CYBER DISINFORMATION OPERATIONS (CDOS) AND A NEW PARADIGM OF NON-INTERVENTION

Wenqing Zhao*

ABSTRACT

In recent years, cyber-meddling has risen as a threat to international order, as cyber activities become increasingly weaponized by states for purposes of interfering with domestic policies of foreign states. Historically, such interference could only be obtained through expensive military measures, which could be deterred by the proscription of established international norms. Cyber-meddling presents a unique challenge, as it falls through the cracks of international laws. This paper seeks an international legal solution for one particular form of cyber-meddling that has become a substantial threat to inter-state peace and security: cyber disinformation operations ("CDOs"). In order to tackle CDOs, this paper argues that a new paradigm of non-intervention needs to be formulated, shifting its focus away from the conventional standard of coercion to a new standard of manipulation.

 J.D., Yale Law School, 2020; B.A., College of William & Mary, 2017. While at Yale Law School, Wenqing co-authored Sinotech, Lawfare's biweekly newsblog on U.S.-China technology policy and national security news.

TABLE OF CONTENTS

I. INTRODUCTION

The rise of the Internet in the past century has fundamentally changed not only our daily lives, but also the relationship between different countries. Interstate communications that have significant bearings on inter-state relationships are increasingly conducted and exchanged through various means on cyber-space. New challenges emerge as cyber activities become weaponized by states, aiming to interfere with domestic policies of foreign states, which historically could only be achieved through military measures largely prohibited by international laws.

This paper seeks an international legal solution for one particular form of cyber-meddling, CDOs, that have gradually become a substantial threat to inter-state peace and security. Section II will introduce the basics of cyber meddling and CDOs through a review of current research reports. Section III will demonstrate how existing international laws fail to provide sufficient solutions to address CDOs. Section IV will illustrate the urgency and necessity of international legal solutions tackling CDOs and will propose one potential solution.

II. CYBER MEDDLING IN THE FORM OF DISINFORMATION

A. What is Cyber Meddling and What are Forms of Cyber Meddling?

Cyber meddling is the peacetime use of cyber means with the intent to meddle in the internal policy-makings of other states in non-military ways to pass those states' rights of political independence while avoiding international norms that prohibit foreign interventions. International laws today grant states bedrock "political independence," the autonomy to make political decisions and policies and to handle domestic and foreign affairs without undue external interferences or threats.[1] The concepts of "political independence" and "territorial integrity" have evolved from their earliest manifestations at the end of World War I in the League of Nations[2] as key instruments to maintain international order and fundamental international legal principles, serving as the basis for measuring legitimacy of international behaviors. International laws, however, are products of the international environment. As the concept

---

[1] Samuel K. N. Blay, *Territorial Integrity and Political Independence*, OXFORD PUB. INT'L. L., https://opil.ouplaw.com/view/10.1093/law:epil/9780199231690/law-9780199231690-e1116 (Mar. 2010).

[2] *Id.* ("Territorial integrity refers to the territorial 'oneness' or 'wholeness' of the State. As a norm of international law, it protects the territorial framework of the independent State and is an essential foundation of the sovereignty of States.")

of "political independence" morphs into its shape today, the international environment also undergoes drastic changes. Globalism, for example, has made inter-state interactions more frequent and complex. Consequently, inter-state influence has become more inevitable, blurring the line between legitimate impacts and illegitimate intrusions into a state's political independence.[3]

The development of cyber networks exacerbates the problem, as it further decouples inter-state communications from territorial attachments. While traditional international laws provide certain norms for inter-state territorial actions, especially those of a military nature, they nonetheless fail to account for the many cyber operations in inter-state communications, which are not territorially based and are much less belligerent compared to territorial actions. Actors, therefore, increasingly employ cyber meddling as a newfound device to influence the internal political processes of other states.

Cyber meddling operations can be multifaceted and can include, among other things: (1) hacking into federal, state, and local policy-making institutions to acquire and exploit classified information; (2) hacking and leaking selective information to influence domestic policies; (3) massive and organized operations of disinformation and trolling through numerous media outlets to shape public opinions, sow conflict and distrust, and effectuate changes in policies; and (4) hacking and tampering the information technology systems of political institutions to obstruct or influence political processes, like altering election ballots.

While cyber means can also be deployed in a more military fashion, such as through hacking and undermining the central military operations and facilitation systems of adversaries,[4] cyber meddling is more covert and less aggressive, aiming at audiences, institutions, and capacities within states. The objectives of cyber meddling frequently converge with those of military cyber actions. By manipulating policy-making processes, cyber meddling likewise pushes the target state to change its policies. However, due to its more discreet and amorphous nature, cyber meddling manages to exist in the twilight zone of international law when more belligerent cyber operations are captured under cyber warfare[5] and the prohibition of use of force. This gap of norm erodes not only the principle of political independence, but also inter-state

---

[3] W. Michael Reisman, *Meddling in Internal Affairs: Establishing the Boundaries of Non-Intervention in a World without Boundaries, in* RESOLVING CONFLICTS IN THE LAW 98, 99 (Chiara Giorgetti & Natalie Klein eds., 2019) ("There is no such thing as the international arena; because there are only other States, to interact with them is, to an inescapable extent, to interfere in them.").

[4] Johann-Christoph Woltag, *Cyber Warfare*, OXFORD PUB. INT'L. L., https://opil.ouplaw.com/view/10.1093/law:epil/9780199231690/law-9780199231690-e280?rskey=s7XhqZ&result=1&prd=OPIL (Aug. 2015).

[5] *Id.*

trusts based on the principle of non-intervention, rendering it inevitable that more conflicts, greater uses of other grey-zone countermeasures, and diminutions of the competence of international law will ensue.

B. Cases of Cyber Disinformation Operations (CDOs)

One particularly troublesome form of cyber-meddling is a CDO, a cyber operation from foreign state actors or affiliates, or those under the control or instruction of state actors and affiliates. CDOs involve massive fabrication of information or creation of identities falsely claimed to be affiliated to citizens or institutions of other states and disseminations of those disinformation to the audience in other states. CDOs are often times part of a larger influence campaign—political propaganda reinvented through the power of disruptive technology and that of social media, such as deepfakes, manipulative algorithms, automation, and big data. Those who aim to influence campaigns hope to sway the general public of another state in order to shape its political decisions.

CDOs take on different forms and usually have at least one of the following components: (1) productions of outright fake digital contents; (2) creation and automation of trolls pretending to be citizens or entities of the targeted state; (3) research and manipulation of social media's "trending" and recommendation algorithms to amplify disinformation, customize audience bases, and flood digital platforms; or (4) deliberate seeding of disinformation in partial truths.

1. Russia's Operation in the U.S. 2016 Presidential Election and 2018 Midterm Election

A notorious example of a CDO is Russia's cyber disinformation campaign for both the 2016 U.S. Presidential election and in the 2018 U.S. midterm election. According to the U.S. Department of Justice's Report on Russian Interference in the 2016 Presidential Election ("Mueller Report"),[6] Russia deployed an "Active Measures" Social Media Campaign, a type of operation usually conducted by Russian security services, aiming to influence international affairs.[7] The Internet Research Agency LLC ("IRA"), a Russian cyber troll factory, generated massive social media accounts by embodying fake U.S. personas, operated those accounts while pretending to be U.S. activists, and promoted the dissemination of political disinformation

---

[6] U.S. DEP'T OF JUST., REPORT ON THE INVESTIGATION INTO RUSSIAN INTERFERENCE IN THE 2016 PRESIDENTIAL ELECTION (VOLUME I OF II) (2019), https://www.justice.gov/storage/report.pdf.

[7] *Id.* at 14.

disparaging candidate Hillary Clinton in favor of candidate Donald Trump over social media.[8] The IRA bought over 3,500 Facebook political advertisements[9] in excess of two million dollars[10] and fabricated public social media pages. The IRA also accrued significant influence on various social media outlets, controlling at least 470 Facebook accounts that reached as many as 126 million people and at least 3,814 Twitter accounts that reached approximately 1.4 million people.[11] In conducting the campaign, the IRA utilized amplification algorithms of many social media platforms, and mass-produced bots to artificially inflate the influence of its fake accounts and disinformation, inserting content into other users' suggested feeds or trending topics even when those users did not subscribe to the IRA's troll pages.[12]

Russia's CDO took place again in the 2018 U.S. midterm election, renamed as Project Lakhta.[13] Project Lakhta involved extensive creations of bots on social media platforms that generated posts touching on controversial political topics. While the 2016 interference was aimed at influencing voters' preference for one candidate over another, the 2018 operation did not pick out any one particular candidate. Rather, the operation sought to exacerbate the existing social and political tensions in the U.S. by writing and promoting diverging viewpoints on the same contentious issue. As the Mueller Report and FBI affidavits on midterm interference put it, the strategic goal of Russia's cyber interference campaign was not only to achieve a particular policy change, but also to sow "division and discord" in the U.S. society and politics.[14] Moreover, the operation also used fake images.

