Harder, Better, Faster, Stronger: International Law and the Future of Online PsyOps

Tim Hwang[1] and Lea Rosen[2]

[1]Data and Society
[2]Pacific Social

January 17, 2017

Abstract

Recent years have seen an explosion of activity from states and non-state actors seeking to manipulate online political discourse at home and abroad. These efforts have leveraged a range of different techniques, from the use of swarms of automated bots to the systemic spreading of misleading or outright fabricated information through social media. Most dramatically, recent revelations at the time of writing have suggested that the use of these techniques by the Russian government may have played a role in swaying the outcome of the 2016 US presidential election. Technological trends seem poised to make these types of online psychological operations (psyops) ever cheaper, more effective, and difficult to attribute in the near future. Given the potential for this new generation of psyops to destabilize the global political environment, what can be done through channels of international law and other forms of coordination to combat or control the impact of these persuasive campaigns?

This paper examines this question in the context of state and non-state actor use of online psyops to undermine other states. It examines the current state of development of these techniques, and projects future capabilities based on recent advances in artificial intelligence and quantitative social science. It then examines a set of applicable international legal frameworks, arguing that the existing body of laws and norms fail to adequately constrain the use of these techniques. Finally, it provides a set of potential interventions for exploration, considering both technical and legal approaches.

Support for this paper was provided by the European Research Council under grant "COMPROP—Computational Propaganda: Investigating the Impact of Algorithms and Bots on Political Discourse in Europe." Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the European Research Council.

Introduction

Recent years have seen an explosion of activity from states and non-state actors seeking to manipulate online political discourse at home and abroad. These efforts have leveraged a range of different techniques, from the use of swarms of automated bots to the systemic spreading of misleading or outright fabricated information through social media. They have been used for a range of different purposes, from hindering the online coordination of dissenters to supporting the effectiveness of military activity on the ground.

On one level, these efforts are not novel. Indeed, psychological operations (psyops) by states and non-state actors have a long-standing history in conflict that precedes the rise of the Internet.[1] Similarly, the use of propaganda to shape domestic public opinion and undermine dissent is an established phenomenon, which echoes contemporary cases seen on the Internet.[2]

However, the unique characteristics of the online environment and advances in technology seem poised to give rise to a new generation of techniques which are greatly expanded in scope and effectiveness. Whereas earlier attempts to manipulate public opinion were expensive, slow, data poor, and attributable, contemporary techniques are cheap, fast, data rich, and difficult to attribute. These factors are likely to make these techniques considerably more destabilizing to the international environment, and increasingly so in the near future.

This paper takes up a specific challenge: addressing the potential risks posed by the use of modern and near future psyops by states and non-state actors to attack the stability of other states. Other issues pre-

1 E.g. Holt, *The Deceivers*; Linebarger, *Psychological Warfare*.
2 E.g. Welch, *The Third Reich: Politics and Propaganda*.

sented by the use of these techniques—for example, by governments on their own citizens, or in the use of these techniques by private actors against other private actors—is beyond the scope of this paper.

To date, the use of psyops for this purpose has not been considered so significant or dangerous to be the subject of international laws and agreements that would limit or combat their use by nations or others. Indeed, as discussed below, proposals to do so have been rejected by the United States and other members of the international community in the past. But, to the extent that psyops continues to improve and become a more effective means of attack, it may soon require renewed attention from the international community to limit its impact.

We will take stock of the existing potential legal mechanisms available to tackle the use of the new generation of online psyops against states, and propose alternatives to the extent that established tools seem unable to address the threat. First, we will review the current landscape, examining the emerging techniques that different actors are using to manipulate public opinion through the Internet. We will then discuss some emerging trends in research that suggest that these techniques will become increasingly more powerful over time, distinguishing it from earlier generations of "traditional" psyops. Second, we will examine some of the existing legal frameworks under which the use of psyops for destabilizing states might be limited or prevented. Third, arguing that these frameworks are insufficient, we propose a set of potential alternative interventions that may help to mitigate the potential negative effect of these technologies. Finally, we conclude with some areas for potential further exploration.

Part I: The Present and Future of Manipulating Discourse Online

In recent years, journalists have uncovered a panoply of techniques that have been deployed to influence discourse online, particularly in the political realm. These have ranged from simple campaigns of spam used to suppress dissenting voices online to sophisticated campaigns of disinformation tightly integrated with conventional military operations and cyberattacks.

On social media, bots—fake user accounts that often autonomously repeat the same or meaningless content—have proven to be a particularly popular method for manipulation. Activists in Turkey and Syria have been subject to bot spamming campaigns, which attempt to drown out oppositional political speech occurring on popular Twitter hashtags.[3] In the US, false accounts have been used to bolster the apparent grassroots support of political candidates, sometimes with a particular emphasis among key constituencies.[4] In Mexico, one recent presidential election cycle featured two opposing groups of bots attempting to contest each other on social media.[5] Bots were also a prominent feature of the online political discussion around the "Brexit" vote in the United Kingdom, helping to rally support around the decision to leave the European Union.[6]

More sophisticated efforts have focused on augmenting the ability for human operators to manage multiple plausible "personas" on social media. Examples include a 2011 contract awarded by the US Department of Defense to create software that would allow an individual to control up to 10 distinct identities appearing to be located in different parts of the world. This was thought to be connected to "Operation Earnest Voice," a campaign to counter violent extremist and enemy propaganda through the Internet.[7] China has also engaged in similar activity, leveraging large numbers of participants to form its "50 Cent Army," a coordinated effort to redirect and derail political discourse online.[8]

Fake identities are only part of the picture. Beyond bot accounts, efforts have also been uncovered which integrate these methods as just one component of more sophisticated strategies for shaping public opinion. One example detailed in the New York Times in 2015 is the Russian "Internet Research Agency," which has been connected to elaborate misinformation schemes that include fabricated videos and realistic clones of actual news sites.[9] These activities saw a particularly dramatic culmination in efforts by the Russian government to sway the 2016 US presidential election, a campaign which combined both psyops and hacking in support of its objectives.[10]

3 Sozeri, "The Rotten Politics Infecting Turkey's Social Media"; Qtiesh, "Spam Bots Flooding Twitter to Drown Info About #Syria Protests [Updated]."
4 Gaffney, "Statistical Probability That Mitt Romney's New Twitter Followers Are Just Normal Users"; Andrews, "Pro-Trump Twitter Bots at Center of Nevada Mystery."
5 Finley, "Pro-Government Twitter Bots Try to Hush Mexican Activists."
6 Dewey, "How Online Bots Conned Brexit Voters."
7 Fielding and Cobain, "Revealed."
8 King, Pan, and Roberts, "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument."
9 Chen, "The Agency."
10 Office of the Director of National Intelligence, "Assessing Russian Activities and Intentions in Recent US Elections"; Kelly, "FBI Agrees With CIA On Russian Interference In Presidential Election."

These types of actions are not limited to comparatively well-resourced government agencies. Bloomberg reported in April 2016 the story of Andrés Sepúlveda, who was involved in a series of private efforts to sway elections in Latin America through a combination of bots, compromising of voting machines, and digital eavesdropping, among other techniques.[11]

The last few years have also seen revelations of more ambitious undertakings that not only attempt to shape activity occurring on existing channels online, but seek to develop entirely new channels, as well. Zunzuneo—a text-based social network launched in Cuba—was revealed in 2014 to be a US-backed scheme to influence public opinion and trigger "smart mobs" in an effort to undermine the Castro regime.[12]

These examples—only a selection of a much larger number of stories that have emerged in recent years—speak to the interest among state and non-state actors in shaping public opinion through online channels.

Emerging Trends

These cases by themselves would not be surprising, as we might expect online channels to be the subject of psyops, in the same way that established media like newspapers and radio have been in the past.[13]

However, two key trends seem poised to converge with these new psyops campaigns in ways that would make them substantially more effective and destabilizing to the global geopolitical environment in the near future. This new generation of "computational propaganda" presents new risks, and invites an analysis of legal and technical interventions that should be arrayed to mitigate the threat.

Trend: Better Fakes

Many of the campaigns discussed above leverage astroturfing—creation of masses of false identities in order to give the impression that an upwelling of opinion exists where it does not in reality. These types of campaigns critically rely on the believability of the identities being launched. If "users" are easily identified as being fake or originating from a single obvious source, the persuasive impact of the campaign is significantly diminished and can be easily flagged by users for removal by the platform.

To date, the realism of campaigns of bot persuasion has been limited. Fake accounts often repeat the same content over and over, and profile pictures and other content are often copied wholesale from elsewhere.[14] When content is generated programmatically, the posts of these fake accounts will often appear with similar syntactic structure or may appear as meaningless collections of words.[15]

Indeed, this regularity is often the means by which reporters have been able to detect and report on these campaigns to date. It has also limited the scale of some types of persuasive campaigns: a prospective planner of these types of psyops must craft unique, believable identities manually, raising the costs and time necessary to set up and execute these tactics.

However, the cost of simulating more believable identities appears to be poised to drop significantly as new techniques emerge. Methods to programmatically generate realistic synthetic faces at scale continue to improve.[16] Similarly, breakthroughs in machine learning are enabling the creation of ever-more authentic-sounding computer speech.[17] The end result is to enable the creation of identities that can look and sound real, without any obvious copying from other sources.

Beyond the capacity for these fake identities to appear real, technology will also lower the costs of enabling them to interact effectively with real users, as well. Advances in the field of machine learning and deep learning are enabling the creation of conversational agents that are substantially more sophisticated than what has been available in the past.[18] This research might be quickly adapted from openly available papers on the topic to power a swarm of fake identities that have the power to interact believably with real users and gain their trust.

