. Although PSYOPS can be valuable, modern warfare is highly decentralized, making it difficult to coordinate effective campaigns on a large scale.[37] These efforts are also resource intensive, requiring trained personnel, equipment, and operational support. The advent of the Internet and the capabilities presented by a simple, high-speed, low-cost platform, not bound by geographic barriers, changes the very foundation of PSYOPS as it has been understood for centuries.

II. THE DEVELOPMENT OF SOCIAL MEDIA: A CHANGE IN KIND, NOT DEGREE

Social media platforms not only increase the speed at which a fake story can spread but also remove the filters employed by reputable news agencies and allow adversaries to directly infiltrate communities. It is "the current embodiment of taking the fight directly to the people."[38] In some ways, this has been a positive development as social media has leveled the

---

[34] David Robert Grimes, *Russian Fake News Is Not New: Soviet Aids Propaganda Cost Countless Lives*, GUARDIAN (June 14, 2017, 7:54 AM), https://www.theguardian.com/science/blog/2017/jun/14/russian-fake-news-is-not-new-soviet-aids-propaganda-cost-countless-lives [https://perma.cc/GRQ7-UU3M].

[35] *Id*.

[36] *See, e.g.*, U.S. DEP'T OF STATE, FOREIGN AFFAIRS NOTE, SOVIET ACTIVE MEASURES: FOCUS ON FORGERIES, figs.5 & 6 (1983), http://insidethecoldwar.org/sites/default/files/documents/Department%20of%20State%20Note%20Soviet%20Active%20Measures%20Focus%20on%20Forgeries%20April%201983.pdf (discussing the use of forgeries to "add to frictions between the United States and its West European allies over the gas pipeline issue" and to make it appear that the U.S. was complicit in plans to overthrow multiple African governments); Dennis Kux, *Soviet Active Measures and Disinformation: Overview and Assessment,* 15 PARAMETERS: J. U.S. ARMY WAR C. 19, 26 (1985), https://www.iwp.edu/docLib/20131120_KuxSovietActiveMeasuresandDisinformation.pdf (discussing Soviet disinformation campaigns during the Cold War as an attempt to "exploit anti-American attitudes.").

[37] *See generally* Steven Collins, *Army PSYOP in Bosnia: Capabilities and Constraints*, PARAMETERS: U.S. ARMY WAR C. Q. 57 (1993) (discussing the difficulty of employing PSYOPS during the Bosnian conflict when tactical commanders needed to operate with high levels of autonomy).

[38] Prier, *supra* note 12, at 75.

playing field and given a voice to those without resources to navigate traditional media. For example, in Raqqa, Syria, a group of young men used social media to document the atrocities being carried out by ISIS and, in doing so, were able to execute a "counter campaign" and bring international attention to the mass suffering taking place in their country.[39]

Despite its potential as a weapon against tyranny, social media supplanting the traditional "media gatekeepers" has "contributed to the spread of misleading and outright fake news that is able to reach wide audiences to a degree unprecedented in modern history."[40] False stories posted on the Internet enter the public domain with the appearance of legitimacy but without any of the traditional checks on accuracy. Take for example a story posted to the self-proclaimed false news website, WTOE 5 News, that claimed Pope Francis had given his support to Donald Trump's presidential candidacy.[41] This story racked up over 960,000 Facebook engagements[42] within a month.[43] It contained patently false claims that would have been discredited in the earliest fact checks by any reputable organization. However, because the Internet allows a user to post without any barriers, WTOE 5 News was able to place this story on social media, allowing it to spread without any additional effort.[44]

The use of social media to propagate disinformation to deliberately manipulate a civilian population during a time of peace gained the attention of the Western media during the 2016 U.S. presidential election.[45] However, the use of disinformation to affect civilian populations during peacetime is not a phenomenon restricted to the United States. The annual Freedom on the Net Report found that "[o]nline

---

[39] *See* Dan Harris et al., *These Friends from Raqqa, Syria, Risk Their Lives to Document ISIS Horrors in Their Hometown*, ABC NEWS (July 6, 2017, 3:56 PM), http://abcnews.go.com/International/friends-raqqa-syria-risk-lives-document-isis-horrors/story?id=48478222 [https://perma.cc/S5RM-KW5M].

[40] DAVID PATRIKARAKOS, WAR IN 140 CHARACTERS: HOW SOCIAL MEDIA IS RESHAPING CONFLICT IN THE TWENTY-FIRST CENTURY 133 (2017).

[41] The website, WTOE 5 News, included a disclaimer stating, "WTOE 5 News is a fantasy news website. Most articles on wtoe5news.com are satire or pure fantasy." *See Pope Francis Shocks World, Endorses Donald Trump for President*, SNOPES, https://www.snopes.com/fact-check/pope-francis-donald-trump-endorsement/ [https://perma.cc/YGS8-L5DU] (last updated July 24, 2016).

[42] The number of "engagements" on Facebook refers to the number of times users interacted with a post by clicking on a link, liking or sharing the content. *See Post Engagement*, FACEBOOK, https://www.facebook.com/business/help/735720159834389 [https://perma.cc/MCU3-43K3].

[43] Hannah Ritchie, *Read All About It: The Biggest Fake News Stories of 2016,* CNBC (Dec. 30, 2016, 2:04 AM), https://www.cnbc.com/2016/12/30/read-all-about-it-the-biggest-fake-news-stories-of-2016.html [https://perma.cc/TD6G-WPA3].

[44] There is no evidence to support a claim that WTOE 5 News was run by a state actor. Rather, it was a satirical website run by private individuals. However, the viral spread of the Pope Francis story is demonstrative of the incredible reach of false narratives on the Internet.

[45] "Social media" refers to Internet-based social networking platforms that include, among others, Facebook, Twitter, Instagram, and Reddit.

manipulation and disinformation" impacted elections in at least eighteen countries in 2016.[46] Since 2004, Russia has been accused of interfering with the affairs of twenty-seven countries using a range of cyber tools, including disinformation.[47] Specifically, the Baltic states, including Estonia, Lithuania and Latvia, were targeted and manipulated by Russian disinformation campaigns.[48]

"Troll factories"—primarily a tool of the Russians—dramatically change the volume at which stories can be manufactured and planted.[49] The most widely publicized of the Russian troll factories involved in the U.S. presidential election was the "Internet Research Agency" (IRA).[50] Writing in twelve hours shifts, individuals at the IRA responded to prompts and posted assigned stories through "thousands of fake social media accounts."[51] IRA workers were assigned to either Russian or English speaking audiences and were required to meet comment and share quotas to increase the chances that a story would go viral.[52] In addition to commenting and sharing, an adversary can take advantage of the concept of "trending" stories to gain more control over how viral a story becomes. If the story is associated with a hashtag and is then shared at a high volume, the story will be included on the trending list on Twitter and other similar sites.[53] These lists drive content and can be powerful for spreading a message "across social clusters."[54] This tool allows an adversary to "weaponise" [sic] a trending topic by taking advantage of the "very media that uncovered it."[55]

---

[46] FREEDOM HOUSE, FREEDOM ON THE NET 2017: MANIPULATING SOCIAL MEDIA TO UNDERMINE DEMOCRACY 3 (2017), https://freedomhouse.org/sites/default/files/FOTN_2017_Final.pdf.