These operations of cyber disinformation and trolling are hardly just incidents between Russia and the U.S. Organized inter-state disinformation campaigns asserting foreign influence are on the rise because these new political weapons are cheaper to conduct, harder to detect, and much harder to inculpate under current international laws. Russia, for instance, has targeted not only the U.S., but also many other states in its cyber disinformation incursions, such as with efforts against Sweden and France confirmed.[15] The

---

[8] *Id.* at 25.
[9] *Id.*
[10] Michael N. Schmitt, *"Virtual" Disenfranchisement: Cyber Election Meddling in the Grey Zones of International Law*, 19 CHI. J. INT'L. L. 30, 35 (2018).
[11] U.S. DEP'T OF JUST., *supra* note 6, at 15.
[12] *Id.* at 26.
[13] Adam Goldman, *Justice Dept. Accuses Russians of Interfering in Midterm Elections,* N.Y. TIMES (Oct. 19, 2018), https://www.nytimes.com/2018/10/19/us/politics/russia-interference-midterm-elections.html.
[14] *Id.*
[15] Margaret L. Taylor, *Combating Disinformation and Foreign Interference in Democracies: Lessons from Europe,* BROOKINGS INST. (July 31, 2019), https://www.brookings.edu/blog/techtank/2019/07/31/combating-disinformation-and-foreign-interference-in-democracies-lessons-from-europe/.

IRA has, among other things, produced more than 115,000 tweets in German, more than 42,000 tweets in Arabic, and more than 18,000 tweets in Italian, an indication of a global propaganda operation.[16] A report published by Princeton researchers also shows that more than twenty countries have fallen into being victims of fifty-three foreign influence efforts from other countries from 2013 through 2018.[17] CDOs are no doubt global, as demonstrated by the following research reports.

2. The Oxford Report: The Global Disinformation Order

Oxford published a report on the global phenomena of organized social media manipulation ("The Oxford Report").[18] The report monitored efforts from governments and political parties to conduct organized social media manipulations and disinformation campaigns from 2016 to 2019, including those campaigns aimed at foreign interference. The Oxford Report examined cyber troop activities in seventy countries, among which Facebook and Twitter had identified at least seven countries that had used the platforms for foreign influence operations.[19] Because of the difficulties in attributing cyber activities to state actors, this list is by no means conclusive. The remaining sixty-three countries, while not yet identified by Facebook and Twitter as involved in foreign influence operations, have engaged in different forms of organized social media manipulation campaigns and tested different strategies and techniques of conducting computational propaganda.[20]

The Oxford Report also highlighted the extensive scope, scale, and precision of CDOs and the danger of networking technologies once weaponized. CDOs are affordable compared to many other aggressive political strategies. A country can automate bots to "amplify narratives or drown out political dissent;"[21] bot accounts have been identified and used in fifty countries. Even human-run fake accounts are inexpensive to experiment with, such that eighty-seven percent of the countries examined adopted the

---

[16] Tim Mak, *Troll Factory Contributes to Russia's Worldwide Interference,* NPR (Dec. 12, 2018, 5:18 AM), https://www.npr.org/2018/12/12/675987838/russias-worldwide-interference.

[17] Diego A. Martin & Jacob N. Shapiro, *Trends in Online Foreign Influence Efforts,* ESOC PUBL'NS (July 8, 2019), https://scholar.princeton.edu/sites/default/files/jns/files/trends_in_foreign_influence_efforts_2019jul08_0.pdf.

[18] Samantha Bradshaw & Philip N. Howard, *The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation,* COMPUTATIONAL PROPAGANDA RSCH. PROJECT (2019), https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf.

[19] *Id.* at 2.

[20] *Id.* at 6.

[21] *Id.* at 11.

strategy.[22] Skills and knowledge of CDOs also diffuse across geographic lines, as several states sent their propaganda officials to more sophisticated states for disinformation training. The affordability and accessibility of CDOs make it appealing for countries to adopt or test their strategies and subsequently expand the scope and scale of the usage and the precision of CDOs. So far, twelve countries have acquired advanced capacity to conduct CDOs; while twenty-six countries have acquired medium capacity with consistent CDO strategies and full-time staff members dedicated to CDOs.[23]

Although a lot of countries investigated by the Oxford Report so far have not been found to engage in inter-state cyber meddling, the risk of them doing so cannot be neglected. The number of countries engaging in CDOs, domestic or foreign, has increased from twenty-eight in 2017, to forty-eight in 2018, to seventy in 2019.[24] This rapid expansion signifies how quickly countries are picking up on the tool of cyber disinformation and makes one wonder how soon it will take for them to extend the application from domestic context to foreign interference. The methods and objectives of domestic CDOs are also strikingly similar to those of inter-state cyber disinformation meddling. More than sixty-eight percent of the countries investigated used disinformation and media manipulation to mislead audiences, trolling and fake accounts to defraud users, and bots and flooding hashtags to amplify messages and content.[25] Equipped with capacities of CDOs and motivated by national interests, countries are likely to get their hands-on inter-state meddling in the near future. The methodological limitations of the Oxford Report (media bias and language)[26] also substantiate the concern that the report might be under-inclusive of the number of countries and the extent of their effort engaging in foreign influence.

3. The Nemr Report: Foreign State-Sponsored Disinformation in the Digital Age

Separately, a 2019 report conducted by Christina Nemr and William Gangware ("The Nemr Report") also examined the appeal of CDOs as an efficient weapon for states to use when interfering with the affairs of other states.[27] Operations to spread fake news are very effective as a false story, on

---

[22] *Id.*
[23] *Id.* at 18-19.
[24] *Id.* at 2.
[25] *Id.* at 15.
[26] *Id.* at 8.
[27] CHRISTINA NEMR & WILLIAM GANGWARE, WEAPONS OF MASS DISTRACTION: FOREIGN STATE-SPONSORED DISINFORMATION IN THE DIGITAL AGE, (Rhonda Shore & Ryan Jacobs eds., 2019), https://www.state.gov/wp-content/uploads/2019/05/Weapons-of-Mass-Distraction-Foreign-State-Sponsored-Disinformation-in-the-Digital-Age.pdf.

average, "reaches 1,500 people six times more quickly than a factual story."[28] Gaming platform algorithms and big data disinformation operations can also customize target audiences and tailor content according to the behavioral patterns of users, making the dissemination of disinformation by bots and trolls more efficient.

The Nemr Report provided more data than the redacted Mueller Report in its overview of foreign disinformation campaigns, illustrating the adeptness and flexibility of different disinformation tricks. Russia's IRA, for instance, did not stop at Facebook and Twitter. Other social media platforms utilized included "Instagram, YouTube, Google+, Vine, Meetup, Pinterest, Tumblr, Gab, Medium, Reddit."[29] Leveraging different features of various platforms, the IRA was able to maximize the reach of its disinformation and audience engagement. For instance, features of Instagram were studied by the IRA, which eventually landed the IRA 187 million engagements (likes and shares).[30]

The Report also cited several other occasions of sovereign states engaging in cyber disinformation campaigns, including operations from China and Iran. In the case of Iranian cyber propaganda, fake news was created and disseminated, among which some had attracted high-level real-life responses. In late 2016, the website AWDnews, leveraged by the Iranian government, published a fake piece claiming that the Israeli government had threatened a nuclear attack if Pakistan sent troops to Syria. Mistaking the fake news as authentic, Pakistan's then-Defense Minister "responded with an actual nuclear threat against Israel."[31]

Lastly, the Nemr Report highlighted the insufficiency of a technology-only approach to combat state-sponsored CDOs and warned readers of the rapid advancement of AI technology that could be misused to further CDOs. The report sensed a significant gap in technology,[32] where the AI-powered altered photos and videos, fake news, and bots were becoming more sophisticated, easier to produce yet harder to detect. Disruptive technology like deepfakes had opened up the Pandora's box for CDOs. It would only be a matter of time before disinformation content begins to migrate "from being largely static (fake articles) to dynamic (video and audio)"[33] to drastically enhancing the credibility of fake news, as candidates or key personnel could now be portrayed by disinformation operations as saying or doing things that they never said or did. The fake detection technology, on the other hand, is left crippled behind, and the gap between

---

[28] *Id.* at 3.
[29] *Id.* at 17.
[30] *Id.*
[31] *Id.* at 24.
[32] *Id.* at 39-41.
[33] *Id.* at 40.

deepfake technology and fake detection technology is expected to widen in the near future. A Gartner report predicted that by 2020, "the abilities of AI to generate counterfeit media will surpass those of AI to identify such media."[34] It indicated that the pinning of all hopes on private social media platforms to detect, monitor, and deter disinformation was also overly-optimistic, since serious structural challenges like the volume of content and encryption of messages existed and impeded social media's capacities to cabin state-sponsored disinformation operations.

The Nemr Report suggested that the ultimate responsibility for countering disinformation should fall on governments, who were the targets of geopolitical adversaries. However, the report believes that the battleground "rests firmly in private hands,"[35] that governments need to more actively regulate social media, and that social media needs to discipline its content. Realizing how many hurdles any mechanism of public regulations and private self-regulations would have to jump through, the report was only able to vaguely suggest a "greater collaboration between technology companies and governments" in the "absence of clear delineations of responsibility."[36]These collaborations would mainly involve information-sharing and fact-finding, though the efficiency and sufficiency of which are highly questionable in the face of technological gaps and structural challenges.