Both of these trends will both lower the costs to creating believable identities and raise the flexibility to which these campaigns can be put. Whereas "bots" might have in the past been largely used to push a sequence of repetitive messages at scale or pad the apparent popularity of a user, these tech-

11 Robertson, Riley, and Willis, "How to Hack an Election."
12 Associated Press, "US Secretly Created 'Cuban Twitter' to Stir Unrest and Undermine Government."
13 E.g. Mulford, "Benjamin Franklin's Savage Eloquence"; Puddington, "Broadcasting Freedom."
14 *Supra*, notes 2-3.
15 *Ibid.*
16 *See* Ohana et al., "HoneyFaces." Observing that the generation of photo-realistic synthetic faces was "very efficient and takes only 1.2903 × 10−4 seconds [per face] on average using Matlab.
17 E.g. DeepMind, "WaveNet."
18 *See*, *e.g.*, Vinyals and Le, "A Neural Conversational Model."they are often restricted to specific domains (e.g., booking an airline ticket

nologies might enable more subtle campaigns of persuasion and rapport building with real users at scale. It will also make these campaigns less detectable, as bot behavior increasingly converges with the normal behavior of human users online.[19]

Trend: Data, Targeting, and Social Physics

The widespread adoption of social media platforms has been one of the defining developments of the modern web. Platforms like Facebook now count over 1.7 billion users worldwide, with Twitter and WhatsApp featuring 313 million and 1 billion users, respectively.[20]

Parallel to this adoption has been the generation of a massive quantity of data about social behavior. Users reveal their preferences and social connections through their activities on these platforms, which in aggregate provide a high resolution, continually updated picture of vast segments of the global population. This abundance of data has produced two notable effects relevant to the future of psyops campaigns.

For one, it has significantly enhanced the capacity to effectively target messages to particular constituencies and even individuals of interest. This has perhaps been most evident in the advertising services offered by social media platforms. Facebook, for instance, offers to advertisers the ability to target ads across a range of highly granular characteristics, including by interest, geography, and connection behavior.[21] These benefits exist even without data or permission from the platform itself—a range of activists, trolls, and other political actors have used publicly available posts and user profiles to target their messaging for recruitment and harassment in recent years.[22] This appears to already be informing psyops campaign activity, with the launch of groups of bots tailored to message and appeal to particular constituencies of interest.[23] Targeting may become ever more granular and effective going forwards.

Secondly, researchers have been leveraging the availability of social data to better understand the mechanics powering group behavioral phenomena like the "viral" spread of content through a network, or the factors encouraging the spread of misinformation.[24] This has given rise to an emerging body of research, dubbed by MIT researcher Alex Pentland as a new "social physics"—a sufficiently advanced, quantitative understanding of social processes that allows for the prediction and manipulation of those processes.[25] Leveraging these techniques, Pentland's lab was able to incentivize and trigger a nationwide search coordinated through a crowd of online collaborators to win the DARPA Grand Challenge in 2009.[26] More recent work relevant to the psyops space has focused on how changes to display of information about peers can influence voting behavior of an individual.[27]

While in the realm of academic inquiry, these results seem to have clear application in the context of online psyops. Many of these results are published openly in research journals, and could be easily leveraged to increase the effectiveness of persuasive campaigns online. This might include the use of predictive statistical models to inform when messaging efforts might be most effective in spreading a message, or leveraging modern community detection algorithms to identify those susceptible to being rallied for a particular cause.[28]

The Next Generation of Psyops

Taken together, these contemporary case studies and the trends that are on the horizon suggest the emergence of a new generation of psyops campaigns that will be broader in scope and more effective than the types of campaigns that took place in the past. To illustrate, we compare the types of campaigns we have been discussing with leafleting—a prototypical and well documented case of "traditional" psyops.[29] Four key differences stand out.

First, traditional methods of psyops campaigning are comparatively more expensive propositions. Whereas leafleting requires printing and distribution by aircraft over a target area, bot campaigns are coordinated online and distribute messages to users anywhere in the world at little or no cost. This makes

19 Subrahmanian et al., "The DARPA Twitter Bot Challenge." Concluding after a large-scale competition to detect bots that the trend was towards "sophistication" in these systems in the near future.
20 Facebook, "Company Info | Facebook Newsroom"; Statt, "WhatsApp Has Grown to 1 Billion Users"; Twitter, "Company | About."
21 Facebook, "Choose Your Audience."
22 *See, e.g.* Wikipedia, "Gamergate Controversy."
23 *Supra,* note 5.
24 *See*, e.g. Cheng et al., "Can Cascades Be Predicted?"; Kumar, Robert, and Leskovec, Jure, "Disinformation on the Web: Impact, Characteristics, and Detection of Wikipedia Hoaxes."
25 Pentland, *Social Physics*.
26 Tang et al., "Reflecting on the DARPA Red Balloon Challenge."
27 Bond et al., "A 61-Million-Person Experiment in Social Influence and Political Mobilization."
28 *See, e.g.,* Wagner et al., *When Social Bots Attack*.
29 Friedman, "Falling Leaves"; Peffer, "Paper Bullets: An Interview with Herbert A. Friedman."

it an affordable strategy for state actors, and allows these types of campaigns to be conducted by small groups of individuals that previously would not have the resources to conduct effective persuasion at scale.

Second, traditional psyops is comparatively slower. Leafleting requires the drafting of content, the printing of the physical paper for distribution, and then their eventual dissemination. The types of bot campaigns that have been seen in recent years, in contrast, can distribute content instantaneously to their audiences and are distributed through sham accounts, which can be quickly generated or purchased.[30] This speed also implies a faster rate of iteration, as well: changing the messaging in a leaflet campaign that is proving to be ineffective might require reprinting the physical documents. In contrast, bot campaigns can begin changing content instantaneously as soon as the operator chooses, enabling a nimbler cadence of operation.

Third, traditional psyops is data poor, while new generation psyops can leverage the abundance of data about social behavior now available online. This has implications for how effective these campaigns might be. The distribution of leaflets provides limited feedback to those launching these campaigns on who received the message, and who was influenced by it.[31] In contrast, bot campaigns taking place on social media platforms like Twitter enable their operators to actively choose who to deliver messages to, and closely monitor whether or not the behavior of those being targeted is changing. Moreover, operators of bots can also leverage the "social physics" techniques discussed above to predict and plan their operations.

Fourth, traditional psyops campaigns are more attributable due to their reliance on more visible or obvious forms of distribution. Whereas leaflets are distributed in an obvious manner (by aircraft, for example), which may reveal the perpetrators of the operation, there is no similar need in the case of bots. Such campaigns can operate through a large group of false identities that bear no signals of ownership. Moreover, messaging can be subtle—bots can mix in their persuasive messaging with more benign forms of content to bolster their believability and obfuscate the progress of a campaign. The end result is that this next generation of psyops campaigns may be less attributable. It can be more difficult to tell who is pushing forwards a strategic persuasion effort, and indeed even when one of these efforts has begun or ended.

These points of comparison speak to a change in degree that might more rightly be considered a change in kind. When combined with advances in the ability to generate more believable identities and an increased understanding of the factors driving group behavior, this next generation of psyops seems poised to enable a variety of actors to effectively shape and persuade at significant scale. The resulting destabilizing effect, particularly when combined with the continuing evolution of cyberwarfare more broadly, may justify more extensive international intervention than has happened in the past around these tactics.

Future Scenarios

By way of making the above analysis more tangible, we present a series of potential future scenarios that leverage the techniques seen in use currently and the emerging trends discussed above to illustrate threats that may be possible as psyops continues to evolve into the near future. These are presented as a series of fictional excerpts from news stories, with analysis provided to indicate how the various techniques we have been discussing might be integrated together in practice.

Realistic Identities - "Shooters Radicalized By Bot, Still at Large"

Chicago remains on lockdown for a seventh day as the manhunt for alleged killer James Colford continues. Law enforcement released the results of a forensic analysis of Colford's computer, confirming earlier reports that he had been radicalized and urged to kill through a series of ongoing conversations with a set of users on social media. Investigators have indicated that this pattern of engagement has been mirrored elsewhere, suggesting that these accounts may have been false identities executing a program. It remains unknown who is releasing these bots, or how many of them remain…

Radicalizing individuals for terroristic acts might be enhanced through the creation of a specialized swarm of "recruitment bots," which leverage available research to identify patterns of user behavior online that indicate that a user might be receptive to messaging urging them to take lethal action. Many bots might be deployed at scale with different personalities, repeatedly identifying and sounding out potential candidates on social media from a variety of different backgrounds and ensuring that the candidates do not have any pre-existing connections with one another to avoid detection.

30 Bilton, "Social Media Bots Offer Phony Friends and Real Profit."

31 *See*, e.g. Oyen and De Fleur, "The Spatial Diffusion of an Airborne Leaflet Message."

When a promising individual is identified, the perpetrators might shift to private channels for deeper conversation driven by human operators, making it more difficult to identify who is behind a given recruitment effort. More sophisticated conversational agents might also establish different recruitment patterns, increasing the challenge of identifying all bots connected to a campaign based on similarity of posted content alone.