[47] Oren Dorell, *Alleged Russian Political Meddling Documented in 27 Countries Since 2004*, USA TODAY (Sept. 7 2017, 9:06 AM), https://www.usatoday.com/story/news/world/2017/09/07/alleged-russian-political-meddling-documented-27-countries-since-2004/619056001/ [https://perma.cc/CJG9-JYPE].

[48] Alexandra Wiktorek Sarlo, *Fighting Disinformation in the Baltic States*, FOREIGN POLICY RESEARCH INST. (2017), https://www.fpri.org/article/2017/07/fighting-disinformation-baltic-states/ [https://perma.cc/PZ7Q-X9ZT]. For a discussion about Russia's approach to mass disinformation, see ADAM SEGAL, THE HACKED WORLD ORDER 184 (2016).

[49] The term "troll factory" or "troll farm" refers to the large, centralized units within the Russian government that focus on producing disinformation and participating on social media at the direction of government authorities. *See generally*, Neil MacFarquhar, *Inside the Russian Troll Factory: Zombies and a Breakneck Pace*, N.Y. TIMES (Feb. 18, 2018), https://www.nytimes.com/2018/02/18/world/europe/russia-troll-factory.html [https://nyti.ms/2C6RZoE] (explaining the mechanics of one of Russia's largest troll factories).

[50] *Id.*

[51] *Id*.

[52] *See id*.

[53] *See* Prier, *supra* note 12, at 52.

[54] *Id.* at 53.

[55] *Id.*

The primary objective of Russian disinformation efforts is to "muddy the waters and cast doubt upon objective truths."[56] Modern Russian efforts have been termed a "firehose of falsehood" because of the speed, volume, and "shameless willingness to disseminate partial truths or outright fictions."[57] During the 2014 Ukrainian presidential election, a Russian-speaking hacker operation called CyberBerkut compromised the website of Ukraine's Central Election Commission and changed the election results so that the winner would be displayed as ultra-right candidate Dmytro Yarosh. [58] Simultaneously, Russian state media, working with CyberBerkut, published the fake results, damaging confidence in the system and the legitimacy of election.[59]

The media attention and subsequent investigation associated with the election of President Trump make the Russian efforts to manipulate the American populace well documented. An indictment released by the U.S. Department of Justice in February 2018 alleges that Russian "specialists"[60] stole the identities of Americans to "more authentically fabricate political sock puppets and avoid detection."[61] The allegations assert that Russian agents "created hundreds of social media accounts and used them to develop certain fictitious U.S. personas[.]"[62] Russian agents directed efforts at populations deemed most likely to influence their communities. For example, several pages registered through Netfinity JSC of Bulgaria were intended to target Vietnam era veterans, with an assumption that those veterans would be trusted local leaders.[63]

While the scope and magnitude of effects of these social media campaigns on the outcome of the 2016 election remains a matter of debate, there is no doubt that the election is illustrative of a new type of psychological operation. In the context of the election:

---

[56] *Understanding Russian "Hybrid Warfare" and What Can Be Done About It*, *Testimony Before the H. Armed Serv's Comm.*, 115th Cong. 3 (2017) (statement of Christopher S. Chivvis, RAND Corporation), https://www.rand.org/content/dam/rand/pubs/testimonies/CT400/CT468/RAND_CT468.pdf.

[57] CHRISTOPHER PAUL & MIRIAM MATTHEWS, THE RUSSIAN "FIREHOSE OF FALSEHOOD" PROPAGANDA MODEL: WHY IT MIGHT WORK AND OPTIONS TO COUNTER IT 1 (2016).

[58] Andy Greenberg, *Everything We Know About Russia's Election-Hacking Playbook*, WIRED (June 9, 2017, 7:00 AM), https://www.wired.com/story/russia-election-hacking-playbook [https://perma.cc/Z7XT-XQ8H].

[59] *Id*.

[60] *See* Indictment at 14, 16, United States v. Internet Research Agency LLC, No. 1:18-cr-00032-DLF (D.D.C. Feb. 16, 2018).

[61] Andy Greenberg, *Russian Trolls Stole Real US Identities to Hide in Plain Sight*, WIRED (Feb. 16, 2018, 5:29 PM), https://www.wired.com/story/russian-trolls-identity-theft-mueller-indictment/ [https://perma.cc/2RLT-SANN].

[62] Indictment, *supra* note 60, at 14.

[63] Natasha Bertrand, *The Fake Facebook Pages Targeting Vietnam Veterans*, ATLANTIC (Apr. 12, 2018), https://www.theatlantic.com/technology/archive/2018/04/foreign-actors-are-still-targeting-veterans-on-facebook-twitter-and-instagram/557882/ [https://perma.cc/U7ZD-YRKS].

> Cyber tools were also used [by Russia] to create psychological effects in the American population. The likely collateral effects of these activities include compromising the fidelity of information, sowing discord and doubt in the American public about the validity of intelligence community reports, and prompting questions about the legitimacy of the democratic process itself.[64]

Every major news story creates an opportunity for Russian exploitation.[65] Professor Mark R. Jacobson explains, "[w]hether it is 'Brexit' or the American election, Russian propaganda still infects U.S. social media networks . . . . [a]nd we see the same sort of divisive propaganda that we saw during the Cold War."[66] Even when fake stories fail to gain traction, there is value in undermining the credibility of the media[67] and sowing internal discord by amplifying debate.[68]

Perhaps the most cogent example of this tactic is the Russian disinformation campaign that took place after the crash of Malaysian Airlines flight MH-17. MH-17 was shot down over eastern Ukraine on July 17, 2014 by pro-Russian separatists.[69] Russia faced harsh and immediate global condemnation for arming the separatists.[70] In an attempt to shift blame to the Ukrainians, the Kremlin employed a global network of trolls to plant false stories, comment on social media threads, and obfuscate the truth.[71] Gaining traction for the false stories was immaterial.[72] The true goals were to cause confusion, create doubt, and undermine trust in traditional media outlets—to do that, "social media's various platforms, and the ability they have endowed upon users to spread narratives, were the perfect tool."[73]

---

[64] CATHERINE A. THEOHARY & CORY WELT, CONG. RESEARCH SERV., IN10635, RUSSIA AND THE U.S. PRESIDENTIAL ELECTION (2017).

[65] *See, e.g.*, Erin Griffith, *Pro-Gun Russian Bots Flood Twitter After Parkland Shooting*, WIRED (Feb. 15, 2018, 2:00 PM), https://www.wired.com/story/pro-gun-russian-bots-flood-twitter-after-parkland-shooting/ [https://perma.cc/ARB8-NB9D].

[66] Linda Qiu, *Fingerprints of Russian Disinformation: From AIDS to Fake News*, N.Y. TIMES (Dec. 12, 2017), https://www.nytimes.com/2017/12/12/us/politics/russian-disinformation-aids-fake-news.html [https://nyti.ms/2l1nGbF].

[67] *See id*.