4. The NYU Report: Disinformation and the 2020 Election

The New York University ("NYU") Stern School of Business also published a report on disinformation operations ("The NYU Report"),[37] projecting disinformation development in the coming years, particularly in the 2020 U.S. Presidential election. The report predicted several main trends of foreign disinformation operations. First, more AI-backed synthetic videos would be used for political disinformation by foreign adversaries of the U.S. and would likely attract significantly wider circulation and assert greater influence on viewers.[38] Second, foreign disinformation operations would start to move their battlegrounds from traditional platforms like Facebook and Twitter, to more private messaging and sharing platforms like Instagram and WhatsApp, adopting new strategies like digital voter suppression (i.e.,

---

[34] *Id.* at 41.
[35] *Id.* at 37.
[36] *Id.*
[37] PAUL M. BARRETT, NYU STERN CTR. FOR BUS. AND HUM. RTS., DISINFORMATION AND THE 2020 ELECTION: HOW THE SOCIAL MEDIA INDUSTRY SHOULD PREPARE (2019), https://issuu.com/nyusterncenterforbusinessandhumanri/docs/nyu_election_2020_report?fr=sY2QzYzI0MjMwMA.
[38] *Id.* at 3.

bullying or confusing voters to refrain from voting).[39] Third, disinformation operations have incubated a whole new industry of election-manipulation services that could come to the aid of states interested in foreign interference. Some companies had even blatantly promoted themselves online as taking "every advantage available in order to change reality according to [their] client's wishes."[40] This phenomenon not only implies an acceleration of collaborations between actors across borderlines, but also means more accessibility of external disinformation capacities to states with currently low CDO capacities. States interested in cyber meddling can now outsource their schemes to professional "social media research firms" to do the tricks.[41] The NYU Report also re-emphasized the importance of combatting disinformation campaigns, as the damages were detrimental: erosion of democratic values, heightened cynicism, and exacerbation of political polarization.

The NYU Report, like the Nemr Report, ended on the note of urging more efforts from social media platforms. It, however, acknowledged the ever-changing permutations of disinformation operations, which greatly increase the challenges for social media self-regulation.[42]

5. From the U.S. Defense of 2020 Election to Global Legal Order Against CDOs

All reports surveyed above fail to recognize that an international challenge of CDOs needs an international approach. These reports do not comment on the wrongfulness of inter-state CDOs and no international solutions are discussed, in spite of the global nature of foreign disinformation operations. In contrast, one of the EU's key 2019 studies on disinformation operations[43] recognized that international relations would be increasingly affected by inter-state disinformation operations, and that international law should be developed to provide more guidance in the future, deterring professionally designed, built, and financed foreign media manipulations. As the study put, "states must not endure criminals systematically misusing intermediaries' services and their infrastructure […] Such actions should be

---

[39] *Id.* at 12.
[40] *Id.*
[41] *Id.* at 11.
[42] *Id.* at 17.
[43] JUDIT BAYER ET.AL., POLICY DEP'T FOR CITIZENS' RIGHTS AND CONSTITUTIONAL AFFAIRS, DIRECTORATE GENERAL FOR INTERNAL POLICIES OF THE UNION, DISINFORMATION AND PROPAGANDA – IMPACT ON THE FUNCTIONING OF THE RULE OF LAW IN THE EU AND ITS MEMBER STATES 125, 135 (2019), https://www.europarl.europa.eu/RegData/etudes/STUD/2019/608864/IPOL_STU(2019)608864_EN.pdf.

fought with the ultimate tools of the law […] international law should be taken into consideration."[44]

The battleground should not be, as the Nemr Report suggests, only in the private hands of social media platforms. Rather, a key component of the solution to this global challenge is the establishment of a new international norm regarding the (il)legality of CDOs. Digital networks have fundamentally changed the landscape of inter-state communications and have been strategically exploited by states to circumvent traditional international norms of non-interventions on an unprecedented scale and in a frightening speed. Without legal restraint, such misuses of disinformation are bound to proliferate in the future, spurred by fast-developing technologies. The Nemr Report is correct about one thing: the global community is in urgent need of "clear delineations of responsibility."[45] International law is duty-bound.

III. CURRENT INTERNATIONAL LAWS DO NOT PROVIDE SUFFICIENT SOLUTIONS TO ADDRESS CDOS – CDOS FALL IN THE GREY ZONE OF INTERNATIONAL LAW

How CDOs fit into the current international legal landscape governing inter-state communications is not a straightforward question. Russia, the most (in)famous power behind CDOs, depicts it as an "Information War," a defensive instead of offensive tactic against threats from external information activity initiated by Western media conglomerates that promote "the geopolitical agenda of the U.S. and its allies at Russia's expense."[46] Defining the information war as "a struggle between two or more states … to destabilize a society and a state through massive psychological conditioning of the population, and also to pressure a state to make decisions that are in the interest of the opponent,"[47] the Russian Ministry of Defense echoes one popular view that CDOs are a form of psychological operations ("PsyOps"), though by categorizing the operation as part of a "war," Russia also suggests that CDOs have a strong military overtone.

Regardless, neither international principles regulating the lawfulness of military measures nor rules about PsyOps can successfully tackle CDOs, as CDOs do not fall squarely into either category. The customary international norm of non-intervention comes the closest, yet the current interpretations and applications of the norm also do not cover CDOs. Another controversial

---

[44] *Id.* at 139.
[45] NEMR & GANGWARE, *supra* note 27, at 37.
[46] Russia's Information War – Propaganda or Counter-Propaganda?, EUR. PARL. DOC. PE 589.810 (2016), http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/589810/EPRS_BRI(2016)589810_EN.pdf.
[47] *Id.*

customary international principle, sovereignty, likewise does not help much. Currently, there are no treaty provisions that specifically take on international behaviors in cyberspace, in spite of increasing global awareness of the need to establish norms. However, academia has contributed some analysis in the Tallinn Manuals on the International Law Applicable to Cyber Warfare 2.0, a non-binding study commissioned by the NATO Cooperative Cyber Defense Center of Excellence and produced by an international group of experts to address how to interpret international laws and apply them to cyber operations. The Tallinn Manuals also did not reach a conclusion regarding the legality of CDOs. To conclude, CDOs live and thrive in the grey zone of international law.

A. Victim States of CDOs Have No Legal Recourse in the United Nations Charter Article 2(4)

The lawfulness of resorting to military measures, the jus ad bellum, is regulated by the Charter of United Nations Article 2(4). Article 2(4) provides that "all members shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state."[48] No definitions are provided in the Charter, however, for the key terms, "use of force" and "political independence," except for a mentioning in the preamble that one purpose of the Charter is to "ensure…that armed force shall not be used, save in the common interest."[49] It is altogether unclear if "use of force" in Article 2(4) implies a use of armed force, and it would be imprudent to draw such a definite conclusion as arguments were made that use of force is "a large category of activities containing a smaller subset of events that qualify as armed subset."[50]

CDOs, while capable of asserting broad influence on the target states, can hardly fall under the category of armed force. Furthermore, a CDO rarely amount to a "use of force," even when the term arguably covers operations less grave than armed force. Regardless of the scope of use of force, incursions into a nation's political independence in the form of CDOs have an "ostensibly peaceful" façade that is non-belligerent in nature, and therefore are "seldom broadly accepted as uses of force."[51] The worry of counting CDOs as "use of force," other than it is counter-intuitive to the plain language, is that the effect of CDOs might be disproportionate to that of belligerent military actions (the

---

[48] U.N. Charter art. 2, ¶ 4.
[49] U.N. Charter pmbl.
[50] Christopher H. Kinslow, *Game of Code: The Use of Force Against Political Independence in the Cyber Age*, ARMY LAW., July-Aug. 2018, at 29, 30, https://www.loc.gov/law/mlr/pdf/07-08-2018.pdf.
[51] *Id.* at 31.

traditional understanding of "use of force"). Consequently, counting CDOs as "use of force" might trigger a state's countermeasures disproportionate to what CDOs deserve, in forms of draconian economic sanctions or even military actions, especially when the effect of CDOs is extremely hard to measure as I will discuss in later sections.

As if to address the proportionality concern, the U.S. has adopted an "effects-based test," where cyber activities would count as use of force when causing "direct physical injury and property damage." [52] While cyber operations, like hacking that causes nuclear explosions, would clearly amount to a use of force, such a test makes nuanced cyber operations like CDOs even harder to inculpate, as the harms done by CDOs (i.e., erosion of democratic values, heightened cynicism, and exacerbation of political polarization) are all indirect and difficult to measure, though often times no less severe.