Controlling Virality - "Leaks and Fake News Create Market Free-Fall, Regulators Struggle to Establish Stability"

The Dow Jones experienced a fourth day of steep declines today as regulators continued to reel from a series of leaks showing deep instability in the nation's leading banks. While the identity of the leaker remains unknown, our investigation has revealed that some of the documents originally believed to genuine have in fact been false. Select journalists appear to have been targeted in a concerted campaign to spread misinformation through the manipulation of their Facebook and Twitter feeds…

Analyzing the data of key influential users might allow the creation of a model that reliably predicts when they might share information to their audiences online. Variables might include the appearance of a given piece of content from several trusted sources, the timing in which they receive content, or the degree to which a piece of content matches certain pre-conceived notions. By monitoring their activity or even probing a target with different types of content, a perpetrator might create statistical models of several important influential users, enabling them to craft a blend of true and fabricated content that is most likely to be shared widely. Here, such techniques are used to spread messages that manipulate the financial markets, with the impact potentially enhanced by algorithmic trading systems, which monitor and make trades autonomously based on content online.[32]

On-the-Ground Attacks - "Hundreds Dead in Bombing; Victims Misled by Spurious Messages Spread Online"

Emergency crews continue to comb through the wreckage downtown tonight in what is the worst terrorist attack in our nation's history. At this hour, more details continue to filter in to us - it appears that the death toll was raised significantly by a widespread hack that enabled the terrorists to send phony emergency messages throughout the city, producing panic and encouraging victims to move towards areas of the highest danger. Messages included fabricated photos of shelters and safe zones in areas targeted by the bombers. Investigators have also noted a network of spambots which successfully promoted these messages to "trending" on several social media platforms late last night. Authorities have advised citizens to be on the lookout for this fake content as the situation continues to unfold…

Cyberwarfare and the next generation of psyops might come together to significantly enhance the damage created by a terrorist attack on the ground. In this scenario, the use of a bombing produces a chance to leverage a technical vulnerability—such as those used to send public service announcements on mobile devices—for influence purposes in ways that maximize lethal impact. This might be accompanied by a parallel campaign of misinformation on social networks that seeks not to persuade with believable fake identities, but simply to manipulate the algorithms that various platforms use to identify and promote "trending" content to their users. Messaging strategies might also be adjusted on the fly by the perpetrators, enabling them to create newly effective messages even as officials attempt to stamp out misinformation spread earlier.

Challenges of Attribution - "Protesters Turn on Each Other in Bloody Street Fight"

Protests against the government appear to have conclusively broken down last night, as three rival factions within the opposition fought each other in a series of increasingly violent clashes. Disagreement has centered around adoption of Proposition A to the opposition demands, a flashpoint between the left and right wings of the party. The proposal was until recently considered a dormant issue by many members, but has once again become a hot topic on several online forums. The President took to the airwaves tonight, denouncing the violence and calling on the National Guard to enforce a curfew. "If they are unable to govern themselves," he said, "How should we expect that they will govern the nation?"

The scenarios discussed thus far are ones where the psyops campaign is focused on persuading a target constituency. However, these tactics may be similarly used to fragment and create conflict within a group, here popular opposition to a government. While this objective is not a new one in the context of psyops, attribution may become an increasing challenge as information campaigns can be deployed from anywhere in the world at low cost and with small teams. Uniquely, modern psyops can also eliminate entirely

32 *See, e.g,* Moore and Roberts, "AP Twitter Hack Causes Panic on Wall Street and Sends Dow Plunging."

the need to directly message their targets: spambots might be used simply to amplify certain controversial viewpoints by manipulating "trending" algorithms within social media platforms to promote conflict.

The upshot of all this is that it would be difficult here to ascertain the source of the campaign—whether from the incumbent president, a group allied with the president, or even a source outside the country that is uncoordinated with the current administration. Indeed, it may be difficult to determine that a campaign is happening at all as these techniques produce more and more accurate simulations of "organic" political activity.

Part II: Responses - Existing Frameworks

As the scenarios discussed above suggest, the next generation of online psyops might be more destructive than the types of techniques used in the past. At the same time, the speed, falling costs, lack of attributability, and richness of data now possible might make these techniques an increasingly attractive tool for states and non-state actors seeking to destabilize adversary governments around the world.

This appears to be already becoming the case. As detailed in a 2015 study released by NATO, Russia has made manipulation of social media a significant complement to its ongoing military operations in Ukraine.[33] Techniques have included the spreading of leaked information, fabricated news stories, and hacking to influence public opinion in Crimea.[34] This appears to be an ongoing strategy—as one NATO colonel has remarked, similar techniques also made an appearance in the 2008 invasion of Georgia.[35] If these strategies become more widespread both within and beyond the context of armed conflict, the global impact may be significant and negative.

Insofar as these online psyops campaigns are perpetrated on states by other states or non-state actors, might the existing frameworks of international law, rules, and norms be a means by which to counter their use? This section provides an analysis, examining modern psyops as described above through the context of existing rules of warfare, arrangements around criminal conduct, and rules concerning telecommunications infrastructures; and considers the possibility of regional coordination using Europe as a case study.

33 Lange-Ionatamishvili and Svetoka, "Strategic Communications and Social Media in the Russia Ukraine Conflict."
34 *Ibid.*
35 Pop, "Nato Colonel Sheds Light on Russia 'Psy-Ops.'"

While we ultimately conclude that the patchwork of legal rules fail to address these new techniques effectively, this review suggests a few alternative avenues that might be productive means of meeting the threat.

Preliminary Consideration: Sovereignty

International law does not directly deal with psyops. As one paper which examined the space in 2007 wrote simply, international law covers psyops "only by analogy and then often in a patchwork fashion."[36] To that end, in order to take advantage of international law to control the use or proliferation of these techniques, the acts or harms we seek to limit or control must be fitted into an existing framework of definitions and meanings. This presents significant challenges because in order for existing mechanisms to apply to psyops, established definitions would need to be reinterpreted in striking new ways.

For one, it may be necessary to require revision of fundamental notions of sovereignty enshrined in the structure of international law. One foundational premise of the modern framework is that international agreements operate through the consent of states, defined as juridical entities with sovereignty over a geographic area.[37] Sovereignty is understood as the full right and power to govern without interference from any outside source or body. It is sometimes termed the right of "self-determination" or "self-rule." Sovereign states are generally understood to be neither dependent on, nor subject to, any other power or state. They act independently and without outside interference.[38]

While there is no universally binding definition of precisely what a state's sovereignty applies to within these geographic bounds, the 1933 Montevideo Convention on Rights and Duties of States set forth the most widely accepted formulation of the criteria of statehood in modern international law. It notes that the state, as a legal person, should possess (a) a permanent population; (b) a defined territory; (c) government; and (d) capacity to enter into relations with other states.[39]

The important thing to note is that the accepted criteria for statehood does not mention the beliefs of the permanent population. There is no element that the population governed by the government and liv-

36 Hollis, "Why States Need an International Law for Information Operations."
37 Shaw, *International Law*, 178.
38 "Sovereign - Definition of Sovereign in English | Oxford Dictionaries."
39 "Montevideo Convention on the Rights and Duties of States," sec. 1.

ing in the defined territory should share a uniform understanding of the world distinct from that shared by other states' populations. A 2005 article considering the international legal limitations on the conduct of psyops noted that traditional notions of sovereignty would need to be "expanded" if they were to include "the hearts and minds of the people."[40]

To that end, in order to leverage existing international laws to limit or control psyops, a state would first have to take the novel position that psyops and its potential effects should be understood to impact its sovereignty in a way that would make it cognizable under international law. This would be a dramatic re-interpretation, taking the position that intrusions by other states into a nation's public opinion should be considered parallel to physical intrusions. Most international treaties address very specific issues in as narrow a way as possible, and it is unclear if such an interpretation would be accepted by other states.

As discussed below in the limited context of war, there are some exceptions governing the use of psychological techniques to affect the morale of a nation's armed forces in order to achieve a military objective. However, there is no existing standard that addresses whether a state's civilian population can be legally targeted by psyops, in war or in peacetime.

Even if we assume that international law might shift to take such a position, the existing international frameworks are insufficient to comprehensively address the specter of harm raised by the type of strong psyops that seems foreseeable in the future.

Framework: Rules of Warfare

Some of the first attempts to understand the destabilizing impact of psyops in the Internet age emerged from researchers working in the military and national security context, where it was observed that information operations presented a possible future where military goals could be accomplished in new ways and, potentially, by new actors. This resulted in a number of studies focusing on how and whether rules of armed conflict might be applied to acts of aggression carried out through psyops.[41]

Rules governing international aggression and conflict are focused on preempting and limiting armed conflict between sovereign states. There are two bodies of law that govern in this space: laws concerning *jus ad bellum* and laws concerning *jus in bello*. *Jus ad bellum* is the law that controls during peacetime and entrance into warfare; *jus in bello* sets out the boundaries of war itself. We review each in turn.

Jus in Bello

Jus in bello law would frame the analysis in a hypothetical case in which the use of psyops by one nation during wartime against an adversary nation was called into question. This situation has been analyzed in several papers from the late 1990s and early 2000s.[42] These pieces apply the law of war to psyops in order to determine whether it is permissible to use it as a weapon in warfare.

The modern law of warfare—also called international humanitarian law or IHL—has since its inception sought to restrict the aim of warfare to the achievement of military objectives. It is a standing norm of IHL that "the only legitimate object which States should endeavor to accomplish during war is to weaken the military forces of the enemy."[43] Because psyops is so customizable and can be limited to military targets, analysts have concluded that careful deployment and use of psyops would not violate the *jus in bello* framework.[44]

Despite the fact that existing frameworks do apply and govern online psyops in warfare, its application does not always produce sensible results. One illustration is in the effort by these existing papers to develop procedural safeguards that would guard against the possibility that the use of this novel 'weapon' would violate the rules of war. This led the U.S. Department of Defense to emphasize, in its 1999 report on international legal issues in information operations, that in spite of the low risk of detection and prosecution, "it is the firmly established policy of the United States that U.S. forces will fight in full compliance with the law of war" even when the "fighting" in question consists of launching a computer network attack far from its target. This implies that, at least within the context of an armed conflict, psyops 'attacks' should, whenever possible, be carried out by authorized military person-

40 Smyczek, "Regulating the Battlefield of the Future: The Legal Limitations on the Conduct of Psychological Operations (PSYOP) under International Law."
41 *See also* Lungu, "War.com: The Internet and Psychological Operations"; U.S. Department of Defense, "An Assessment of International Legal Issues in Information Operations." (hereinafter "DOD"); Hollis, "New Tools, New Rules: International Law and Information Operations.".
42 Smyczek, "Regulating the Battlefield"; Hollis, "New Tools"; DOD; Johnson, "Is It Time for a Treaty on Information Warfare?"
43 Hollis, "Why States Need an International Law for Information Operations."
44 Lungu, "War.com," 13.

nel.[45] This conclusion is based in the existing rules, which state that only *lawful combatants* may engage in armed conflict and require that lawful combatants distinguish themselves from noncombatants by the wearing of a uniform, among other requirements.[46]

Of course, modern online psyops are accomplished by operators that are not physically visible to their targets. While in the past a requirement to wear a physical uniform would help to aid in distinguishing combatants from noncombatants, to do so now seems to have little impact on the psyops "battlefield" of a social media platform or message board. This suggests that, while existing frameworks are applicable in broad terms, further interpretation and norm development will be necessary in the space to establish the visible badges of affiliation that should apply for these types of activities online in warfare. To date, these have not yet been established.