[68] During the 2017 National Football League protests, U.S. Senator James Lankford accused Russian trolls of "taking both sides of the argument . . . to try to raise the noise level of America and make a big issue seem like an even bigger issue as they are trying to push divisiveness in this country." Dustin Volz, *Senator Says Russian Internet Trolls Stoked NFL Debate*, REUTERS (Sept. 27, 2017, 6:40 PM), https://www.reuters.com/article/us-usa-congress-cyber-russia/senator-says-russian-Internet-trolls-stoked-nfl-debate-idUSKCN1C237J [https://perma.cc/79CE-HBB2].

[69] PATRIKARAKOS, *supra* note 40, at 163–64.

[70] *Id.* at 164.

[71] *Id.* ("Conspiracy theory, denial, and blame-shifting: the building blocks of Russian propaganda laid bare, shared and tweeted into infinity.").

[72] *Id.* at 165.

[73] *Id.*

Soviet efforts to manufacture rumors and plant fake news stories are most analogous to modern social media manipulation efforts. KGB defector Vasili Mitrokhin testified that, in 1975, the KGB planted 5,510 stories and controlled ten Indian newspapers.[74] The aforementioned Operation Infektion began as a single manufactured story in a small Indian newspaper.[75] Through a series of strategic steps as well as a resurgence of concern about the spread of AIDS, the rumors became remarkably widespread—appearing in the major newspapers of upwards of fifty countries.[76] In his examination of Operation Infektion, Thomas Boghardt explains:

> Once the AIDS conspiracy theory was lodged in the global subconscience [sic], it became a pandemic in its own right. Like any good story, it traveled mostly by word of mouth, especially within the most affected sub-groups. Having effectively harnessed the dynamics of rumors and conspiracy theories, Soviet bloc intelligence had created a monster that has outlived its creators.[77]

Even at its height, Soviet disinformation was reliant on traditional means of communications, requiring control of news agencies or an ability to evade the filters of reputable organizations. In contrast, social media offers low-cost, simple, and effective weapons that are not limited in the ways traditional tools are limited. Social media capabilities are beyond anything available prior to the digital age, the scope of which is difficult to understate.[78] This change is not simply a change in degree; it is a change in kind and requires a new understanding of applicable law.

III. CURRENT LEGAL FRAMEWORK GOVERNING PSYCHOLOGICAL OPERATIONS

The use of information as a weapon is a basic tenant of warfare and has historically been a legal tactic.[79] The use of ruses is a long-accepted

---

[74] Thomas Boghardt, *Soviet Bloc Intelligence and Its AIDS Disinformation Campaign*, 53 STUD. INTELLIGENCE 1, 6 (2009).
[75] *See* Taylor, *supra* note 32.
[76] *Id.*
[77] Boghardt, *supra* note 70, at 19.
[78] *See generally* Samanth Subramanian, *Inside the Macedonian Fake-News Complex*, WIRED (Feb. 15, 2017), https://www.wired.com/2017/02/veles-macedonia-fake-news/ [https://perma.cc/8W6Y-YX99] (describing how a post made by a teenager publishing fake news in Veles, Macedonia can be posted in Facebook groups and shared by more than one thousand users).
[79] *See, e.g.*, SUN TZU, *Laying Plans, in* THE ART OF WAR 35 (Lionel Giles trans., 1910) ("[a]ll warfare is based on deception").

technique rooted in norms of chivalry.[80] The chivalric code was based on a sense of loyalty born out of a duty to uphold personal oaths. For example, a knight who swore an oath of loyalty to a king could not leverage that oath to gain access and harm that king. However, a knight could mask his true intentions, absent an oath, to gain a strategic objective over an adversary. In other words, if no oath was violated, there was no illegal act.

The distinction between permissible deception and unlawful perfidy has remained unchanged since the days of knighthood.[81] The Prohibition of Perfidy is codified in Article 37 of Protocol 1 to the Geneva Conventions ("Article 37").[82] Article 37 makes illegal "[a]cts inviting the confidence of an adversary to lead him to believe that he is entitled to, or is obliged to accord, protection under the rules of international law applicable in armed conflict, with intent to betray that confidence."[83] This includes acts such as feigning an intent to surrender or "feigning . . . . non-combatant status."[84] However, tactics, specifically including "the use of camouflage, decoys, mock operations, and misinformation," are not prohibited because they do not invite the confidence of the enemy "with respect to protection under that law."[85] The U.S. Army interprets permissible tactics to include "transmitting false or misleading radio or telephone messages, [and] deception of the enemy by bogus orders purporting to have been issued by the enemy commander."[86]

Although the mandates of the Geneva Conventions were sufficient to clarify the limits of acceptable behavior for nearly forty years, new threats emerging at the dawn of the digital age required the international community to reimagine some aspects of general and customary international law. Conversations surrounding emerging norms in cyberspace became especially intense after a massive cyberattack against Estonia in 2007. In 2013, the International Group of Experts (IGE) convened to produce the *Tallinn Manual on the International Law Applicable to Cyber Warfare* ("Manual").[87] The Manual attempted to apply traditional law of war concepts to the use of cyber weapons and pronounced that the concepts of *jus ad bellum* and *jus in bello* are

---

[80] Thomas C. Wingfield, *International Law and Information Operations*, *in* CYBERPOWER AND NATIONAL SECURITY 538 (Franklin D. Kramer, Stuart H. Starr, & Larry K. Wentz eds., 2009).

[81] Treachery is another common term for perfidy. There is no significant difference between the terms in the literature.

[82] Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), art. 37, June 8, 1977, 1125 U.N.T.S. 3.

[83] *Id*.

[84] *Id*.

[85] *Id*.

[86] Wingfield, *supra* note 80, at 539.

[87] *See* TALLINN MANUAL ON THE INTERNATIONAL LAW APPLICABLE TO CYBER WARFARE 1–2 (Michael N. Schmitt ed., 2013) [hereinafter TALLINN MANUAL].

applicable in cyber space.[88] As the group of experts explains, "[t]his means that that [sic] cybe[r] events do not occur in a legal vacuum and states both have rights and bear obligations under international law."[89]

An update to the original Manual, *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* ("Manual 2.0") attempts to maintain several of the concepts discussed above, including the distinction between ruses and perfidy, in cyberspace.[90] Rule 122 states that "[i]n the conduct of hostilities involving cyber operations, it is prohibited to kill or injure an adversary by resort to perfidy."[91] The Group of Experts agreed that only those "perfidious acts intended to result in death or injury" are prohibited under international law.[92] As defined in Rule 122, "[a]cts that invite the confidence of an adversary to believe that he or she is entitled to, or is obliged to accord, protection under the law of armed conflict . . . constitute perfidy."[93] In the context of social media, a state could violate the prohibition of perfidy if it uses social media to invite this reliance. For example, perfidy would be using social media to convince an adversary to come to a location to meet with the International Committee of the Red Cross when, in reality, the meeting is a planned ambush.[94] The Manual 2.0 permits the use of ruses in cyber operations.[95] This explicitly includes psychological operations.[96]

The Manual 2.0 looks to the United Nations Charter Article 2(4) definition of "use of force" to shape cyber operations. Rule 68 of the Manual 2.0 invokes the language of Article 2(4), requiring that "[a]ll Members . . . refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any State, or in any other manner inconsistent with the Purposes of the United Nations."[97] Rule 69 explains that there is no clear definition of what constitutes a cyber use of force, but the evaluation is typically done through a "scale and effects" test.[98] To qualify as a use of force, the cyber operation would need to rise to the level of a comparable non-cyber operation.[99] For example, the use of a cyber operation to damage a power