Scholars have attempted to update the use of force paradigm to account for emerging cyber operations like CDOs. However, Article 2(4) seems ill-suited as a solution to CDOs. The "use of force" in Article 2(4) has long been read as containing physical violence or resulting in physical harm. This is distinct from other forms of coercion that generally fall under the customary international norm of non-intervention; such interpretation also represents the longstanding position of the U.S.[53] There were also significant historical debates over whether or not to clearly define the term "use of force." One such contention took place in the preparation of UN General Assembly Resolution 2625, the Declaration on Principles of International Law Friendly Relations and Co-operation among States adopted in 1970.[54] The Report of the Special Committee on Resolution 2625[55] unfolded the debate over the definition of "force" in detail. Multiple states had jointly proposed to insert a broad definition of the term "force" in Resolution 2625, such that "force" would include not only military force, but also "all forms of pressure, including those of a political and economic character, which have the effect of threatening the territorial integrity or political independence of any state."[56] Some suggested that forms of pressure exercised on ideological and religious grounds should thus be prohibited.[57]

However, no agreement was reached by members to either include a definition of "force" in the Resolution 2625 statement or to expand "use of

---

[52] *Id* at 32.
[53] *Id.* at 33.
[54] G.A. Res. 2625 (XXV), at 121 (Oct. 24, 1970).
[55] U.N. Special Comm. on Principles of Int'l Law Concerning Friendly Rel. and Co-Operation Among States, *Report of the Special Comm. on Principles of Int'l Law Concerning Friendly Rel. and Co-Operation Among States,* U.N. Doc. A/7326 (1968), https://digitallibrary.un.org/record/856001.
[56] *Id.* at 14.
[57] *Id.* at 24.

force" to cover any form of pressure against the political independence or territorial integrity of a state. Some representatives would only understand the term "force" to mean exclusively "armed force."[58] Consequently, no definition of "force" was provided in the Declaration on Friendly Relations. The standard for non-belligerent inter-state actions was instead provided in the principle of non-intervention, a principle that I will further discuss.

In any case, in spite of the fact that ambiguity remains in whether pressure exercised through cyber means could amount to "use of force," it is unlikely that General Assembly members would agree on elevating cyber meddling to the status of "use of force." Article 2(4)'s prohibition against "use of force" has so far been a standard of physical violence, not contemplating actions that are short of military natures. To re-propose the expanded definition of "use of force" to cover CDOs is currently unattainable, given that the extended definition would result in significant and unwanted ramifications on the legality of other forms of economic and political pressures, and on the rights of self-defense and countermeasures.

While historically egregious inter-state actions almost certainly would involve physical violence and result in physical damages in order to achieve certain objectives, time has changed a state's available inventory to accomplish the same goals non-belligerently. Old concepts no longer work, and new rules need to be developed. As Chris Kinslow, an author who wrote on cyber operations and Article 2(4) put, instead of "continuing to finesse the armed attack standard into greater feats of contortion, the legally responsible course of action is to admit that the world has indeed changed…"[59] The question therefore becomes this: in a world where cyber capacities have fundamentally changed the means to achieve political ends, what new concepts and rules need to be brought into place and what international law-making modality could facilitate the creation and adoption of those concepts and rules?

In this process, existing international legal norms based on older concepts could inform us of the gaps in the law, instruct us on the feasibility of creating new concepts, and lend us power to model the creation and adoption of new rules on the successful process of old ones. While Article 2(4) and the "legislative history" of the term "use of force" show us that it is a stretch to stuff CDOs under the umbrella of "use of force," the norm of non-intervention would seem facially closer as a viable solution to CDOs, though it eventually turns out to be insufficient as well.

---

[58] *Id.*

[59] Kinslow, *supra* note 50, at 33.

B. Customary International Laws Like the Norm of Non-Intervention and the Principle of Sovereignty, Under Their Current Interpretations, Also Do Not Cover CDOs

1. The Principle of Non-Intervention Covers Operations That Do Not Amount to the "Use of Force," but Traditionally Wrongful Acts Under the Principle of Non-Intervention Requires an Element of "Coercion"

Current customary international law provides the principle of non-intervention, which is the duty on states to not intervene, directly or indirectly, in the internal or external affairs of any other state. The duty of non-intervention, being jus cogens, is "one of the fundamental duties of the State," [60] developed and codified in multiple UN General Assembly resolutions and affirmed as "part and parcel of customary international law" by the International Court of Justice on the merits of Nicaragua v. United States of America.[61] Though not explicitly mentioned, the non-intervention principle is rooted in the UN Charter, as relevant UN resolutions refer to the Charter as the foundation for such a principle. Many argue that the principle of non-intervention is drawn from the principle of state sovereignty in Article 2(1).[62] The obligation not to intervene is said, according to Lassa Oppenheim, to be "the corollary of every State's right to sovereignty, territorial integrity, and political independence." [63] Others believe that the principle of non-intervention is derived and expanded from Article 2(4)'s "use of force" provision, as force is one important form of prohibited interventions.[64]

Surveying the history of the principle of non-intervention, one can indeed see two interwoven trends: (a) the evolvement from no non-intervention principle at all, to the one-sided outlawing of foreign intervention, to finally a global recognition of the principle as the notion of Article 2(1)'s "sovereignty" matured;[65] and (b) the expansion of the principle from restricting mainly inter-state military actions mirroring the Article 2(4) "use of force," to prohibiting other forms of interferences. On one hand, state borders around the globe were gradually cemented through and post two World Wars, unoccupied land claimed and new nations established, a

---

[60] Philip Kunig, *Intervention, Prohibition of*, *in* MAX PLANCK ENCYC. OF PUB. INT'L LAW, https://spacelaw.univie.ac.at/fileadmin/user_upload/p_spacelaw/Kunig_Intervention_Prohibition_of.pdf.
[61] *Id.*
[62] Kinslow, *supra* note 50, at 48.
[63] Sir Michael Wood, *Non-Intervention (Non-interference in domestic affairs)*, *in* ENCYCLOPEDIA PRINCETONIENSIS, https://pesd.princeton.edu/node/551.
[64] Kunig, *supra* note 60, at 4.
[65] *Id*.

territorial norm that preserves the notion that existing state borders and sovereignties needs to be in place to maintain "international peace and security."[66] The norm of non-intervention fulfills such purpose and is essential to ensure that "nations live together in peace with one another;"[67] outlawing inter-state forcible means, it guarantees a certain level of stability and amity among states. On the other hand, flourishing globalism, with its increasing interstate exchange of ideas, necessitates a norm of communication that, on top of forbidding explicit use of force, also draws a line between permissible interaction and impermissible interference (those impressible interferences are often times achieved through economic, political, and diplomatic means that does not necessarily amount to threats or use of force). The broader understanding of intervention was motivated by a historical "increasing co-operation among nations which made it possible to interfere much more subtly and effectively without the use of force."[68]

Such an expanded notion of intervention was first introduced in UN General Assembly Resolution 2131.[69] The resolution in its preamble explicitly mentioned the General Assembly's concern of "direct or indirect forms of interference" threatening the sovereignty and political independence of States and declared that all forms of indirect intervention "constitute a violation of the Charter of the United Nations." Apart from officially finding a category of non-violent interference measures distinct from prohibited armed interventions, but equally condemnable under the principle of non-intervention, the Resolution also brought the notion of "duress or coercion" into the non-intervention legal landscape.

While the Resolution imposed a general prohibition on "all forms of interference" against the personality of another state in Article 1 and acknowledged in Article 5 that "every state has an inalienable right to choose its political, economic, social and cultural systems, without interference in any form by another State," it seemed to narrow the scope of impermissible intervention to those measures that "coerce another State" to exercise its sovereign rights subordinated in Article 2.[70] Does the more specific provision in Article 2 requiring an element of coercion serve as an inherent supplementary interpretation of what counts as "interference" in Article 1 and 5? Or do Article 1 and 5 have stand-alone force to encompass measures not facially coercive? Nowhere in the resolution were intervention or interference defined; thus, the ambiguity is hard to avoid. It nevertheless preserves the possibility that the non-intervention principle could cover more than coercive

---

[66] G.A. Res. 2625, *supra* note 54, at 2.
[67] *Id.*
[68] Kunig, *supra* note 60, at 3.
[69] G.A. Res. 2131 (XX), at 1 (Dec. 23, 1965).
[70] *Id.* at 3.

measures suggested by Article 2, as otherwise Article 1 and Article 5 would be redundant.

Edward McWhinney, an international legal scholar, thinks the ambiguities of Resolution 2131 as ways to accommodate contradictory imperatives between different international players, so to ensure the eventual passage and the adoption of the Resolution. Resolutions of diverging interests lie in the "arts of diplomatic-legal drafting," the intentionally succinct and vague languages that "facilitate normative ambiguity as to their later interpretation or concrete application."[71] The application of the general principle of non-intervention to non-military interventions in concrete cases is therefore uncertain, preserving rooms by design for future "give-and-take."[72] This legislative history mirrors that of Resolution 2625, where members reached a consensus on the high-level abstract principle without detailing controversial definitions or secondary rules, thereby preserving room for more specified standards based on consents in legal instruments like treaties.