Jus ad Bellum

A second scenario covered by the laws of warfare is one in which a state uses psyops to manipulate the population of another state in order to achieve an objective while not at war. *Jus ad bellum* is the body of law that determines when it is permissible for a state to enter into armed conflict with another state.

Self-Defense

States are customarily recognized as having an inherent, sovereign right of self-defense. The concept of sovereign self-defense—and the attendant concept of just war—are enshrined in the Geneva Conventions and the Charter of the United Nations, two of the most significant multilateral agreements concerning the conduct of international disputes. Both instruments assume that there are situations in which a state is entitled to use force to defend itself, and in which a state is justified in using that force.

What acts trigger this right of self-defense, however, has been an amorphous and shifting concept in international law. The Geneva Conventions do not provide any real definition and there is no universal authority for classifying "conflicts." The Geneva Conventions demand simply that the existence of an international conflict be determined on a factual, case-by-case basis.

Article 51 of the United Nations Charter makes an "armed attack" the condition for the justified exercise of the right of self-defense[47] but does not provide a definition of the term. In 1950, the Soviet Union proposed that the UN General Assembly define the concept of "aggression" "as accurately as possible."[48] The proposal was referred to the International Law Commission for study, and in 1974 the General Assembly adopted a resolution that set out a nonexhaustive list of acts which qualify as acts of aggression.[49] These acts include "invasion or attack by the armed forces of a State of the territory of another State," "bombardment by the armed forces of a State against the territory of another State," and the "blockade of the ports or coasts of a State by the armed forces of another State."[50]

Force has historically been understood as kinetic, military force—that is, a physical incursion into the defined territory of another sovereign state. Therefore, the right of self-defense has been historically understood as justified in response to a threat to a state's *physical*, territorial sovereignty. A country that believes it has been the victim of a psyops offensive could take advantage of *jus ad bellum* if it wishes to claim that a retaliation based in sovereign self-defense is justified.

However, to do so would require taking two tenuous positions. First, the defending country must take the position that psyops is covered by existing standards governing self-defense. This means claiming that use of psyops is aggression, or a use of force, under existing law and custom. As discussed, this would require a rethinking of basic notions of sovereignty that it is unclear the international community would accept.

Also presenting a significant challenge would be the necessity for the state to take a position on what an appropriate defensive response would be, including when—or whether—an armed attack would be an appropriate defensive response to a psyops offensive.

The laws of armed conflict incorporate a fundamental concept known as "proportionality." Even in the case of justified sovereign self-defense, states are bound by norms that prohibit excessive retaliation. A response in self-defense must apply only as much force as is necessary and proportional to address the problem at hand.

45 DOD, 8.

46 *See generally* Geneva Convention; DPI, *Charter of the United Nations and Statute of the International Court of Justice*.

47 DPI, *Charter of the United Nations and Statute of the International Court of Justice*, chap. 7. "Nothing in the present Charter shall impair the inherent right of individual or collective self-defence if an armed attack occurs against a Member of the United Nations".

48 International Law Commission, "Question of Defining Aggression."

49 UN General Assembly, "A/Res/29/3314."

50 *Ibid.*

What form and level of force is implied by proportionality in responding to psyop aggression is unclear. It is possible to imagine that states will respond "in kind"—that retaliatory psyop would be seen as proportional and justified. Sanctions and the expulsion of diplomats, the approach taken by the United States in response to Russian efforts during the 2016 election, might also emerge as an accepted "proportionate" response.[51] However, a state claiming its right to self-defense could also invoke a more conventional use of force. Ultimately, there are no established norms for when or whether misinformation or propaganda campaigns might trigger the right of self-defense, and no real guidelines for determining what proportionality looks like in that context.

The potential for abuse is high, as targets might assert a high level of impact from a psyops campaign as pretextual justification for armed conflict. This is particularly in a context where attribution of an offensive online psyops is difficult to place. Again, as in the *jus in bello* case, existing frameworks might apply, but they leave undefined important questions that are likely to mitigate the instability created by advancements in online psyops.

"Propaganda for War"

Beyond international practices concerning self-defense, the International Covenant on Civil and Political Rights (ICCPR) may provide some limited control of online psyops. Adopted by the United Nations General Assembly in 1966, the ICCPR is a treaty adopted by 168 different countries that commits parties to protect the civil and political rights of their citizens.[52] Compliance is monitored by the United Nations Human Rights Committee (HRC), which evaluates complaints and can request that nations party of the treaty provide a remedy for breaches of the ICCPR commitments.

In particular, Article 20 of the ICCPR states that "[a]ny propaganda for war shall be prohibited by law" and further that "[a]ny advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law."[53] This is a broad provision: the UN Office of the High Commissioner for Human Rights has commented that Article 20 is targeted at "all forms of propaganda threatening or resulting in an act of aggression or breach of the peace contrary to the Charter of the United Nations."[54] Moreover, this prohibition applies regardless of whether or not the advocacy has "aims which are internal or external to the State concerned."[55]

To that end, the ICCPR may be applicable to states that engage in online psyops that are directed towards the specific ends of triggering warfare, violence, or discrimination. While this captures some of the potential applications that have been observed in the realm of online psyops, the series of examples and hypothetical scenarios described above should make clear that these techniques can cause significant harm without necessarily inciting violence. Online psyops may attempt to undermine faith in a target government, destabilize markets, and spread misinformation that leads citizens into danger, all without necessarily falling within the type of incitement campaigns contemplated under Article 20.

Even when engaging in online psyops falling within the ambit of the ICCPR, practical considerations may inhibit the effectiveness of the treaty in combatting these activities. As one study of Article 20 in 2007 concluded, "[m]any states have refused to give effect to this provision...primarily on the grounds that 'propaganda for war' is not adequately defined, and that it constitutes an unacceptable threat to the right of freedom of expression."[56] With incomplete implementation of the ICCPR in the national laws of the countries party to the treaty, it may be difficult to adequately enforce these rules in practice.

Moreover, the HRC has little power to shape the behavior of states and has declared itself to understaffed.[57] Even if a breach is found in response to a complaint, the primary recourse of the HRC is to simply request reports from the party nation concerning measures taken to remedy the violation. The HRC also depends on the voluntary participation of the party nations to receive reports about compliance with the ICCPR. As of 2015, the HRC reported that 79 countries were overdue on these reports, with 42 countries being more than 10 years overdue or having never

51 Northam, "Obama Expels 35 Diplomats, Imposes Retaliatory Sanctions Against Russia For Hacking."
52 Article 19 - which concerns freedom of expression - may have application to psyops which are conducted by a government against its own citizens, though this analysis lies beyond the scope of this paper. *See* "ICCPR."
53 Ibid.
54 Office of the High Commissioner for Human Rights, "Prohibition of Propaganda for War and Inciting National, Racial or Religious Hatred (Art. 20): . 29/07/1983. CCPR General Comment No. 11."
55 Ibid.
56 Kearney, *The Prohibition of Propaganda for War in International Law*, 19.
57 UN Human Rights Commission, "Report of the Human Rights Commission."

submitted a report.[58] One researcher has declared that the body is simply, "basically weak and ineffectual."[59]

Considering Non-State Actors

There are further challenges here—it is also important to consider that both *jus in bello* and *jus ad bellum* are legal constructs which apply to states, leaving uncovered an entire category of conflicts in which non-state actors target states. It has been observed that, given the historical context of the UN Charter, the drafters likely limited their consideration to governments since "at that time, they alone had armed forces at their disposal which could launch an armed attack."[60]

That is no longer the case in the case of psyops. It is possible for non-state actors to employ tools and techniques of psyop that are as complex and sophisticated as those employed by state actors. To that end, other legal tools will be required to potentially address the challenge presented by these technologies, and the existing international framework around criminal conduct may be relevant in this regard.

Framework: Rules Concerning Criminal Conduct

Since the rules of war do not readily apply to situations involving non-state aggressors, the next logical step is to look to international criminal prohibitions. A private individual or group that causes this type of destabilizing harm could, perhaps, be guilty of a crime and subject to prosecution.

As a general matter, the baseline assumption of international law is that sovereign states have total control over their own criminal laws, as an inherent element of self-governance. In the absence of any agreement, a police force entering into a sovereign territory and arresting a member of another state's population could be considered aggression. These agreements aim to limit international conflict by clarifying how and when a state's law enforcement activity is permitted to extend into another state's territory. Otherwise, case-by-case diplomatic agreements are usually the only other means to achieve these kinds of outcomes.

There are three primary types of agreement among states designed to facilitate the investigation and prosecution of crime. First, there are extradition agreements, which provide that participating states will assist other participants in finding wanted criminals and transferring them for prosecution. Second, there are mutual legal assistance treaties—also called judicial assistance agreements—which oblige member states to help each other with criminal investigations by gathering and providing to the prosecuting state any evidence that might be found within a member state's territory. Third, there are international crimes, broadly defined as crimes prosecutable by any nation or by international bodies.

General Challenges

Across all three of these types of agreements, there are significant challenges to addressing the challenge of online psyops. For one, psyops is not precisely covered across any existing legal assistance, extradition treaty, or framework around international crime. Only a few mutual legal assistance treaties apply broadly to *all* law enforcement investigations and prosecutions. Most are targeted narrowly at specific types of criminal behavior of concern, such as transnational drug trafficking, money laundering, or arms dealing. To that end, the psyop activity in question would have to substantially resemble the criminal acts prohibited by the laws themselves. This leads to a landscape of highly limited applicability, as discussed below.