---

[88] *Foreword* to CYBER WAR: LAW AND ETHICS FOR VIRTUAL CONFLICTS, at v (Jens David Ohlin, Kevin Govern & Claire Finkelstein eds., 2015).
[89] Michael Schmitt, *Tallinn Manual Research*, NATO COOPERATIVE CYBER DEFENCE CENTRE OF EXCELLENCE, https://ccdcoe.org/research.html [https://perma.cc/2NQA-EEV3].
[90] *See* TALLINN MANUAL 2.0 ON THE INTERNATIONAL LAW APPLICABLE TO CYBER OPERATIONS 491–96 (Michael N. Schmitt & Liis Vihul eds., 2017) [hereinafter TALLINN MANUAL 2.0].
[91] *Id*. at 491.
[92] *Id.* at 492.
[93] *Id.* at 491.
[94] *See id*.
[95] *Id.* at 495.
[96] *Id.* at 496.
[97] U.N. Charter art. 2, ¶ 4; TALLINN MANUAL 2.0, *supra* note 90, at 329.
[98] TALLINN MANUAL 2.0, *supra* note 90, at 330–31.
[99] *Id*. at 330.

grid or to interfere with flight control operations would qualify as use of force. Rule 69 explains that "non-destructive cyber psychological operations intended solely to undermine confidence in a government" would *not* qualify as a use of force.[100] Under the existing model, the use of information to deceive and manipulate a civilian population on social media during peacetime falls short of a use of force and would escape the confines of Article 2(4) and Rule 69.

In relying on existing international law principles and rules, Manual 2.0 explains that, in cyberspace, even an act falling short of a use of force may be illegal under international law if it is a violation of sovereignty or a breach of the non-intervention principle.[101] In *Nicaragua v. United States*, the I.C.J. defined a prohibited intervention as one that is "bearing on matters in which each State is permitted, by the principle of State sovereignty, to decide freely."[102] The prohibition of intervention refers to actions that infringe upon the Westphalian definition of sovereignty and the concept of *domaine réservé.*[103] *Domaine réservé* refers to matters not governed by international law[104] and includes "the choice of a political, economic, social and cultural system, and the formulation of foreign policy."[105] Intervention requires an element of coercion. Use of force or providing military assistance during a civil war are examples of "obvious" coercion.[106] However, actions short of a use of force may be considered a violation of the principle of non-intervention even if the U.N. Charter does not prohibit the act. The I.C.J. has not yet limited the extent to which a state may act in cyberspace to influence the population of another sovereign.

Manual 2.0 attempts to address this grey area—a violation of non-intervention that falls short of a use of force. Manual 2.0 draws a line between intervention and interference: "[I]nterference refers to acts by States that intrude into affairs reserved to the sovereign prerogative of another State, but lack the requisite coerciveness . . . to rise to the level of intervention."[107] Even in cyberspace, coercion remains the determining factor between lawful interference and illegal intervention. Although the I.C.J. has stopped short of defining coercion short of use of force, the United States Supreme Court has characterized coercion as "the point at which pressure turns into compulsion."[108] Manual 2.0 suggests that the use of denial of service attacks intended to force a government to make a

---

[100] *Id.* at 331.
[101] *Id.* at 330.
[102] Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S), Judgment, 1986 I.C.J. Rep. 14, 108, ¶ 205.
[103] *See supra* Introduction at 37–38; *see also* TALLINN MANUAL 2.0, *supra* note 90, at 314 (identifying the term "domaine réservé").
[104] TALLINN MANUAL 2.0, *supra* note 90, at 314.
[105] Nicar. v. U.S., 1986 I.C.J. at 108, ¶ 205.
[106] *Id*.
[107] TALLINN MANUAL 2.0, *supra* note 90, at 313 (internal quotation marks omitted).
[108] South Dakota v. Dole, 483 U.S. 203, 211 (1987) (internal quotation marks omitted).

decision it would not otherwise make would be coercive and therefore unlawful intervention.[109] It is possible that a social media disinformation campaign could be so widespread and persuasive that it meets the requisite coercive requirement.

Regardless of its potential utility, Manual 2.0 is not law. Even the authors of Manual 2.0 acknowledge that, while valuable, it does not represent the views of states and therefore, does not constitute international law.[110] State Department Legal Adviser Brian Egan explained that, "[t]he United States has unequivocally been in accord with the underlying premise of [the Manuals], which is that existing international law applies to State behavior in cyberspace," but the United States does "not necessarily agree with every aspect of the Manuals."[111] The Manuals represent the opinion of a group of experts and are intended to guide decision makers; however, to effectively regulate the behavior of state actors, international agreements between the states themselves must be reached.

IV. THE WAY AHEAD: A CALL FOR INTERNATIONAL ACTION

A. DEVELOPING CUSTOMARY INTERNATIONAL LAW IN CYBERSPACE

There have been many calls for a comprehensive cyber treaty that would attempt to regulate behavior in cyberspace in a manner analogous to the way the United Nations Charter on the Law of the Sea (UNCLOS) regulates maritime behavior.[112] UNCLOS was based on centuries of conduct on the high seas, much of which had risen to the level of customary international law or been formalized in treaties.[113] In many ways, UNCLOS simply codified years of state practice and provided an arbitral forum to resolve disputes. No such customary law or historic precedent yet exists in cyberspace.

In the absence of formal international agreements, states depend on customary international law to regulate behavior and maintain order.

---

[109] *See* TALLINN MANUAL 2.0, *supra* note 90, at 315.

[110] *Id.* at 2 ("It is essential to understand that *Tallinn Manual 2.0* is not an official document, but rather the product of two separate endeavours [sic] undertaken by groups of independent experts acting solely in their personal capacity. The Manual does not represent the views of the NATO CCD COE, its sponsoring nations, or NATO. Nor does it reflect the position of any other organisation [sic] or State . . . .").

[111] Brian J. Egan, Legal Adviser, U.S. Dep't of State, Remarks on International Law and Stability in Cyberspace (Nov. 10, 2016) (transcript available at https://www.law.berkeley.edu/wp-content/uploads/2016/12/egan-talk-transcript-111016.pdf).

[112] *See* United Nations Convention on the Law of the Sea, Dec. 10, 1982, 1833 U.N.T.S. 397.

[113] BARRY E. CARTER & ALLEN S. WEINER, INTERNATIONAL LAW 814 (Vicki Been et al. eds., 6th ed. 2011) ("It is important to have an appreciation of the historical development of the law of the sea to understand the provisions of the LOS Convention. Some of its provisions reflect the customary international law and treaties existing at the time . . . .").