Resolution 2625—the Friendly Relations Declaration—recycled most of the non-intervention language in Resolution 2131, though in a slightly stronger tone as it states that "armed intervention and all other forms of interference… against the personality of the State" are violations of international law.[73] Like in the case of defining "use of force," debates arose during the legislative process as whether to expand the scope of consensus regarding the non-intervention principle reached in Resolution 2131. The Special Committee report showed that the drafting committee was tasked with "consideration of addition proposals with the aim of widening the area of agreement of General Assembly Resolution 2131,"[74] though eventually no agreement was reached on any additional proposals. Some delegates expressed their wish to clarify the forms of pressure and the definition of "non-intervention," but the line drawing was extremely hard as General Assembly members struggled with notions of interference, intervention, coercion, and use of force. Whether to subject one notion under the category of another, to provide separate standards for different levels of aggressions, or to remain silent, were important policy decisions to make, particularly when the norm of non-intervention conflicted with some other established international norms (for instance, when interventions happened in the context of supporting internationally recognized objectives like self-determination or protections of human rights).

---

[71] *Id.* at 2.
[72] *Id.*
[73] G.A. Res. 2625 (XXV), *supra* note 54, at 123.
[74] U.N. Special Committee on Principles of International Law Concerning Friendly Relations and Co-operation Among States, *supra* note 55, at 4.

Resolution 2625, like Resolution 2131, left key terms of intervention undefined, an issue that was later taken up in UN General Assembly ("UNGA") Resolution 36/103 (The Declaration on the Inadmissibility of Intervention and Interference in the Internal Affairs of States), [75] an unsuccessful attempt to expand the scope of intervention. This new resolution declared that "no state or group of states has the right to intervene or interfere in any form or for any reason whatsoever in the internal and external affairs of other States,"[76] asserting not only a duty of non-intervention, but also a duty of non-interference. It then proceeded to an enumerated list comprised of states' rights and duties, among which the most relevant one was the duty to "abstain from any hostile propaganda for the purpose of intervening or interfering in the internal affairs of other States." [77] This Declaration, unbinding as the former two, did not gain wide recognition, as it was passed against the will of many member states and "[did] not reflect general international opinion on the topic."[78]

While the UNGA has not provided further clarification on the definition of intervention yet, an influential case from the International Court of Justice, Nicaragua v. United States of America,[79] put forward a standard of coercion as the definition of intervention that is generally accepted as international consensus. An intervention is considered as prohibited only if it is:[80]

> one bearing on matters in which each State is permitted, by the principle of State sovereignty, to decide freely... Intervention is wrongful when it uses methods of coercion in regard to such choices, which must remain free ones. The element of coercion […] defines […] and indeed forms the very essence of prohibited intervention […]

For many, this two-element definitional test adequately restates the current customary international law of non-intervention and is recognized by the Tallinn Manual in that an operation is wrongful intervention only if it affects a State's domaine reserve and is coercive,[81] though what is a State's domaine reserve and what counts as coercive measures are still up to substantial debate. It is also worth noticing that the ICJ cited Resolution 2625 throughout the decision as evidence of an established customary international norm of non-intervention, but when it introduced the element of coercion to

---

[75] G.A. Res. 36/103, at 78 (Dec. 09, 1981).
[76] *Id.*
[77] *Id.*
[78] Kunig, *supra* note 60, at 6.
[79] Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), Judgment, 1986 I.C.J. Rep. 14, ¶ 205 (June 27).
[80] *Id.* ¶ 205.
[81] Kunig, *supra* note 60, at 2.

the norm, Resolution 2625 was not referred to. Rather, the court appealed to "the generally accepted formulations" of non-intervention to conclude that coercion must be present for an intervention to be wrongful. Consequently, the coercion standard was later argued by some as not rooted in the principle of non-intervention but created by the ICJ only to narrowly address the facts and issues in Nicaragua.

2. CDOs are Not Currently Considered Coercive

CDOs are not currently considered coercive, and therefore are unlikely to be covered by the customary international law of non-intervention. CDOs with goals to interfere with domestic policy-making or elections of another state would usually pass the first element of the non-intervention test, as policy-making and elections are considered as a state's domaine reserve – areas of state activities that "are internal or domestic affairs of a State and are therefore within its domestic jurisdiction or competence."[82] However, the second element of the test—coercion—presents a significant problem for CDOs.

To start with, experts in the Tallinn Manual 2.0 specifically made a distinction between lawful cyber interference and unlawful cyber intervention because the former is not coercive. The Tallinn Manual 2.0 observed that the term "coercion" was not defined in international law, but read it as "an affirmative act designed to deprive another State of its freedom of choice […] to force that State to act in an involuntary manner or involuntarily refrain from acting in a particular way."[83] Though coercive actions need not to be physical or result in physical damages, they must not only seek for a change of conduct, but also effectuate the change of conduct that "deprives the State of control over the matter in question."[84]

The Tallinn Manual 2.0 also differentiated coercion from "persuasion, criticism, public diplomacy, propaganda… [that] merely involve[s] […] influencing the voluntary actions of the target State."[85] As an example, a state-sponsored public information cyber campaign "designed to persuade another State" would not rise to a violation of prohibited intervention.[86] Michael Schmitt, the general editor of the Tallinn Manual 2.0, in his article, Virtual Disenfranchisement: Cyber Election Meddling in the Grey Zones of

---

[82] Katja S. Ziegler, *Domaine Reserve, in* MAX PLANCK ENCYC. OF PUB. INT'L LAW, https://opil.ouplaw.com/view/10.1093/law:epil/9780199231690/law-9780199231690-e1398?rskey=Lnml6f&result=1&prd=OPIL.

[83] MICHAEL N. SCHMITT & LIIS VIHUL, TALLINN MANUAL 2.0 ON THE INT'L L. APPLICABLE TO CYBER OPERATIONS 749 (Cambridge Univ. Press 2017).

[84] *Id.* ¶ 20.

[85] *Id.* ¶ 21.

[86] *Id.*

International Law,[87] further discusses where cyber disinformation campaigns fall on the intervention spectrum. Acknowledging that cyber disinformation campaigns are more than persuasion, influence, or propaganda because of their deceptive nature, Schmitt nevertheless suggests that the coercion standard as it currently is accepted can barely catch CDOs, because the causation between disinformation efforts and the results is hard to prove.[88] Coercion might be manifested in the "subordination of sovereign will" of a state, in the sense that the election result would not have otherwise changed but for the CDO efforts. Establishing such a causal link is, however, extremely difficult, even if one seeks only for indirect causations. A catch-22 emerges: more indirect causations would "move the activity along the continuum in the direction of interference and away from intervention" because they are less coercive, [89] but more direct causations present more insurmountable evidentiary barriers. Taking a step back, the threshold question of whether coercion requires a "direct causal nexus between the act in question and the coercive effect" is itself largely unresolved, as a number of the Tallinn Manual 2.0 experts believe that direct causation is necessary for the establishment of coercion.[90]

In spite of this difficulty of establishing a causal link of coercive force between the conduct of CDOs and their resulting effects, scholars motivated by the goal of regulating CDOs especially in the cases of election-meddling have attempted to contextualize "coercion" in the cyber backdrop.

Schmitt, for instance, suggests that indirect causation could serve as a proxy to the notion of coercion in cases of CDOs. He, together with a majority of the Tallinn Manual 2.0 experts, takes the position that no direct causation is required to satisfy the element of coercion when evaluating the wrongfulness of an international act under the norm of non-intervention.[91] If indirect causation could satisfy the "causal facet of coercion,"[92] then as a matter of law, CDOs could breach the norm of non-intervention even if they haven't factually altered the election result, as the use of coercive methods is all that is needed to qualify an action as intervention regardless of its eventual success. Realizing that indirect causations frame the acts in question more like what Tallinn Manual 2.0 considers as acceptable interferences instead of impermissible interventions, Schmitt attempts to distinguish CDOs from other acts of indirect causations by emphasizing CDOs' "covert nature" and the "extent to which they distorted the accepted U.S. electoral dynamic."[93]

---

[87] Schmitt, *supra* note 10, at 32.
[88] *Id.* at 50.
[89] *Id.* at 51.
[90] *Id.*
[91] *Id.*
[92] *Id.* at 52.
[93] *Id.*

However, this theory of indirect causation is self-defeating, as Schmitt's attempt to shove CDOs into the concept of coercion distorts coercion to the extent that the concept is no longer recognizable as the one presented by the ICJ in Nicaragua. If an indirect causation could serve as a matter-of-law proxy to coercion, then the threshold artificially constructed in Nicaragua and designed to exclude "non-coercive" soft economic and diplomatic measures so to make the norm of non-intervention more acceptable, will be dismantled to ground. Schmitt tries to make a distinction between CDOs and other soft measures, highlighting the stealth and the disruptive impact of CDOs, but such a move salvages hardly anything. Stressing the distortion caused by CDOs is circulatory, as the evaluation of distortion presupposes an evaluation of causation. Secrecy, on the other hand, is largely at odds with the common understanding of coercion associated with compulsion, pressure, deprivation of free will, and subordination of agency. Secrecy cannot conceptually add to the aggravation of an act measured according to the metric of coercion, when coercion inherently connotes a relationship of domination (i.e., assertion of power) and subjugation (i.e., acknowledgment of power).