Even if there was common agreement in the international community that a new rule around criminal conduct which would enable prosecution for the execution of psyops was needed, there are a number of significant challenges. First, mutual legal assistance treaties and extradition treaties often require states parties to make the definitions of relevant criminal acts uniform across their separate legal codes. To date, an international consensus around the elements of psyops-related crimes is not established and would need to be developed to implement such a rule.

Second, it is unclear if such a legal framework could see sufficiently broad adoption. Many of the existing rules around criminal conduct have limited geographic applicability, applying solely to the nations that have become party to a given agreement. Notably, some of these rules are not legally binding on significant state actors such as the United States, which has declined to submit to the jurisdiction of any international court.

However, online psyops can be inexpensively deployed from any location with a connection to the Internet. Either a new agreement defining the elements of the criminal act of psyops, or a very broad agreement covering all types of criminal activity, would need to be ratified by a large number of states in order for crim-

58 Ibid., Chapter 3.
59 Mutua, "Looking Past the Human Rights Committee."
60 Zemanek, "Armed Attack."

inal prohibitions to functionally limit the use or proliferation of these techniques. Both would break with the historical pattern seen in these types of agreements.

Limited Applicability

Despite these challenges, there is some limited applicability in a few narrow cases under the existing framework. These rules would only apply to psyops in circumstances where the act also met all the other elements of the specific crime covered by them.

For example, the International Convention for the Suppression of Acts of Nuclear Terrorism makes it a crime to threaten to use nuclear material "under circumstances which indicate the credibility of the threat."[61] One might imagine circumstances under which a non-state actor engaged in an online psyops campaign that centered around the creation of panic about the presence of nuclear devices within the target country. Fabricating "fake news" and ensuring its circulation in a manner which created a credible threat might enable the use of this convention to prosecute these actors.

There are also international crimes. These are violations of customary international law carried out by private individuals, as opposed to by states. As in the examples above, a psyops would have to include all the other required elements of the crime in order to be prosecuted under international law.

The most infamous of these international crimes is genocide, where there is an applicable legal precedent finding liability for the deliberate manipulation of a society through information. In 2003, the International Criminal Tribunal for Rwanda (ICTR) found three individuals guilty of the crime of genocide, based on the specific content aired on their radio and television station. Their radio and television station, RTLM, "called on listeners to seek out and take up arms against the enemy...defined to be the Tutsi ethnic group."[62] Substantial evidence was produced in that case to support the prosecution's claim that much of the anti-Tutsi violence in Rwanda was directly attributable to the media content broadcast by RTLM. Assuming that sufficient causal evidence could be gathered, parallel cases in the future may be treated similarly by international bodies.

61 "ICSANT."
62 *Prosecutor v. Nahimana.*

Framework: Rules Concerning Telecommunications Infrastructures

A final potential international framework for managing the risk from online psyops as it continues to evolve is international telecommunications law. The primary governing body in this context is the International Telecommunications Union (ITU), a specialized agency within the United Nations focusing on information and communication technologies.[63] The ITU engages in promulgating rules for telecommunication systems, including online channels that are host to the emerging techniques around online psyops.

While ostensibly a relevant body, the ITU will likely not serve as an effective bulwark against the increasing sophistication of online psyops. For one, it remains disputed the extent to which the services flowing through the Internet and public policy questions fall under the standards-setting jurisdiction of the body. The International Telecommunications Convention of 1982 remains the primary instrument of international law governing telecommunications infrastructure, though it has been amended and expanded over time.[64]

In 2012, an effort was made to incorporate the Internet more explicitly into the treaty language of the ITU under the justification that it traveled through the telecommunications networks traditionally under the purview of the organization.[65] However, this was strongly rejected by the United States, Germany, Japan, Canada, and others, arguing that to do so would upend the bottom-up governance of the network currently in force.[66]

These countries have refused to sign the new treaty, leaving major stakeholders—many of whom are host to the platforms being leveraged to execute the latest generation of psyops—beyond the reach of the ITU. The ITU itself remains largely focused on issues of technical standards and access with regards to the Internet: as of the time of writing, its primary working group on public policy questions surrounding the Internet are focused on domain name governance and development aspects of the technology.[67] It also seems possible that Russia and China, both of whom appear

63 "Constitution and Convention of the International Telecommunication Union."
64 International Telecommunication Union, "Nairobi Convention."
65 International Telecommunication Union, "Final Acts - World Conference on International Telecommunications (Dubai 2012)."
66 Pfanner, "Citing Internet Standoff, U.S. Rejects International Telecommunications Treaty."
67 International Telecommunication Union, "Council Working Group on International Internet-Related Public Policy Issues."

to have made investments in online psyops, would resist efforts to regulate this activity through the ITU despite being signatories to the new 2012 rules.[68]

One Potential Model: Regional Coordination

While considerable challenges exist to adapting current international legal frameworks to meet the challenges presented by contemporary online psyops, it may be possible that regional coordination may prove to be more effective. Countries in a given region may share more common interests with regards to the threats posed by the use of these techniques from particular state and non-state actors, and may have similar preferences with regards to what should and should not be permitted in their use, as well.

As an illustrative case study, we examine how the legal framework surrounding three international bodies in Europe—the North American Treaty Organization (NATO), the Organization for Security and Co-operation in Europe (OSCE), and the European Union (EU)—has been and might potentially contend in the future with the threat from online psyops. While the result of our analysis is that many of the broader challenges discussed above also appear here, these organizations may become promising platforms for facilitating narrower coordination on these issues going forwards. Particularly in the case of NATO and the EU, preliminary action to respond to online psyops has already begun, though typically in response to specific threats rather than as part of a general regulatory approach.

NATO

Central to the North Atlantic Treaty—NATO's foundational treaty—is Article 5, which articulates the principle of collective defense that "an armed attack against one or more of them in Europe or North America shall be considered an attack against them all."[69] This principle derives legal power from Article 51 of the United Nations Charter, which enshrines the principles of sovereignty and self-defense discussed above.[70]

In that respect, the question of whether an online psyops campaign could trigger action under NATO Article 5 is closely tied to the question of whether or not these campaigns might be considered equivalent to an "armed attack" under customary international law. As described above, this has not traditionally been the case. "Armed attack" has typically referred to kinetic, military force, not acts of strategically targeted persuasion. This corresponds to the record of Article 5, which has been affirmatively invoked only once, on September 12, 2001, in the immediate aftermath of coordinated terrorist attacks against the US.[71] For the reasons described above, assertion of the legal equivalence of kinetic attacks and online psyops would raise thorny challenges around sovereignty and proportionality of response, making such an act unlikely. NATO appears to have proceeded along these lines: Article 5 has not yet been invoked despite the fact that several member states have been targeted by online psyops of varying sophistication.[72]

However, even without affirmative invocation of Article 5, NATO might still play a role in developing a response to the emerging techniques in online psyops. Indeed, the principle of collective defense animates a great deal of other NATO activity such as the reinforcement of the defenses of other member states, which can request assistance under the treaty.

There are a number of ongoing initiatives in this respect. In January 2014, NATO launched its Strategic Communications Centre of Excellence (StratCom COE)—an independent body supported by 11 countries and focusing on disseminating expertise on psyops, information operations, public affairs, and other related topics.[73] As of the time of writing, much of the recent published work of StratCom COE has focused on tracking the techniques used by Russia in Ukraine, Syria, and Moldova.[74] Reports in early 2016 also suggested that the NATO Military Committee was considering a policy of strategic communications to combat Russian "weaponization of information."[75] Further efforts to coordinate a strategic response to counter Russian activities in this space seem likely, beyond thornier questions around whether a future campaign might be grounds for triggering a response under Article 5.

OSCE

Founded in 1973 as the Conference on Security and Co-operation in Europe, the OSCE is a regional security organization with participation from 57 states

68 *Supra*, notes 7-8.
69 NATO, "The North Atlantic Treaty."
70 *Supra*, notes 39-40.
71 "NATO - Topic: Collective Defence - Article 5."
72 Gotev, "Commission." Naming Poland, the Czech Republic, Slovakia, and Hungary all as countries targeted by Russian psyops.
73 NATO Strategic Communications Centre of Excellence, "About Strategic Communications."
74 NATO STRATCOM, "Publications | StratCom."
75 Emmott, "NATO Looks to Combat Russia's 'Information Weapon.'"

in North America, Europe, and Asia. In the past, the OSCE has played a role in issues as diverse as arms control, human rights, policing strategies, counter-terrorism and economic and environmental activities.76

OSCE may be a potential forum for greater coordination and action around meeting the challenge posed by online psyops. The organization currently houses the OSCE Representative on Freedom of the Media, an institution which has a mandate to "protect and promote media freedom in all 57 OSCE participating States."77 To date, these activities have included "observing media developments as part of an early warning function and helping participating States abide by their commitments to freedom of expression and free media."78

Momentum currently exists for the Representative to take a broader role in policy debates around the evolution of online psyops. In 2014, the OSCE Representative issued a communiqué expressing alarm at the use of propaganda in the Ukrainian conflict and advocating for participating OSCE states to "stop manipulating media; stop information and psychological wars."79 The Representative has also promoted regulation of propaganda on the basis of Article 20 of the ICCPR and in the founding charter of the OSCE, which expressed a commitment to create "a climate of confidence and respect among peoples consonant with their duty to refrain from propaganda for wars of aggression."80 In 2016, the Representative called for a renewed international dialogue on propaganda for war given the new technological environment.81

Perhaps the greatest challenge to the OSCE taking a more significant role in these debates is its organizational structure. Since its founding charter was not a treaty but a political commitment, decisions of the OSCE depend on the consensus of its members and do not have legally binding effect.82 To that end, the organization would depend on its ability to produce a common agreement about these issues among participating states. In doing so, the OSCE may confront many of the legal challenges discussed above in the attempt to create more concrete action on these issues.