Customary international law refers to the set of international norms observed by states out of a sense of legal obligation.[114] A principle or rule becomes customary international law when it is the general and consistent practice of states, stemming from a sense of obligation and evidenced in *opinio juris*.[115] Although the prohibition on the use of disinformation on social media has not risen to the level of customary international law, norms are beginning to emerge.[116]

International organizations and regional agreements have begun to address the role that developing telecommunications technologies play in the area of foreign relations. In 2015, the United Nations Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of Internet Security (ITC Group) attempted to develop a framework for cyber behavior, without a strict focus on the law of war. In their report (ITC Report), the ITC Group explains that international law is binding on states in their use of ITC systems and that "States must observe, among other principles of international law, State sovereignty, sovereign equality . . . and non-intervention in the internal affairs of other States."[117]

Similarly, in the winter of 2015, representatives from China, Kazakhstan, Kyrgyzstan, the Russian Federation, Tajikistan, and Uzbekistan began to address developments in the field of "information and telecommunications in the context of international security."[118] The representatives authored a non-binding code of conduct, in which they agreed "[n]ot to use information and communications technologies and information and communications networks to interfere in the internal affairs of other States or with the aim of undermining their political, economic and social stability."[119] Nations with greater influence have also weighed in. In April 2017, leaders of the Group of Seven (G7) countries released a declaration explaining that they "are increasingly concerned about cyber-enabled interference in democratic political processes," and

---

[114] CURTIS A. BRADLEY, INTERNATIONAL LAW IN THE U.S. LEGAL SYSTEM 139 (2d ed. 2015).

[115] *Id*.; *see* CARTER & WEINER, *supra* note 113, at 116-18.

[116] In his statement following Russian cyber operations "aimed at the U.S. election," President Obama accused Russia of violating "established international norms of behavior." Press Release, Office of the Press Secretary, Statement by the President on Actions in Response to Russian Malicious Cyber Activity and Harassment (Dec. 29, 2016), https://obamawhitehouse.archives.gov/the-press-office/2016/12/29/statement-president-actions-response-russian-malicious-cyber-activity [https://perma.cc/W4C5-38KS].

[117] U.N. Secretary-General, *Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, ¶ 28, U.N. Doc. A/70/174 (July 22, 2015) [hereinafter U.N. Secretary-General, *Group of Governmental Experts*].

[118] Letter Dated 9 January 2015 from the Permanent Representatives of China, Kazakhstan, Kyrgyzstan, the Russian Federation, Tajikistan and Uzbekistan to the United Nations Addressed to the Secretary-General, at 1, U.N. Doc. A/69/723 (Jan. 13, 2015).

[119] *Id*. at 5.

calling for an international dialogue to shape international law in cyberspace.[120]

The G7 is taking a prudent step in calling for an open and public dialogue on the applicability of international law to cyberspace, as this dialogue will directly contribute to the formation of norms and expectations amongst states. A series of non-binding agreements at the United Nations, as well as the efforts at the regional and individual state level, directly contribute to building a consensus in the *opinio juris* prohibiting state use of social media to propagate disinformation.[121] Such a consensus would be a step toward establishing the use of disinformation on social media as a violation of customary international law.

B. THE ROLE OF INTERNATIONAL AGREEMENTS

The current legal framework governing PSYOPS is insufficient to account for the fundamental change in the scope and power of weaponized social media. The underlying premise behind the legal framework, or lack thereof, regulating traditional PSYOPS was that the enemy could be expected to defend itself.[122] Social media changes this expectation; regardless of the size of a nation's cyber force, little can be done to contain a viral post or tweet. A new paradigm needs to be developed to address this change, not in degree, but in kind. Traditional understandings of the prohibition of intervention and requisite coercion, as codified in treaties, are not broad enough to encompass state use of social media.

An effective treaty targeting state use of social media to spread disinformation must accomplish several aims. For a treaty to be effective, it must make clear that deliberate attempts by a state actor to manipulate the population of another sovereign through the use of disinformation on social media is an unlawful intervention in the affairs of the sovereign of the kind prohibited by customary international law, codified in numerous international agreements, and expressed in the decision in *Nicaragua*.[123] A multilateral agreement should be premised on an understanding that social media is a powerful tool not constrained by borders and that the spread of false narratives on social media has a negative impact on the stability of the entire international community. A successful treaty will also need to

---

[120] Group of Seven, *G7 Declaration on Responsible States Behavior in Cyberspace*, at 1, 3 (Apr. 11, 2017), http://www.esteri.it/mae/resource/doc/2017/04/declaration_on_cyberspace.pdf.

[121] Note that the Tallinn Manual does not represent *opinio juris* because it does not reflect the views of any state; it is simply the view of the International Group of Experts, a non-state entity.

[122] *See* Wingfield, *supra* note 80, at 538–39.

[123] *See, e.g.*, Convention on Rights and Duties of States Adopted by the Seventh International Conference of American States, art. 8, Dec. 26, 1933, 49 Stat. 3097, 165 L.N.T.S. 19; Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S), Judgment, 1986 I.C.J. Rep. 14, 106, ¶ 202 ("The principle of non-intervention involves the right of every sovereign State to conduct its affairs without outside interference.").

clarify the language used to refer to state use of social media, distinguish between peacetime and wartime behavior, address the role of non-state actors and proxies, and outline permissible countermeasures. The United Nations is an ideal forum for such an agreement as member states are already bound by the prohibition on the use of force and principle of non-intervention. Alternatively, a narrow protocol could be attached to the Geneva Conventions. Until such an agreement can be reached, the United States has an important role to play in the development of norms.

Ideally, this prohibition would be contained within a larger comprehensive framework, regulating conduct in cyberspace in a manner similar to what the Manual 2.0 models. Such a comprehensive framework or regime would have tremendous value in "encourag[ing] certain uses of cyberspace and discourag[ing] others."[124] However, beginning with a narrowly tailored solution to a distinct problem could represent a realistic starting point for future negotiations. While understanding that the use of social media to spread fake new stories is only one tool in an adversarial cyber playbook, a limited international agreement would be a productive step.

In the absence of a framework, a range of descriptive language is used to label disinformation campaigns on the Internet. The most extreme commentators refer to this attempt to interfere in sovereign acts as an act of war, invoking the language of Article 2(4) and Article 51 of the United Nations Charter.[125] Some commentators refer to these acts as criminal, inviting the attention of the Department of Justice and a variety of international tribunals.[126] Other commentators refer to these efforts as

---

[124] RICHARD HAASS, A WORLD IN DISARRAY: AMERICAN FOREIGN POLICY AND THE CRISIS OF THE OLD ORDER 246 (2017).

[125] *See* Louis Nelson, *Cardin: Russia's Election Meddling Is 'an Act of War*,' POLITICO (Nov. 1, 2017, 11:03 AM), https://www.politico.com/story/2017/11/01/russia-meddling-us-elections-ndi-event-244414 [https://perma.cc/2KFN-R67L] ("Cyber is an attack against our country. When you use cyber in an affirmative way to compromise our democratic, free election system, that's an attack against America . . . [i]t's an act of war."); Morgan Chalfant, *Former DNC Chair: Russian Election Hacking an 'Act of War*,' HILL (Mar. 29, 2017, 1:36 PM), http://thehill.com/policy/cybersecurity/326350-former-dnc-chair-calls-russian-election-hacking-an-act-of-war [https://perma.cc/RBG8-YNMT] (quoting former DNC Chair Donna Brazile: "I've never agreed with Dick Cheney in my entire life, but when he said this was an act of war, I have to agree with the former [V]ice [P]resident. It was an act of war."); Morgan Chalfant, *DHS Chief Issues Stern Warning to Russia, Others on Election Meddling, Cyberattacks,* HILL (Apr. 17, 2018, 2:47 PM), http://thehill.com/policy/cybersecurity/383571-homeland-security-chief-warns-of-seen-and-unseen-consequences-to [https://perma.cc/4P63-NDX6] (DHS Secretary Kirstjen Nielsen characterizing Russian cyber operations as attempts "to attack our democracy" and warning that "the U.S. government will consider all options 'seen and unseen' for responding to malicious attacks in cyberspace").