Schmitt's approach also begs the question of what counts as sufficient indirect causation (what is a "causal facet of coercion"?); shifting the interpretive burden from the standard of coercion to that of indirect causation hardly solves anything, if not exacerbating the confusion and ambiguity. The international community, having a hard time consenting to the broad interpretation of intervention in UNGA Resolution 36/103, is unlikely to buy the idea of coercion with indirect causation, as it undermines the strength of a common ground understanding without providing more clarity.

Harriet Moynihan, an associate fellow of the UK-based international research institute Chatham House, proposes[94] another way of stretching the standard of coercion in the cyber context. She starts with the position that the ICJ's dicta on the principle of non-intervention, particularly the coercion standard, should not be read restrictively[95] as the ICJ in Nicaragua only seek to "define the aspects of the [non-intervention] principle which appear to be relevant to the resolution of the dispute," addressing acts of interventions that involve direct use of force or assistance to use of force. Indirect support of military or aggressive actions in Nicaragua is inscribed as coercive accomplice, which to qualify as an offense against the non-intervention principle, must fall under one ground of prohibition necessitated by the essence of non-intervention. The ICJ, therefore, defined non-intervention with the standard of coercion, but left unfinished the work of refining the notion of

---

[94] HARRIET MOYNIHAN, CHATHAM HOUSE, THE APPLICATION OF INTERNATIONAL LAW TO STATE CYBERATTACKS: SOVEREIGNTY AND NON-INTERVENTION (2019).
[95] *Id.* at 27.

coercion, as precision was not needed in applying the crude coercion standard to the particular facts in Nicaragua. In Nicaragua, assistance of military action, because of its proximity to the use of force, without much controversy falls under the general sense of coercion. But in the context of CDO, proximity to military actions is absent.

Moynihan contemplates the elasticity of the coercion standard when applying it in a different context unrelated to forcible or military actions. She thinks that there is no reason why "a flexible approach to coercion" should not apply in the cyber context, as the ICJ in dictating the standard of coercion still preserved some room for contextualization.[96] Coercive behavior, she argues, may be understood as "pressure applied by one state to deprive the target state of its free will in relation to the exercise of its sovereign rights in an attempt to compel an outcome in a matter reserved to the target state."[97] Lowering the bar of coercion from the traditional formulation of "dictatorial interference" resulting in the "subordination of the will of one sovereign to another" to simply pressure, Moynihan nevertheless maintains that an element of compulsion is necessary, and that to evidence existence of compulsion, there should be "actual or potential effects" on the target state's free will.[98]

Even Moynihan's broad coercion standard is no solution for CDOs, as pressure and compulsion are hard to find in a CDO. For a start, CDOs, due to their secretive nature, do not affirmatively assert pressure on their target audience or states; on the contrary, the last thing that CDOs want is for their audiences to feel anything other than a genuine belief of autonomy and free-will when processing the fraudulent information, even if they are in fact being manipulated by CDOs to achieve the ends of others. Assertion of pressure brings attention, and therefore is antithetical to the goals of CDOs. To trick and defraud, one needs to stay under the radar. Absent the action of affirmative assertion, CDOs can only be said to compel or pressure due to their behavior-changing effects, reverting the analysis back to the causation and scale-of-effect test, a nearly impassable road for inculpating CDOs. Not only will a national survey of "would your actions or conclusions change if you haven't seen any disinformation propagated by CDOs" be pragmatically impossible, but also people will not be able to quantify the influence from disinformation compared to that from their other decision-making factors, even if such a "but-for" survey can be produced, conducted, and collected.

In spite of the belief of many scholars' that the means and techniques used by a state "to coerce another state in relation to the exercise of the latter's state powers can be various and nuanced,"[99] it is rather difficult for them to

---

[96] *Id.* at 31.
[97] *Id.*
[98] *Id.* at 33.
[99] *Id.* at 29.

develop a coherent theory that adequately fills up the gap between the essentially forcible, or at least clearly proselytizing nature, of "coercion" and the nebulous social-conditioning of CDOs, the effects of which are nearly impossible to evaluate, let alone to prove. While Moynihan might not have CDOs on her mind when expanding the notion of coercion in the cyber context, Schmitt's struggle in categorizing CDOs shows a fundamental hardship in spelling out the evils of CDOs using the vocabulary of coercion.

CDOs warrant a different standard that are not pivoted on the idea of coercion, to facilitate a much-needed international conversation. As I will propose in Section IV, some of the essential vices of CDOs that go against the spirit of non-intervention are not "coercions" but rather "manipulations and frauds," which when magnified by the power of cyber and AI technology, creates global suspicion and distrust detrimental to international peace and security. When the use of armed forces and other conspicuous forms of pressure are closely monitored by the global community and regulated by international law, secret, cheap, and unregulated CDOs are bound to propagate as reports in Section II suggest. Because of the stealthy nature of CDOs, which makes it so hard to bar them under the coercion standard, CDOs will likely create a prisoner's dilemma among the states due to worries of asymmetrical information. The threat of possible overhanging CDOs could spur pre-emptive defensive cyber mechanisms, the legality of which is also questionable. In the absence of cyber declarations or cyber treaties that signal certain consensus among states, deter potential violators, and assure global communities of the integrity of information, CDOs and their fast-developing permutations will continue to spread in the cyber chaos and legal vacuum, imposing significant cost and instability on inter-state relationships.

3. The Principle of Sovereignty Cannot Sufficiently Address Non-Coercive CDOs

While the norm of non-intervention steps in to fill a gap in international law when a coercive act is considered wrongful but does not arise to the level of Article 2(4) "use of force," some scholars argue that since the coercion standard of non-intervention is also a high one incapable of covering many problematic yet "non-coercive" conducts, the principle of sovereignty should be relied on in adjudicating if such non-coercive conducts are wrongful.[100]

The principle of sovereignty is generally considered violated when "a state exercises its authority in another state's territory without consent in relation to an area over which the territorial state has the exclusive right to

[100] MOYNIHAN, *supra* note 94, ¶ 94.

exercise its state powers independently," or as Tallinn Manual 2.0 put, when a state launches cyber operations that "prevent or disregard another State's exercise of its sovereign prerogatives."[101]

The principle of sovereignty, however, is more controversial compared to the norm of non-intervention as to whether it entails a primary rule of international law and enjoys the status of customary international law.[102] Schmitt describes such a contention as the first grey zone of the principle of sovereignty when applied to cyberspace, as some parties argue that "sovereignty is but a foundational principle that yields no sovereignty-specific primary rule of international law."[103] This position was initially proposed by three former and current senior U.S. Department of Defense officials, and has since been challenged by Schmitt and other Tallinn Manual 2.0 experts as inconsistent with extensive state practices in "non-cyber context" that treat the principle of sovereignty as a primary rule. However, substantial debates remain in the works of other scholars about "the extent to which the notion of territorial sovereignty applies to cyberspace at all," as territorial sovereignty is typically associated with physical incursion and the use of force.[104] Again, the question of where to draw the line of wrongfulness regarding inter-state activities that fall short of forcible nature surfaces, but this time involving a principle that is traditionally territorial in nature, even less acclimated to the non-territorial cyber context than the norm of non-intervention.

Schmitt, together with the majority of experts in Tallinn Manual 2.0, adopts the view that "sovereignty is the basis for a primary rule of international law by which the cyber operations of one state can violate the sovereignty of another," citing key language from Corfu Channel, the ICJ's first case, that "respect for territorial sovereignty is an essential foundation of international relations."[105] This results in Rule 4 of the Tallinn Manual 2.0—a state must not conduct cyber operations that violate the sovereignty of another States. While this controversial cyber-principle of sovereignty seemingly captures non-coercive cyber activities that the principle of non-intervention is incapable of covering, it is nevertheless substantially limited by its territorial elements.

Schmitt points out the second grey zone of the principle of sovereignty when dealing with "remote cyber operations conducted from outside the target State,"[106] as those cyber operations are unattached to the

---

[101] SCHMITT & VIHUL, *supra* note 83, Rule 4 ¶ 1.
[102] Michael N. Schmitt, *Grey Zones in the International Law of Cyberspace*, 42 YALE J. INT'L. L. 1-2 (2017).
[103] *Id.* at 4.
[104] MOYNIHAN, *supra* note 94.
[105] SCHMITT & VIHUL, *supra* note 83, at 5.
[106] *Id.* at 6.

territory of the target state; this grey zone makes the application of Tallinn Manual 2.0 Rule 4 to CDOs, a form of remote and covert cyber operations, extremely difficult. All experts agree that remote cyber operations causing "physical damage or injury" in the target state would be a clear violation of the state's sovereignty, and a majority of the experts believe that "loss of functionality of cyber infrastructure" located in another state would suffice. But no consensus can be reached as to "whether, and if so, when, a cyber operation that results in neither physical damage nor the loss of functionality" could amount to a violation of sovereignty,[107] except for when a State's cyber operation "interferes with or usurps the inherently governmental functions of another State."[108]

Not surprisingly, the term "inherently governmental functions" could not be defined, as experts could only agree to provide some examples of interference or usurpation of inherently governmental functions without defining the term. CDOs based on those examples do not fall under the meaning of the term. One particularly relevant comparison is made by the experts between "official communications among a State's leadership" and "when a State posts information on terrorist organizations on a website."[109] The experts think that the former situation is inherently governmental while the latter is not, because in the latter situation, "other entities, such as non-governmental organizations, also engage in it."[110] This literal approach to the term "inherently governmental functions" clearly excludes those activities targeted by CDOs, which necessarily involves non-governmental entities, namely the public.