EU

One potential path for greater coordination around online psyops may be for action to be taken in response to acute threats or concerns, rather than a general agreement which would cover broadly the use of techniques for strategic persuasion. This appears to be the evolving approach in the EU, which has been most focused on the deployment of online psyops by Russia and ISIS to advance their aims.

Responding to a call by the European Council in March 2015 to "challenge Russia's ongoing disinformation…[and prepare] an action plan on strategic communication," the European External Action Service (EEAS)—the EU's foreign ministry—formally established a team to monitor and counter Russian psyops activities in Eastern Europe.83 The EEAS published an Action Plan later that year which detailed an effort to increase public awareness of disinformation campaigns, build networks of allied communicators, and support increased pluralism in the Russian language media space, among other initiatives.84 These efforts have broadened over time, with the EEAS informing the European Parliament and Council in 2016 that it intended to work with Member States to increase their capacity to "deliver proactive strategic communications and optimise use of media monitoring and linguistic specialists."85 To date, this work has included the creation of projects like The Disinformation Review, which seek to counter pro-Russian misinformation as new stories emerge.86

The European Parliament has supported these moves, passing a resolution in November 2016 which supported the continuation and expansion of these activities.87 The Parliament also highlighted the use of propaganda by ISIL/Daesh and Al-Qaeda to "promote its political, religious, social, hateful and violent narratives" and recruit members.88

76 Organization for Security and Co-operation in Europe, "OSCE | Organization for Security and Co-Operation in Europe."
77 Organization for Security and Co-operation in Europe, "Media Freedom and Development | OSCE."
78 Ibid.
79 Mijatovic, "Communiqué by OSCE Representative on Freedom of the Media on Propaganda in Times of Conflict | OSCE."
80 Richter, "The Relationship between Freedom of Expression and the Ban on Propaganda for War"; OSCE Representative on Freedom of the Media, "Propaganda and Freedom of the Media."
81 Ibid.
82 Mijatovic, "Recommendations Following the Expert Meeting Propaganda for War and Hatred and Freedom of the Media."
83 Gotev, "Tiny EU Task Force Set up to Counter Russian Propaganda."
84 European External Action Service, "Action Plan on Strategic Communication."
85 European External Action Service, "JOINT COMMUNICATION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL - Joint Framework on Countering Hybrid Threats a European Union Response."
86 European External Action Service, "EU vs Disinformation | Don't Be Deceived, Question Even More."
87 Committee on Foreign Affairs, EU strategic communication to counteract anti-EU propaganda by third parties.
88 Ibid.

It urged action to counter these campaigns, as well. It remains to be seen how this resolution shall be implemented in practice, particularly in light of statements by Russia that it will retaliate in response to efforts to "curb the activities of the Russian media on EU member states."89 However, at the time of writing, it appears that the EU may be poised to take further targeted action on the issue of online psyops, particularly as it is used by a specific set of strategic actors of concern.

Part III: Responses - Alternative Approaches

As discussed above, the existing framework of international laws, rules, and norms have typically not covered psyops explicitly in the past. One result of this standing silence on the issue is that while there are a relatively small set of specific circumstances under which existing rules would apply, by and large these activities are not addressed comprehensively by the international community and are only beginning to be addressed regionally.

Reticence to address this issue head-on may have been appropriate in the past, particularly in an earlier era in which psyops tactics were less measurable in their impact and conflict in physical battlefields was the central focus. However, to the extent that the trajectory of technology raises the concern that online psyops will become increasingly powerful and destabilizing in coming years, it may be appropriate for nations to develop new approaches to limit and control the use of these technologies by state and non-state actors. This section describes and evaluates the pros and cons of a series of different proposals that might be productive to explore as interventions in the space.

Improving Transparency: Investigatory Groups and Agreements

While modern psyops techniques discussed above take a number of different forms, one common challenge they present is the issue of attribution. It can be difficult to ascertain when a campaign is ongoing, and—if it is—the identity of the perpetrators. As in the case of cyber warfare, limited attribution might act as a particular incentive for actors to engage in these campaigns, since it provides a means of undermining adversaries without loss of international reputation or credibility.

To that end, interventions that help to increase the level of transparency in the space might help to disincentivize their use. One approach may be for nations concerned about these techniques to collaboratively fund the creation of an independent, international investigative agency that would conduct monitoring and forensics work to uncover online psyops campaigns and their perpetrators. This agency would bring together computer security experts, journalists, quantitative social scientists, law enforcement, and others to develop best of breed detection methodologies for online psyops techniques as they appear. At the core of this center would be a series of reports helping to "name and shame" actors engaging in these campaigns and exposing emerging strategies being used in the space. This agency could also act in an expert advisory capacity to help supporting nations develop defensive capabilities against these techniques.

In lieu of an entirely new organization or program, nations might also develop collaborative intelligence agreements to assist each other in detection, analysis, and resistance against these campaigns. These may be easier to achieve, since the creation of a truly independent investigatory agency with the mandate to publish findings publicly would raise the risk that one of the supporting nations might be exposed in their own psyops operations. However, these bilateral or multilateral agreements might only achieve limited transparency to the international community at large, reducing the desirable disincentive to engage in these techniques at all.

Enhancing Public Robustness Against These Techniques

Psyops aims at public opinion. By spreading misinformation or creating the appearance of a mass constituency, perpetrators of these operations attempt to shape perceptions and influence group behavior. To that end, one means of reducing the destabilizing impact of online psyops may be to find ways of "inoculating" the public at large against these techniques so they are less effective.

This intervention might take two primary forms. For one, nations might work collaboratively to encourage greater media literacy—helping the public understand the possibility that online platforms might be leveraged for the purposes of psyops campaigns. Simple awareness that these techniques can be deployed, and knowledge about some of the common patterns, may play a role in limiting their effectiveness in shaping belief and behavior in the near term.

Secondly, nations might invest in user-friendly, open-source tools that help users to navigate informational

89 Samuels, "EU Votes to Fight Back against Russian 'Propaganda Warfare.'"

sources online. Tools might include the creation of a browser extension that helps to evaluate whether a "trend" observed online has suspicious provenance, or gives users the ability to quickly examine the past behavior of an account engaging in persuasive behavior online. Users might also play a role in helping to identify suspicious behavior, helping to alert researchers or investigative organizations like the one described above to emerging campaigns that may be in progress.

This approach is largely limited by adoption. It is unclear if media literacy campaigns or tools would reach sufficient numbers of the public to influence the overall effectiveness of modern psyops, particularly as the techniques become more sophisticated over time. While this may help in at least reducing the impact of more rudimentary campaigns that rely on simple bots, it may not serve as a long-term solution.

Changing Platform Behavior

The internal policies of online platforms play a large role in shaping the ecosystem in which online psyops campaigns take place and defining if they will be more or less effective. Facebook policies, for instance, were considered to have contributed to an economic ecosystem around "fake news" that may have shaped the 2016 US election result.90 Similarly, the relatively liberal policy of Twitter towards bots has also been seen as one reason the platform has played host to a rich ecosystem of them in recent years.91

Shaping these policies may play an important role in making modern psyops campaigns more or less attractive to state and non-state actors. Nations might collaborate to find ways of encouraging or pressuring platforms to change policies towards this end. These interventions might include more proactive interventions on the part of the platform to detect and halt misinformation, more stringent requirements on connecting accounts to real-world identities, increased disclosure to authorities about activity on the platform, or more active banning of accounts found to be engaging in these campaigns.

Of course, these interventions come with their own challenges. For one, companies are likely to resist many of these changes, since they will tend to reduce the ease with which new users are able to join the platform and thereby slow platform growth. Platforms may also not wish to become engaged in policing this type of campaigning and taking on the editorial role of determining what is and is not factual.92 Secondly, even if the platforms were willing to take on this role, heavy-handed or poorly crafted policies may work to limit or chill political speech on these platforms, which may outweigh the benefits of curtailing online psyops campaigns.

Post Hoc Interventions

The interventions discussed above attempt to prevent online psyops campaigns prospectively before they are even launched, or hinder them while they are in progress. However, there are a series of approaches that may be tried that influence the state of play after a campaign has been completed. These may be useful to consider in part because the full damage produced by a campaign may not be known until it is completed, and because it may take time for investigators to attribute the techniques to a particular actor.

As discussed above, for non-state actors, the international community might move towards agreements that enable criminal extradition for individuals and groups engaging in psyops campaigns towards other nations. This might raise the stakes for groups attempting to execute campaigns of government destabilization across borders, particularly if any of the above techniques increase the ability for nations to attribute these efforts effectively. Nations may also have aligned interests in these cases to constrain third-parties who may be offering sophisticated psyops techniques for hire—as in the Sepúlveda case mentioned above—and ensuring that governments maintain a relative monopoly on the most advanced approaches in the space.

For state actors, the international community might recognize a system by which nations may be able to retrieve some form of compensation for the harms generated by psyops campaigns. This would recognize that the use of these techniques on some level is challenging to preclude entirely, and that attempts to shape online platforms themselves may not be particularly effective in hindering these campaigns. In the alternative, the international community would establish a cognizable monetary penalty that targeted countries could pursue that would raise the costs of supporting and deploying a defined set of impermissible persuasive techniques. Disputes of this kind could then be integrated into the inter-

90 Lee, "Facebook's Fake News Problem, Explained"; Silverman, "This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook."
91 Hernandez, "Why Can't Twitter Kill Its Bots?"

92 *See, e.g,* Lessin, "Facebook Shouldn't Fact-Check." Article suggesting that users may not want them to take on this editorial role, either.

national system of courts, leveraging existing and credible institutions to help enforce this agreement.

Bans and Disclosure Regimes

Perhaps the most dramatic interventions would involve categorical bans by the international community from using certain types of persuasive techniques, and mandatory disclosure regimes that would attempt to enforce such a ban or at least track national behavior around certain types of online psyops.