[126] *See, e.g.*, Helen Klein Murillo & Susan Hennessey, *Is It a Crime?: Russian Election Meddling and Accomplice Liability Under the Computer Fraud and Abuse Act*, LAWFARE (July 13, 2017, 10:24 AM), https://www.lawfareblog.com/it-crime-russian-election-meddling-and-accomplice-liability-under-computer-fraud-and-abuse-act [https://perma.cc/P7RT-KSKZ] (discussing available criminal charges directed at individuals involved with the Trump campaign).

military operations,[127] psychological operations,[128] or covert acts—each label inviting a different range of appropriate responses.[129] The current absence of clarity creates a legal "grey zone" in which a willing adversary can take advantage of competing interpretations of "poorly demarcated" international law principles and rules.[130] Although disinformation campaigns on social media are roughly analogous to traditional psychological operations, the differences in scale and scope necessitates a more precise label. Multinational agreements could adopt the language that already exists in the United Nations documents discussed above, prohibiting the use of information technologies to interfere with the affairs of the sovereign. These agreements must also distinguish between peacetime and wartime to clarify when normally prohibited acts could be employed as countermeasures.

Another option under an international agreement would be to adopt the position of the most aggressive commentators and invoke the power of the language of Article 2(4) and label social media disinformation operations on social media as a use of force.[131] A use of force label is problematic because the use of force invites a proportional use of force in response. The response is not restricted in kind. That is, there is a legally plausible argument that an adversary could respond to a social media campaign with a kinetic strike. Adopting use of force as the appropriate language also creates a difficult question regarding the determination of a threshold. In other words, the international community would need to reach an agreement on how to decide when a social media disinformation campaign reaches a level that qualifies as a use of force and what entity would be authorized to make that determination.

The nature of cyber operations makes it feasible for states to evade attribution and act through proxies with greater ease than with traditional

---

[127] Robert Chesney & Danielle K. Citron, *Disinformation on Steroids: The Threat of Deep Fakes*, COUNCIL ON FOREIGN RELATIONS (Oct. 16, 2018), https://www.cfr.org/report/deep-fake-disinformation-steroids [https://perma.cc/P833-WMAR]; Robin Emmott, *Britain, Baltics, Seek Italian Support for EU Cyber Sanctions*, REUTERS (Oct. 15, 2018, 6:56 AM), https://www.reuters.com/article/us-eu-cyber-sanctions/britain-baltics-seek-italian-support-for-eu-cyber-sanctions-idUSKCN1MP17O [https://perma.cc/HZ6F-QG53].

[128] Massimo Calabresi, *Inside Russia's Social Media War on America*, TIME (May 18, 2017, 3:48 PM), http://time.com/4783932/inside-russia-social-media-war-america/ [https://perma.cc/YS9E-X8MG].

[129] *See, e.g.*, Morgan Chalfant, *Former CIA Director: Don't Call Russian Election Hacking 'Act of War,'* HILL (Apr. 11, 2017, 4:29 PM), http://thehill.com/policy/cybersecurity/328344-former-cia-director-dont-call-russian-election-hacking-act-of-war [https://perma.cc/2QJM-8UY5] (quoting former CIA Director General Michael Hayden referring to Russian operations as a "covert influence campaign" and cautioning against the use of the phrase "act of war").

[130] Michael N. Schmitt, *Grey Zones in the International Law of Cyberspace*, 42 YALE J. INTL. L. ONLINE 1, 1 (2017).

[131] *See supra* note 125.

tools.[132] Any international agreement must address the use of proxies to carry out prohibited uses of social media as agents of the state and require states to regulate non-state behavior within their territory. The ITC Report acknowledges this necessity, stating that "States must not use proxies to commit internationally wrongful acts using [information and communications technologies], and should seek to ensure that their territory is not used by non-State actors to commit such acts."[133] An effective treaty would also encourage states to regulate behavior within their territories to prevent non-state actors from propagating false information for profit.[134]

Decentralized terrorist activity provides a similar challenge in that non-state actors often carry out acts without regard to borders. States have recognized the importance of encouraging one another to regulate behavior of non-state actors within their own territory by extending state responsibility to include agents and instrumentalities of the state as well as non-state actors. The International Convention for the Suppression of the Financing of Terrorism provides a relevant example of a provision designed to address this challenge. Article 18 states:

> Parties shall cooperate in the prevention of the offences [sic] set forth in article 2 by taking all practicable measures, inter alia, by adapting their domestic legislation, if necessary, to prevent and counter preparations in their respective territories for the commission of those offences [sic] within or outside their territories.[135]

States have been held civilly liable for sponsorship of terrorism directly and indirectly, through agents and instrumentalities.[136] By construing a similar provision to address the spread of disinformation, states will be

---

[132] Logan Hamilton, Note, *Beyond Ballot-Stuffing: Current Gaps in International Law Regarding Foreign State Hacking to Influence a Foreign Election*, 35 WIS. INT'L L.J. 179, 199 (2017) ("[T]o an extent almost unheard of with physical attacks, cyber-actions can be masked, through such means as proxy servers, virtual private networks ('VPNs'), the TOR software and network, botnets, or other measures that disguise the true origin of a cyber action.").

[133] U.N. Secretary-General, *Group of Governmental Experts*, *supra* note 117, at ¶ 28.

[134] This type of language would target the type of activity coming out of troll farms in Veles, Macedonia. *See supra* note 78. Language *encouraging* state behavior should be distinguished from language *mandating* state behavior. In regulating legal entities within their territories, treaty language should simply *encourage* to avoid unnecessarily requiring that states limit free and open use of the Internet.

[135] International Convention for the Suppression of the Financing of Terrorism, art. 18, Apr. 2002, 2178 U.N.T.S. 197.

[136] *See, e.g.*, Linde v. Arab Bank, 384 F. Supp. 2d 571, 576, 581, 583 (E.D.N.Y. 2005).

incentivized to legislate domestically to better counter disinformation and avoid liability.[137]

Even where agreements are not reached, there is value in participating in ongoing public debate between states. As Brian Egan explains, "[s]tating such views [on the scope of international law in cyberspace] publicly will help give rise to more settled expectations of State behavior and thereby contribute to greater predictability and stability in cyberspace."[138] Adding some predictability by "preserv[ing] and even expand[ing] areas of cooperation amid inevitable areas of disagreement" will decrease the chances of a rapid escalation in response and increase the chances of avoiding major conflict.[139]

Any attempt to control the use of social media to manipulate a foreign population through the spread of disinformation must be balanced with the need to protect a human right to access information. The West has "seemingly developed an almost messianic belief in the democratizing power of the Internet."[140] The United Nations has also emphasized the need to have open access to information. Article 19 of the Universal Declaration of Human Rights states that the fundamental right of freedom of expression includes the freedom "to seek, receive and impart information and ideas through any media and regardless of frontiers."[141] During his testimony to Congress in 1998, Bill Gates, in extolling the virtues of the Internet, highlighted its difficulty to control as a key component in its power to give a voice to the voiceless.[142] Developing international agreements that define the limits of permissible behavior with narrow prohibitions will preserve the promises of a democratized Internet while preventing the adversarial use of social media during peacetime.