This conclusion is consistent with Schmitt's analysis in "Virtual" Disenfranchisement.[111] There, Schmitt argues that the holdings of elections are "a paradigmatic example of an inherently governmental function,"[112] but whether or not different meddling activities related to elections qualify as "interference" or "usurpation" of the election-holding function needs a case-by-case evaluation, especially "by virtue of their effect on an election."[113] For instance, operations would "plainly qualify" if they "altered election data" or "rendered it impossible for voters in a particular district to cast their votes."[114] This is presumably because, according to the above illustrations given by the Tallinn Manual experts in trying to define "inherently governmental functions," those operations happen at a stage that only governmental

---

[107] SCHMITT & VIHUL, *supra* note 83, at 21.
[108] *Id.*
[109] *Id.* at 22.
[110] *Id.*
[111] Schmitt, *supra* note 10.
[112] *Id.* at 45.
[113] *Id.*
[114] *Id.* at 46.

functions of elections are involved, mainly the state's ability to accurately collect the votes already casted by voters. By contrast, if the influence of cyber operations happens before the stage of collection (i.e., in conditioning voters' minds), then the operations do not tamper with the governmental aspect of the election, rendering those operations seemingly innocent in the eyes of the sovereignty principle.

Schmitt does not explicitly make the distinction between the governmental aspect of elections and the civil aspect of elections as pertaining to different treatments under the principle of sovereignty as the Tallinn Manual experts as a group do, but it is clear to him that not every election-related cyber operation violates the principle of sovereignty, especially those operations that engage in mere "election propaganda."[115] While Schmitt comments on Russia's 2016 trolling operations as activities that arguably "tipped the scales and therefore constituted unlawful interference,"[116] he concedes that such a position is extremely uncertain. Russia at the end of the day "conducted its operations in the grey zone of the law of sovereignty," and therefore avoided the "international community's opprobrium for violating international law."[117]

4. International Rules About PsyOps Fail to Discipline CDOs

Lastly, people have also suggested treating CDOs as a new form of psychological operation and looking into international solutions regarding PsyOps.[118] But international laws related to PsyOps are insufficient to deal with CDOs.

PsyOps are usually associated with the dissemination of propaganda designed to influence the mind of the adversary, often times with false rumors and disinformation. PsyOps also seek to incite discord among an adversary's society so that the "enemy population [could] revolt against its government."[119] In recent years, PsyOps have been increasingly conducted in the cyber domain as supplementary to military operations. Although international law does not provide any official definition for PsyOps, relevant international regulations always put PsyOps in the context of belligerent operations—war, terrorism, or instigation of civil strife. PsyOps used in war are governed by jus in bello, where PsyOps used for the purpose of aggression,

---

[115] *Id.*
[116] *Id.* at 47.
[117] *Id.* at 48.
[118] Ashley C. Nicolas, *Taming the Trolls: The Need for an International Legal Framework to Regulate State Use of Disinformation on Social Media*, 107 GEO. L. J. ONLINE 36 (2018).
[119] Kalliopi Chainoglou, *Psychological Warfare*, *in* MAX PLANCK ENCYC. OF PUB. INT'L LAW, https://opil.ouplaw.com/view/10.1093/law:epil/9780199231690/law-9780199231690-e385?rskey=T8lgNZ&result=2&prd=OPIL.

terrorism, or instigation of civil strife are regulated by several UNGA resolutions, including Resolution 2131 and 2625.

On one hand, when we analyze PsyOps in Warfare, there is always the threshold question if we can fit CDOs in those existing frameworks and vocabularies about belligerent PsyOps: whether or not CDOs even amount to PsyOps in the context of military operations, or for purposes of aggression, terrorism, or instigation of civil strife. It is rather obvious that jus in bello does not apply to CDOs; CDOs as I have defined in the beginning of this paper happen in peace-time where there is no declared state of war, and CDOs would not amount to a use of force either. It is also far-fetched to say that general CDOs are deployed for the purposes of terrorism or instigation of civil strife, as both require a clear element of violence or threat of violence. Therefore, doctrines governing PsyOps supplementary to military operations or aggressions, do not apply.

Peace-time PsyOps directed against civilian populations, on the other hand, are more complex, as international laws are largely blank on the issue.[120] As modern peace-time PsyOps almost inevitably employ cyber measures, the question arises as whether peace-time PsyOps should be collapsed into CDOs and evaluated according to other "existing mechanisms,"[121] like the principle of sovereignty and the principle of non-intervention, or whether CDOs should be evaluated according to popular views regarding conventional peace-time PsyOps (i.e., propaganda). The first approach would be to eliminate the intermediary notion of PsyOps and direct us back to the use of force and non-intervention analysis.

The second approach is adopted by Ashley C. Nicolas in her article Taming the Trolls,[122] where she found that "throughout history, psychological operations have been inherently limited in scope and considered legal insofar as they did not constitute perfidy or violate the prohibition of intervention."[123] Nicolas attempts to dissect the doctrine of perfidy but finds the current understanding of perfidy incapable of covering CDOs. Just like there is a distinction between permissible interference and unlawful intervention, there is also a distinction between "permissible deception and unlawful perfidy."[124] Prohibition of Perfidy, codified in Article 37 of Protocol 1 to the Geneva Conventions, outlaws "acts inviting the confidence of an adversary to lead him to believe that he is entitled to, or is obliged to accord, protection under the rules of international law applicable in armed conflict, with intent to betray

---

[120] Tim Hwang & Lea Rosen, *Harder, Better, Faster, Stronger: International Law and the Future of Online PsyOps* 7 (Univ. of Oxford, Working Paper No. 1, 2017), https://pdfs.semanticscholar.org/fc59/e18abf00b30944392fe0242bac26a222bdc9.pdf.
[121] *Id.*
[122] Nicolas, *supra* note 118.
[123] *Id.* at 38.
[124] *Id.* at 48.

that confidence."[125] However, tactics including "the use of camouflage, decoys, mock operations, and misinformation" are not prohibited, as they do not invite an adversary to believe that he will be accorded protection under the rules of international law.[126] Therefore, CDOs, under the current understanding of perfidy, would not be prohibited.

Nicolas finds the perfidy standard too outdated to cover emerging cyber threats that involve "fundamental change in the scope and power of weaponized social media;" she also finds that the Tallinn Manual 2.0 permits the use of ruses in cyber operations that "explicitly include psychological operations."[127] Nicolas then goes on to examine UN Article 2(4)'s "use of force" standard and the norm of non-intervention, and concludes that those international legal standards are inadequate for CDOs as PsyOps as well.

At the end of her study, Nicolas proposes that a new paradigm needs to be developed to "address this [fundamental] change, not in degree, but in kind."[128] International consensus should be established through global or regional declarations, treaties, and protocols. While Nicolas is correct that new criteria needs to be introduced to address the problem of CDOs, her categorization of CDOs as PsyOps and her subsequent analysis is flawed for a couple of reasons. First, Nicolas' analysis of how Article 2(4) and the norm of non-intervention applies to CDOs does not require her categorizing CDOs as PsyOps. If the norm of non-intervention could already sufficiently address CDOs, framing CDOs as PsyOps would add nothing to the argument if the notion of PsyOps needs to be evaluated according to the norm of non-intervention in the first place. Second, the doctrine of perfidy is even less of a candidate for constructing a new paradigm for CDOs than Article 2(4) and the norm of non-intervention are. The definition of perfidy as it currently stands explicitly and unambiguously contains an element of "armed attack," a threshold impossible for CDOs to pass before the doctrine of perfidy would become applicable.

IV. A NEW PARADIGM OF COMBATTING CDOS SHOULD CONCENTRATE ON IMPERMISSIBLE MANIPULATION AND FRAUD

A. Combatting CDOs is an Imperative Duty of the International Legal Order

CDOs have increasingly become a global phenomenon. A new form of powerful weapon invented initially for domestic information suppressions

---

[125] Protocol Additional to the Geneva Conventions of 12 August 1949.
[126] *Id.*
[127] *Id.* at 49.
[128] *Id.* at 53.

and manipulations, CDOs eventually were deployed by states in inter-state relationship to achieve what traditionally only forcible measures can achieve—disturbances of foreign politics and changes of domestic policies in foreign countries. While CDOs facially produce no casualties or physical harms, they nevertheless generate far-reaching negative consequences than their innocuous appearances seem to suggest. Most of all, the covert nature of CDOs sow's interstate mistrust and doubt resulting from a state's anxiety, fear, and paranoia in its inability to detect and respond to foreign initiated CDOs. The manipulative nature of CDOs undermines the integrity of global information, exacerbating the challenges of fact-checking already grim in a multilingual environment. Leveraging modern technologies, the influence of CDOs can potentially be extended to a base, the scope of which any single forcible weapon throughout history has been unable to reach. CDOs, therefore, present challenges that are simultaneously technical, political, and military, the solutions of which are by no means simple and should in no way be just domestic and private.