However, such an approach seems unlikely. Psyops has been a long-standing component of national military strategy, and it is difficult to cleanly distinguish the modern generation of online techniques from ones in the past, even though they may grow more effective over time. Even assuming international willingness to take up such a ban, this inherent ambiguity may make it difficult to enforce, or at the very least easy to evade. Moreover, these agreements would leave aside the issue of non-state actors using these techniques, an important part of the strategic landscape as the cost of executing the campaigns continues to fall. For these reasons, it does not seem likely that this will be an effective means of approaching the potential instability presented by these technologies.

Conclusion

Concerns around the international impact of online psyops is not new. Responding to a 1999 UN Secretary-General call for comments on developments in telecommunications in the context of international security, Russia wrote of its concern around "information weapons," defined as including "use of information to the detriment of a State's defence, administrative, political, social, economic or other vital systems, and the mass manipulation of a State's population with a view to destabilizing society and the State."93 At the time, it called for an international legal basis for identifying and creating a means of "[p]reventing the threat of the use of information of technologies and means to influence social consciousness" with a view towards destabilization.94

These proposals did not advance at the time, but in light of the modern context, it may be necessary for the international community to reexamine these calls for action. The sphere of online activity has only expanded since the early 2000s, and the techniques for manipulation of public opinion have only become more sophisticated. As psyops continues to advance, the trajectory of the technology seems poised to encourage greater adoption of these techniques by state and non-state actors, and for their disruptive impact to grow over time. Particularly in the wake of the revelations surrounding Russian involvement in the 2016 US presidential election, identifying effective approaches and common norms to address the use of these techniques may now be more urgent than ever.

As the discussion around cyberwarfare and the international norms around it continues to evolve, it is critical that the development and deployment of persuasive arms become part of the discussion. Any single intervention will not serve as a "silver bullet" given the great number of potential actors and techniques at play, but further investigation—along the lines outlined above and beyond—should be prepared as the ecosystem continues to evolve rapidly.

93 UN General Assembly, "Developments in the Field of Information and Telecommunications in the Context of International Security."

94 *Ibid.*

Bibliography

Andrews, Natalie. "Pro-Trump Twitter Bots at Center of Nevada Mystery." *WSJ*, February 26, 2016. http://blogs.wsj.com/washwire/2016/02/25/pro-trump-twitter-bots-at-center-of-nevada-mystery/.

Associated Press. "US Secretly Created 'Cuban Twitter' to Stir Unrest and Undermine Government." *The Guardian*, April 3, 2014, sec. World news. https://www.theguardian.com/world/2014/apr/03/us-cuban-twitter-zunzuneo-stir-unrest.

Bilton, Nick. "Social Media Bots Offer Phony Friends and Real Profit." *The New York Times*, November 19, 2014. https://www.nytimes.com/2014/11/20/fashion/social-media-bots-offer-phony-friends-and-real-profit.html.

Bond, Robert M., Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime E. Settle, and James H. Fowler. "A 61-Million-Person Experiment in Social Influence and Political Mobilization." *Nature* 489, no. 7415 (September 13, 2012): 295–98. doi:10.1038/nature11421.

Chen, Adrian. "The Agency." *The New York Times*, June 2, 2015. https://www.nytimes.com/2015/06/07/magazine/the-agency.html.

Cheng, Justin, Lada Adamic, P. Alex Dow, Jon Michael Kleinberg, and Jure Leskovec. "Can Cascades

Be Predicted?" In *Proceedings of the 23rd International Conference on World Wide Web*, 925–936. WWW '14. New York, NY, USA: ACM, 2014. doi:10.1145/2566486.2567997.

Committee on Foreign Affairs. EU strategic communication to counteract anti-EU propaganda by third parties, Pub. L. No. P8_TA-PROV(2016)0441(n.d.).http://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P8-TA-2016-0441&language=EN&ring=A8-2016-0290.

"Constitution and Convention of the International Telecommunication Union." *ITU*. Accessed January 17, 2017. http://www.itu.int:80/en/history/Pages/ConstitutionAndConvention.aspx.

DeepMind. "WaveNet: A Generative Model for Raw Audio." *DeepMind*. Accessed January 16, 2017. https://deepmind.com/blog/wavenet-generative-model-raw-audio/.

Dewey, Caitlin. "How Online Bots Conned Brexit Voters." *The Washington Post*, June 27, 2016. https://www.washingtonpost.com/news/the-intersect/wp/2016/06/27/how-online-bots-conned-brexit-voters/?utm_term=.92b1b1bead1d.

DPI. *Charter of the United Nations and Statute of the International Court of Justice*, 2015.

Emmott, Robin. "NATO Looks to Combat Russia's 'Information Weapon': Document." *Reuters*. January 27, 2016. http://www.reuters.com/article/us-nato-reform-idUSKCN0V51RU.

European External Action Service. "Action Plan on Strategic Communication." European External Action Service, June 22, 2015. http://archive.eap-csf.eu/assets/files/Action%20PLan.pdf.

———. "EU vs Disinformation | Don't Be Deceived, Question Even More." Accessed January 17, 2017. https://euvsdisinfo.eu/.

———. "JOINT COMMUNICATION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL - Joint Framework on Countering Hybrid Threats a European Union Response," June 4, 2016. http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52016JC0018.

Facebook. "Choose Your Audience." *Facebook Business*. Accessed January 16, 2017. https://www.facebook.com/business/products/ads/ad-targeting.

———. "Company Info | Facebook Newsroom." Accessed January 16, 2017. http://newsroom.fb.com/company-info/.

Fielding, Nick, and Ian Cobain. "Revealed: US Spy Operation That Manipulates Social Media." *The Guardian*, March 17, 2011, sec. Technology. https://www.theguardian.com/technology/2011/mar/17/us-spy-operation-social-networks.

Finley, Klint. "Pro-Government Twitter Bots Try to Hush Mexican Activists." *WIRED*. Accessed January 16, 2017. https://www.wired.com/2015/08/pro-government-twitter-bots-try-hush-mexican-activists/.

Friedman, Herbert A. "Falling Leaves." *PsyWar.Org*. Accessed January 16, 2017. https://www.psywar.org/fallingleaves.php.

Gaffney, Alexander Furnas and Devin. "Statistical Probability That Mitt Romney's New Twitter Followers Are Just Normal Users: 0%." *The Atlantic*, July 31, 2012. http://www.theatlantic.com/technology/archive/2012/07/statistical-probability-that-mitt-romneys-new-twitter-followers-are-just-normal-users-0/260539/.

Geneva Convention (1949). https://www.icrc.org/en/document/geneva-conventions-1949-additional-protocols.

Gotev, Georgi. "Commission: Russian Propaganda Has Deeply Penetrated EU Countries." *EurActiv.com*, July 14, 2016. https://www.euractiv.com/section/global-europe/news/thurs-commission-official-russian-propaganda-has-deeply-penetrated-eu-countries/.

———. "Tiny EU Task Force Set up to Counter Russian Propaganda." *EurActiv.com*, August 28, 2015. https://www.euractiv.com/section/global-europe/news/tiny-eu-task-force-set-up-to-counter-russian-propaganda/.

Hernandez, Daniela. "Why Can't Twitter Kill Its Bots?" *Fusion*, September 21, 2015. http://fusion.net/story/195901/twitter-bots-spam-detection/.

Hollis, Duncan B. "New Tools, New Rules: International Law and Information Operations." In *The Message of War: Information, Influence and Perception in Armed Conflict*, edited by T. McKeldin and G. David, 2008.

———. "Why States Need an International Law for Information Operations." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, January 17, 2008. https://papers.ssrn.com/abstract=1083889.

Holt, Thaddeus. *The Deceivers*, n.d.

"International Convention for the Suppression of Acts of Nuclear Terrorism." United Nations, 2005.

https://treaties.un.org/doc/db/terrorism/english-18-15.pdf.

"International Covenant on Civil and Political Rights," n.d. http://www.ohchr.org/Documents/ProfessionalInterest/ccpr.pdf.

International Law Commission. "Question of Defining Aggression." Accessed January 17, 2017. http://legal.un.org/ilc/summaries/7_5.shtml.

International Telecommunication Union. "Council Working Group on International Internet-Related Public Policy Issues." *ITU*. Accessed January 17, 2017. http://www.itu.int:80/en/council/cwg-internet/Pages/default.aspx.

———. "Final Acts - World Conference on International Telecommunications (Dubai 2012)." International Telecommunication Union, 2012. http://www.itu.int/en/wcit-12/Documents/final-acts-wcit-12.pdf.

———. "International Telecommunication Convention." General Secretariat of the International Telecommunication Union, 1982. http://www.itu.int/dms_pub/itu-s/oth/02/09/S020900000B5201PDFE.PDF.

Johnson, Philip A. "Is It Time for a Treaty on Information Warfare?" *International Law Studies* 76, no. Computer Network Attack and International Law (2002): 439–55.

Kearney, Michael. *The Prohibition of Propaganda for War in International Law*. OUP Oxford, 2007.

Kelly, Mary Louise. "FBI Agrees With CIA On Russian Interference In Presidential Election." *NPR*, December 16, 2016. http://www.npr.org/2016/12/16/505892960/fbi-agrees-with-cia-on-russian-interference-in-presidential-election.

King, Gary, Jennifer Pan, and Margaret E. Roberts. "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument." *American Political Science Review*, n.d.

Kumar, Srijan, West Robert, and Leskovec, Jure. "Disinformation on the Web: Impact, Characteristics, and Detection of Wikipedia Hoaxes," n.d. https://cs.stanford.edu/people/jure/pubs/hoax-www16.pdf.

Lange-Ionatamishvili, Elina, and Sanda Svetoka. "Strategic Communications and Social Media in the Russia Ukraine Conflict." In *Cyber War in Perspective: Russian Aggression against Ukraine*. Tallinn: NATO CCD COE Publications, 2015.