C. THE ROLE OF THE UNITED STATES

1. On the World Stage . . .

Prior to, and during, the formation of an international agreement, the United States has an important role to play in developing norms. As the

---

[137] A pragmatic obstacle to imposing civil liability for use of disinformation on social media would be assessing damages. Discussing the practicality of a civil liability regime is beyond the scope of this Note.

[138] Egan, *supra* note 111.

[139] HAASS, *supra* note 124, at 220; *see also* Egan, *supra* note 111 (discussing the importance of restoring predictability to avoid the risks of "misperception and escalation").

[140] PATRIKARAKOS, *supra* note 40, at 134.

[141] G.A. Res. 217 A (III), Universal Declaration of Human Rights, art. 19 (Dec. 10, 1948).

[142] *Testimony of Bill Gates*, CNNMONEY (Mar. 3, 1998, 8:58 AM), http://money.cnn.com/1998/03/03/technology/gatestest/ [https://perma.cc/8RZX-6AB9] ("In any case, it is preposterous to think that any one company could ever control access to the Internet. The openness of the Internet is inherent in its architecture.").

largest economy in the world,[143] one of the most powerful voices on the world stage, and the home of major social media companies,[144] the United States has the authority to move any international agreement toward success or failure. In reality, any international agreement, absent U.S. ratification, is unlikely to have a significant quantifiable effect. Bad actors are unlikely to abide by the dictates of an agreement absent the weight of the United States. However, a treaty ratified by the United States would communicate a strong message to the international community that the use of disinformation on social media, especially through the use of fake identities, is unacceptable. If this agreement is coupled with an international effort to refute disinformation and domestically legislate to criminalize or impose civil liability for certain behaviors, the effect of disinformation could be dramatically reduced.[145]

Even if a multinational treaty is not possible, the United States should consider entering into bilateral treaties with willing nations that demarcate acceptable behavior on social media. Such bilateral agreements could begin by allowing for the use of social media campaigns via bilateral treaty but to require distinction or registration. It would be possible to require state actors to identify themselves, even if only in code. Nations that have strong alliances with the United States may be willing to enter into more comprehensive treaties explicitly prohibiting the use of disinformation and setting limits on other permissible behaviors. Regardless of the scope or number of parties bound by any international agreement, there is value in discussing prohibited behavior in a way that will restore some predictability to the international order.[146]

---

[143] *See* Int'l Monetary Fund, *Gross Domestic Product: Current Prices*, WORLD ECON. OUTLOOK DATABASE (Apr. 2018), https://bit.ly/2PpCJX3 [https://perma.cc/MND4-BCDP] (The United States' GDP in 2018, according to IMF staff estimates, will be roughly $20,400 billion. The next closest country is China, with an estimated GDP of roughly $14,000 billion.).

[144] Twitter, Facebook, Instagram, and Reddit are all headquartered in Northern California. *See* Careers: San Francisco, TWITTER, https://careers.twitter.com/en/locations/san-francisco.html [https://perma.cc/GLM4-PZ7V] (last visited Sept. 28, 2018); Facebook HQ, FACEBOOK, https://www.facebook.com/Facebook-HQ-166793820034304/ [https://perma.cc/69Y9-UH4V] (last visited Sept. 28, 2018); Instagram HQ: About, FACEBOOK, https://www.facebook.com/pg/instagramhq/about/?ref=page_internal [https://perma.cc/3E8U-K9VK] (last visited Sept. 28, 2018); Reddit (@reddit), TWITTER, https://twitter.com/reddit [https://perma.cc/2ENU-KKWM] (last visited Sept. 28, 2018).

[145] In November 2015, the European Union established the EU vs. Disinformation campaign, focused on exposing Kremlin disinformation. The campaign publishes "Disinformation Review" as an effort to highlight false stories and counter with factual information. *See generally Disinformation Review,* EU VS DISINFO, https://euvsdisinfo.eu [https://perma.cc/CV65-9HQD] (the Disinformation Review is dubbed the "flagship product of the EU vs Disinformation campaign").

[146] *See* Egan, *supra* note 111 ("In the context of a specific cyber incident, this uncertainty [created by the lack of an 'applicable legal framework'] could give rise to misperceptions and miscalculations by States, potentially leading to escalation and, in the worst case, conflict.").

2. . . . And at Home

In the United States, objections to a multinational treaty are rooted in a concern about limiting options in the cyber battlespace.[147] Even in the realm of covert actions, the President is bound by the Constitution and statutes of the United States.[148] If the United States ratified an international agreement limiting legal acts in cyberspace, the President would be bound by the *implementing legislation* of a non-self-executing treaty. However, the executive branch of the U.S. Government (U.S.G.) has previously interpreted Article II of the Constitution to allow the President to disregard a non-self-executing treaty absent implementing legislation.[149] Similarly, the U.S.G. has asserted that it is not bound by customary international law.[150] The reluctance of the United States to commit to international law is a source of disquiet amongst the international community. Although U.S.G. officials express a willingness to comply with international law "to the extent possible," it is difficult to assess how U.S.G. actions will impact the effectiveness of any international agreement.[151] If the United States refused to ratify an agreement, it would seriously undermine the effectiveness of its terms.[152]

To defend against disinformation attacks on social media, the United States would be well advised not to wait for an international agreement but rather to develop domestic solutions. The United States can make changes to legislation that would help curtail the effectiveness of information operations. One step would be to amend the Foreign Agent Registration Act (FARA) to make it applicable extraterritorially. FARA was originally "intended to restrict the importation of foreign political propaganda, in particular Nazi-sponsored materials" by requiring foreign agents to register with the Attorney General and label the political propaganda prior

---

[147] *See* Ido Kilovaty & Itamar Mann, *Towards a Cyber-Security Treaty,* JUST SECURITY (Aug. 3, 2016), https://www.justsecurity.org/32268/cyber-security-treaty [https://perma.cc/6S5T-5HAV] (discussing the national security concern that the "ceding of authority" to an international body via a comprehensive cyber treaty would "reduce the US's flexibility in employing its presumed relative technological advantage freely.").

[148] *See* 50 U.S.C. § 3093(a)(5) (1947).

[149] *See, e.g.*, Authority of the Federal Bureau of Investigation to Override International Law in Extraterritorial Law Enforcement Activities, 13 Op. O.L.C. 163, 183 (1989) (asserting that the President has inherent constitutional authority to authorize an extraterritorial arrest even if that arrest violates international law) [hereinafter Authority to Override].

[150] *See, e.g.*, *id*. *But see* Paquete Habana, 175 U.S. 677, 700 (1900) (holding that United States must comply with customary international law "where there is no treaty and no controlling executive or legislative act or judicial decision").

[151] S. SELECT COMM. ON INTELLIGENCE, 113TH CONG., ADDITIONAL PREHEARING QUESTIONS FOR MS. CAROLINE D. KRASS UPON HER NOMINATION TO BE THE GENERAL COUNSEL OF THE CENTRAL INTELLIGENCE AGENCY 7 (Comm. Print 2013), http://fas.org/irp/congress/2013_hr/121713krass-preh.pdf ("As a general matter, and including with respect to the use of force, the United States respects international law and complies with it to the extent possible in the execution of covert action activities.").

[152] *See supra* section IV.C.1.

to publication.[153] Expanding FARA to apply to those spreading propaganda on social media from abroad would be consistent with the original purpose of the legislation.[154] FARA is notoriously ineffective and difficult to enforce.[155] It is unlikely that those charged with violations of FARA applied extraterritorially would ever actually face prosecution in the United States. However, even if changes to FARA did not have a quantifiable effect, congressional action would contribute to evolving norms and expectations.

Recent actions by Special Prosecutor Robert Mueller and his team demonstrate the effectiveness of an indictment, even in cases in which prosecutions are unlikely.[156] An indictment, filed in July 2018, targets several Russian intelligence officers who hacked a variety of e-mail accounts, including those of the Democratic Party, in order to obtain information that was later used to manipulate the voting populace.[157] Although the targets are unlikely to ever see the inside of a courtroom in the United States, by writing a "speaking indictment," Special Prosecutor Mueller was able to educate the public as to the tactics of the Russian cyber operatives through a public filing.[158] An educated voter population is more likely to recognize disinformation campaigns, thereby reducing adversarial capabilities.

---

[153] Timothy Zick, *Territoriality and the First Amendment: Free Speech at—and Beyond—Our Borders*, 85 NOTRE DAME L. REV. 1543, 1576 (2010); 22 U.S.C. §§ 612, 614.

[154] As to the labeling requirement, the Supreme Court explained that the congressional requirement did not violate the First Amendment, it "simply required the disseminators of such material to make additional disclosures that would better enable the public to evaluate the import of the propaganda." Meese v. Keene, 481 U.S. 465, 480 (1987). In the wake of the 2016 election, the Department of Justice took steps toward using FARA to help regulate the behavior of press entities with foreign ties. The U.S. subsidiary of RT (a Russian-backed English language news outlet) and a "related production company" registered as foreign agents in late 2017 after receiving an order from DOJ. DOJ has also required media outlets affiliated with China, Japan, and South Korea to register under the Act. *See* Josh Gerstein, *DOJ Told RT to Register as Foreign Agent Partly Because of Alleged 2016 Election Interference*, POLITICO (Dec. 21, 2017, 10:42 AM), https://www.politico.com/story/2017/12/21/russia-today-justice-department-foreign-agent-election-interference-312211 [https://perma.cc/R6WT-X8M8].

[155] *See* Charles Lawson, Note, *Shining the 'Spotlight of Pitiless Publicity' on Foreign Lobbyists? Evaluating the Impact of the Lobbying Disclosure Act of 1995 on the Foreign Agents Registration Act*, 29 VAND. J. TRANSNAT'L L. 1151, 1164–67 (1996) (discussing FARA enforcement problems).

[156] *See generally* Eric Lach, *Why You Should Read the Latest Mueller Indictment Yourself*, NEW YORKER (July 13, 2018, 6:24 PM), https://www.newyorker.com/current/guccifer-indictment-robert-mueller [https://perma.cc/J3GC-5RL7] (describing the contents of the indictment accusing "twelve Russian military-intelligence officers of interfering in the 2016 U.S. Presidential election").

[157] *See generally* Indictment, United States v. Netyksho, No. 1:18-cr-00215 (D.D.C. July 13, 2018) (outlining the charges against the defendants).

[158] *See* Sarah Grant et al., *Russian Influence Campaign: What's in the Latest Mueller Indictment,* LAWFARE (Feb. 16, 2018, 10:55 PM), https://www.lawfareblog.com/russian-influence-campaign-whats-latest-mueller-indictment [https://perma.cc/T8NA-M4G4].

Freedom of speech and freedom of the press are bedrock values in the United States.[159] Any legislation that appears to curtail those freedoms or makes the federal government the sole arbiter of truth will be met with stiff resistance. It is imperative that, in addressing this complex problem, the government works closely with private industry to protect constitutional guarantees, adequately address the threat, and ensure U.S. companies are not unintentionally exposed to liability. Rather than silencing speech, the federal government should work with U.S. based social media companies to require account identification verification or public disclosure of state sponsorship. A disclosure or verification requirement would undercut a state's ability to operate false accounts in order to manipulate the populace in the way the Internet Research Agency did during the 2016 election.[160] Under any scheme in which the government imposes civil liability against media companies for hosting disinformation, care must be taken to ensure social media companies are shielded in the event a foreign state covertly uses social media platforms to spread disinformation.

CONCLUSION

Advancing the law to regulate state behavior online is tremendously complex and of great importance. Developing technology exponentially increases the threat posed by social-media-based disinformation and emphasizes the need for a workable legal framework. Recently, technology company Adobe developed a tool called VoCo that allows users to manufacture voice recordings that are indistinguishable from live speech.[161] New technology also allows for the flawless creation of videos with superimposed faces, which enables users to create a false reality in video recordings.[162] Aside from the threat of fake reporting on matters of international and domestic concern, there is also a real threat that these technologies could be used during a crisis to interfere with first responders or impede evacuations. Weaponized social media has the potential to dramatically increase the lethality of attacks and level of chaos in the world.

---

[159] U.S. CONST. amend. I.

[160] Grant, *supra* note 158 (explaining "that the Internet Research Agency used [stolen] bank and PayPal accounts to fund online advertising and organize political events in the U.S. linked to its fake personas and pages—including, in some cases, paying people to appear at rallies.").

[161] Matthew Gault, *After 20 Minutes of Listening, New Adobe Tool Can Make You Say Anything*, VICE: MOTHERBOARD (Nov. 5, 2016, 3:00 PM), https://motherboard.vice.com/en_us/article/jpgkxp/after-20-minutes-of-listening-new-adobe-tool-can-make-you-say-anything [https://perma.cc/739H-JAHF].

[162] *See* Emma Bowman & Lawrence Wu, *In an Era of Fake News, Advancing Face-Swap Apps Blur More Lines*, NPR (Feb. 3, 2018, 8:37 AM), https://www.npr.org/2018/02/03/582767531/in-an-era-of-fake-news-advancing-face-swap-apps-blur-more-lines [https://perma.cc/KP9V-APVV].

The time has come for the global community to coalesce around norms and rules to regulate the behavior of states in cyberspace. A prudent first step is to address the spread of disinformation by state actors on social media attempting to influence foreign populations during peacetime. The speed and reach of such social media efforts make modern disinformation different in kind than the disinformation of the Cold War and represents a prohibited intervention in the sovereign affairs of a foreign state. Ideally, a prohibition on these efforts would be contained with a comprehensive framework for state cyber operations, but there is value in a narrowly tailored multinational agreement as a first step towards establishing rule of law on social media.