Lee, Timothy B. "Facebook's Fake News Problem, Explained." *Vox*, November 16, 2016. http://www.vox.com/new-money/2016/11/16/13637310/facebook-fake-news-explained.

Lessin, Jessica. "Facebook Shouldn't Fact-Check." *The New York Times*, November 29, 2016, sec. Opinion. https://www.nytimes.com/2016/11/29/opinion/facebook-shouldnt-fact-check.html.

Linebarger, Paul. *Psychological Warfare*. Second., 1954. http://www.gutenberg.org/files/48612/48612-h/48612-h.htm.

Lungu, Angela M. "War.com: The Internet and Psychological Operations." *Joint Force Quarterly* 28 (September 2001): 13–17.

Mijatovic, Dunja. "Communiqué by OSCE Representative on Freedom of the Media on Propaganda in Times of Conflict | OSCE," April 15, 2014. http://www.osce.org/fom/117701.

———. "Recommendations Following the Expert Meeting Propaganda for War and Hatred and Freedom of the Media." Recommendations of the OSCE Representative on Freedom of the Media. Vienna, March 1, 2016. http://www.osce.org/fom/225351?download=true.

"Montevideo Convention on the Rights and Duties of States." Multilateral Treaty. Montevideo, Uruguay, December 26, 1933. Wikisource.

Moore, Heidi, and Dan Roberts. "AP Twitter Hack Causes Panic on Wall Street and Sends Dow Plunging." *The Guardian*, April 23, 2013, sec. Business. https://www.theguardian.com/business/2013/apr/23/ap-tweet-hack-wall-street-freefall.

Mulford, Carla. "Benjamin Franklin's Savage Eloquence: Hoaxes from the Press at Passy, 1782." *Proceedings of the American Philosophical Society* 152, no. 4 (2008): 490–530.

Mutua, Makau W. "Looking Past the Human Rights Committee: An Argument for De-Marginalizing Enforcement." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 1998. https://papers.ssrn.com/abstract=1527474.

NATO. "The North Atlantic Treaty." *NATO*. Accessed January 17, 2017. http://www.nato.int/cps/en/natohq/official_texts_17120.htm.

"NATO - Topic: Collective Defence - Article 5." Accessed January 17, 2017. http://www.nato.int/cps/en/natohq/topics_110496.htm.

NATO STRATCOM. "Publications | StratCom." Accessed January 17, 2017. http://www.stratcomcoe.org/publications.

NATO Strategic Communications Centre of Excellence. "About Strategic Communications," n.d. http://www.stratcomcoe.org/about-strategic-communications.

Northam, Jackie. "Obama Expels 35 Diplomats, Imposes Retaliatory Sanctions Against Russia For Hacking." *All Things Considered*, December 29, 2016. http://www.npr.org/2016/12/29/507436692/obama-expels-35-diplomats-imposes-retaliatory-sanctions-against-russia-for-hacki.

Office of the Director of National Intelligence. "Assessing Russian Activites and Intentions in Recent US Elections," January 6, 2017. https://www.dni.gov/files/documents/ICA_2017_01.pdf.

Office of the High Commissioner for Human Rights. "Prohibition of Propaganda for War and Inciting National, Racial or Religious Hatred (Art. 20): . 29/07/1983. CCPR General Comment No. 11.," 1983. http://www.ohchr.org/Documents/Issues/Opinion/CCPRGeneralCommentNo11.pdf.

Ohana, Mor, Orr Dunkelman, Stuart Gibson, and Margarita Osadchy. "HoneyFaces: Increasing the Security and Privacy of Authentication Using Synthetic Facial Images." *arXiv:1611.03811 [Cs]*, November 11, 2016. http://arxiv.org/abs/1611.03811.

Organization for Security and Co-operation in Europe. "Media Freedom and Development | OSCE." Accessed January 17, 2017. http://www.osce.org/media-freedom-and-development.

———. "OSCE | Organization for Security and Co-Operation in Europe." Accessed January 17, 2017. http://www.osce.org/whatistheosce.

OSCE Representative on Freedom of the Media. "Propaganda and Freedom of the Media." Non-paper of the OSCE Office of the Representative on Freedom of the Media. Vienna: OSCE, 2015. http://www.osce.org/fom/203926?download=true.

Oyen, Orjar, and Melvin L. De Fleur. "The Spatial Diffusion of an Airborne Leaflet Message." *American Journal of Sociology* 59, no. 2 (1953): 144–49.

Peffer, John. "Paper Bullets: An Interview with Herbert A. Friedman." *Cabinet Magazine*, Fall/Winter 2003. http://www.cabinetmagazine.org/issues/12/pefferFriedman.php.

Pentland, Alex. *Social Physics: How Good Ideas Spread-The Lessons from a New Science*. F First Edition Used edition. New York: Penguin Press, 2014.

Pfanner, Eric. "Citing Internet Standoff, U.S. Rejects International Telecommunications Treaty." *The New York Times*, December 13, 2012. http://www.nytimes.com/2012/12/14/technology/14iht-treaty14.html.

Pop, Valentina. "Nato Colonel Sheds Light on Russia 'Psy-Ops.'" *EU Observer*, January 22, 2015. https://euobserver.com/foreign/127174.

Prosecutor v. Nahimana, No. Case No. ICTR-99-52-T (n.d.).

Puddington, Arch. "Broadcasting Freedom: The Cold War Triumph of Radio Free Europe and Radio Liberty." *Cultural History*, January 1, 2000. http://uknowledge.uky.edu/upk_cultural_history/4.

Qtiesh, Anas. "Spam Bots Flooding Twitter to Drown Info About #Syria Protests [Updated]." *Global Voices Advocacy*, April 18, 2011. https://advox.globalvoices.org/2011/04/18/spam-bots-flooding-twitter-to-drown-info-about-syria-protests/.

Richter, Andrei. "The Relationship between Freedom of Expression and the Ban on Propaganda for War," n.d.

Robertson, Jordan, Michael Riley, and Andrew Willis. "How to Hack an Election." *Bloomberg.com*, March 31, 2016. https://www.bloomberg.com/features/2016-how-to-hack-an-election/.

Samuels, Gabriel. "EU Votes to Fight Back against Russian 'Propaganda Warfare.'" *The Independent*, November 24, 2016. http://www.independent.co.uk/news/world/europe/eu-approves-resolution-to-fight-back-against-russian-propaganda-warfare-a7436036.html.

Shaw, Malcolm N. *International Law*. Cambridge University Press, 2003.

Silverman, Craig. "This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook." *BuzzFeed*, November 16, 2016. https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook.

Smyczek, Peter M. "Regulating the Battlefield of the Future: The Legal Limitations on the Conduct of Psychological Operations (PSYOP) under International Law." *The Air Force Law Review* 57 (2005): 211–40.

"Sovereign - Definition of Sovereign in English | Oxford Dictionaries." *Oxford Dictionaries | English*. Accessed January 17, 2017. https://en.oxforddictionaries.com/definition/sover-

eign.

Sozeri, Efe Kerem. "The Rotten Politics Infecting Turkey's Social Media." *The Daily Dot*, March 30, 2016. http://www.dailydot.com/layer8/turkey-social-media-yeni-safak-facebook-twitter-manipulation/.

Statt, Nick. "WhatsApp Has Grown to 1 Billion Users." *The Verge*, February 1, 2016. http://www.theverge.com/2016/2/1/10889534/whats-app-1-billion-users-facebook-mark-zuckerberg.

Subrahmanian, V. S., Amos Azaria, Skylar Durst, Vadim Kagan, Aram Galstyan, Kristina Lerman, Linhong Zhu, et al. "The DARPA Twitter Bot Challenge." *Computer* 49, no. 6 (June 2016): 38–46. doi:10.1109/MC.2016.183.

Tang, John C., Manuel Cebrian, Nicklaus A. Giacobe, Hyun-Woo Kim, Taemie Kim, and Douglas "Beaker" Wickert. "Reflecting on the DARPA Red Balloon Challenge." *Commun. ACM* 54, no. 4 (April 2011): 78–85. doi:10.1145/1924421.1924441.

Twitter. "Company | About." *Twitter About*. Accessed January 16, 2017. https://about.twitter.com/company.

UN General Assembly. "Definition of Aggression, U.N. General Assembly Resolution 29/3314," December 14, 1974. http://www.un-documents.net/a29r3314.htm.

———. "Developments in the Field of Information and Telecommunications in the Context of International Security." United Nations, August 10, 1999. http://www.un.org/en/ga/search/view_doc.asp?symbol=A/54/213.

UN Human Rights Commission. "Report of the Human Rights Commission." Report of the Human Rights Committee. United Nations, November 24, 2015. https://documents-dds-ny.un.org/doc/UNDOC/GEN/G15/177/83/pdf/G1517783.pdf?OpenElement.

U.S. Department of Defense. "An Assessment of International Legal Issues in Information Operations," May 1999.

Vinyals, Oriol, and Quoc Le. "A Neural Conversational Model." *arXiv:1506.05869 [Cs]*, June 18, 2015. http://arxiv.org/abs/1506.05869.

Wagner, Claudia, Silvia Mitter, Markus Strohmaier, and Christian Körner. *When Social Bots Attack: Modeling Susceptibility of Users in Online Social Networks*, n.d.

Welch, David. *The Third Reich: Politics and Propaganda*. Second. Routledge, 2002. http://psi312.cankaya.edu.tr/uploads/files/Welch,%20Third%20Reich--Politics%20and%20Propaganda,%202nd%20ed.PDF.

Wikipedia. "Gamergate Controversy." *Wikipedia*, January 9, 2017. https://en.wikipedia.org/w/index.php?title=Gamergate_controversy&oldid=759140893.

Zemanek, Karl. "Armed Attack." *Max Planck Encyclopedia of Public International Law*, October 2013. http://opil.ouplaw.com/view/10.1093/law:epil/9780199231690/law-9780199231690-e241.