The reason as to why combatting CDOs is an imperative duty of international law is multifold. It includes the insufficiency of private and domestic measures, the severity of CDOs' destabilizing effects resulting largely from the lack of applicable international rules, and the demand of states on development of cyber norm consensus. To start with, domestic laws, private self-regulations, and technological solutions are not sufficient for CDOs, as these measures alone, absent any international legal order, are inadequate to either deter foreign initiated CDOs, or to alleviate inter-state mistrusts. To see the insufficiency of deterrence from domestic measures, one needs only to be reminded of the fact that CDOs are international operations conducted remotely in a foreign jurisdiction out of the reach of the target state. Many, if not most, of the personnel involved in CDOs will likely never set foot on the land of the target states, thereby unscathed by any civil or criminal penalty that domestic laws might propose to impose on them for conducting unlawful CDOs. Evidence would also be hard to collect in the course of their indictments.

Consequently, the indictment of foreign nationals for cyber-crimes and the enforcement of punishments are extremely difficult, as the criminal case against Konstantin Kilimnik—a Russian political consultant implicated in Russia's 2016 US election interference—shows. Although a central part of the theory of the Mueller Report, Kilimnik could not be indicted for his alleged lobbying activities and ties to the Russian intelligence agencies, and was charged mostly with obstruction of justice. He remains out of the hands

of U.S. law enforcement.[129] The U.S., in 2018, charged another Russian national, Elena Khusyaynova, for midterm election meddling, with "conspiracy to defraud the United States in violation of Title 18, United States Code, Section 371."[130] Again, nothing more could be done: three days after the U.S. unsealed the court documents that charged Elena Khusyaynova with conspiracy, she appeared on a Russian media outlet voicing her surprise at the charges against her.[131] Unlike espionage activities that sometimes involve physical infiltration of target states and therefore subject the spies to risks of arrest and other legal enforcements from the target states, CDOs can be done entirely from out of the territory and jurisdiction of target states, drastically diminishing any deterrence effect that could result from sanctions on individuals and entities, especially when those individuals and entities are entirely foreign-based and have no commercial interactions with the target states.

The U.S. has, in the past, adopted several economic sanctions on individuals and entities engaging in cyber activities through the Office of Foreign Assets Control ("OFAC"). OFAC was authorized by Executive Order 13694, later amended by Executive Order 13757, to sanction individuals and entities related to "interfering with or undermining election processes or institutions."[132] Those economic sanctions mainly involve "block[ing] the property and interests in property of persons" that are determined by the Treasury in consultation with the Attorney General and the Secretary of State as having conducted or aided in the conduct of impermissible cyber activities enumerated in the Executive Orders.[133] Once determined, those individuals, entities, vessels, and aircrafts would go on OFAC's Specially Designated Nationals and Blocked Persons List (SDN); their assets are then frozen and any U.S. person are "generally prohibited from dealing with them." Such economic sanctions on individuals and entities are the major form of U.S. sanctions on Russia, which do not target the Russian state directly.

According to a report from the Congressional Research Service, a public policy research institute of the U.S. Congress, their effectiveness, however, has been hotly debated as "the relationship between sanctions and

---

[129] Allison Pecorin, *What you need to know about the indictments against Konstatin Kilimnik,* ABC NEWS (Feb. 20, 2019, 2:58 PM), https://abcnews.go.com/Politics/indictments-konstantin-kilimnik/story?id=61148969.

[130] Criminal Complaint at 1, United States v. Elena Alekseeyna Khusyaynova, No. No. 1:18-MJ-464 (E.D. Va. Sept. 28, 2018).

[131] Quinta Jurecic, *Where in the World is Elena Khusyaynova?,* LAWFARE (Oct. 26, 2018, 10:17 AM), https://www.lawfareblog.com/where-world-elena-khusyaynova.

[132] Exec. Order No. 13694, 3 C.F.R. 13694 (2015).

[133] OFFICE OF FOREIGN ASSETS CONTROL, DEP'T OF THE TREASURY, CYBER-RELATED SANCTIONS PROGRAM (2017), https://home.treasury.gov/system/files/126/cyber.pdf.

changes in Russian behavior is difficult to determine."[134] Some concerns of sanction ineffectiveness include Russia's willingness to incur the cost of sanctions and the possibility that those sanctions may "target individuals that have less influence on Russian policymaking than the United States assumes;"[135] but directors of CDOs might not even have any asset in the target state to be seized in the first place. In any case, a domestic economic sanction largely hammers only individuals and entities but not states, especially when the state could not be proved to have violated international law. The deterrence effects of such sanctions are questionable, if not severely insufficient, particularly in the context of CDOs that can be conducted completely remotely without any operator or director having a commercial relationship with the target state.

Commentators and researchers have also advocated for private regulations (i.e., platform responsibilities) and technological solutions as defensive measures to combat CDOs. The Nemr Report, for instance, outlines current efforts adopted by social media platforms in containing the spread of foreign disinformation and presses online media platforms to take on more responsibility in countering CDOs.[136] These measures, however, are largely retroactive and not proactive, aiming to mitigate potential harms done by CDOs yet incapable of detection, deterrence, or prevention. The effectiveness of the measures is likewise questionable. While those actions may now make it harder for foreign states to influence politics through CDOs as efficiently as they used to do, the measures by no means present insurmountable barriers for CDOs. CDOs, as various reports in Section II suggest, are cheap to conduct and do not require a lot of technical or hardware capacities. Disinformation operations are labor-intensive instead of asset intensive in comparison to military operations.[137] Furthermore, cyber disinformation can be copied and spread at the tip of one's fingers such that whatever posts taken down by media platforms can bounce back under a different troll at any time. States where labor is inexpensive would therefore hardly be deterred by these platforms' measures, as those measures do not seem to substantially add to the cost of CDOs.

Furthermore, platforms can easily be subjected to the criticism of "censoring free speech and unfairly targeting political views," dis-incentivizing them to tackle trolls and disinformation for fear of losing neutrality[138] if CDOs are not otherwise declared wrongful by international

---

[134] CORY WELT ET AL., CONG. RESEARCH SERV., R45415, U.S. SANCTIONS ON RUSSIA (2020), https://fas.org/sgp/crs/row/R45415.pdf.
[135] *Id.* at 3.
[136] NEMR & GANGWARE, s*upra* note 27.
[137] *Stemming the Tide of Global Disinformation,* COUNCIL ON FOREIGN RELATIONS (Oct. 11, 2019), https://www.cfr.org/event/stemming-tide-global-disinformation.
[138] NEMR & GANGWARE, *supra* note 27, at 29.

laws. Lastly, platforms' self-discipline measures cannot keep up with the ever-evolving forms of CDOs. For instance, the use of deepfake technology in generating CDOs is one thing that a platform by itself is incapable of fighting against, as counter-fake technologies severely lag behind deepfake technologies. What technology alone cannot deal with, the fence and prowess of law must step in to provide preventative deterrence.

International law therefore should take on itself the task to set norms for cyber-related inter-state activities. As those activities gradually become one of the most utilized forms of inter-state influence, reformulations of conventional international principles in cyber context are inevitable. On one hand, CDOs, because of their unique features of stealth and manipulation, have the potential to be the emblematic test case for a new international legal paradigm. On the other hand, the lack of such an applicable international legal paradigm contributes substantially to why CDOs are becoming rampant. While it is crucial to point out the insufficiency of private and domestic measures in addressing the challenges of CDOs, to see why a new paradigm of international law is an indispensable piece of the CDO puzzle, I want to re-examine the consequences of unregulated CDOs if we fail to adopt a new paradigm of international law.

Drawing on the observations on CDOs and existing international laws, I offer the following possible events: (1) current international legal doctrines, including UN Article 2(4)'s "use of force," the principle of non-intervention, and the prohibition of perfidy, all presume certain levels of aggression and coercion, notions both derived from a narrative of traditional armed conflicts and made to address measures effectuating results similar to those achieved by armed conflicts; (2) technology and cyber capacity have gradually made it unnecessary to effectuate political results through facially aggressive or coercive means on the physical territory of a foreign state; (3) old standards of aggression and coercion consequently become ill-suited to address covert cyber operations, leaving a gap in international law; (4) the twilight zone of international law in turn incentivizes states to exploit the gap and to engage in shady cyber practices like CDOs; (5) during this process, CDOs that are fraudulent and covert greatly undermine interstate trust because of states' worry of asymmetrical information, which also creates significant inefficiency as states engage in expensive remedial measures of discerning and filtering disinformation; (6) the lack of an international consensus on cyber countermeasures also exacerbates interstate instability as states may adopt over-defensive tactics that are not proportionate in scope; and (7) as technology and cyber capacity continues to advance while international law lags behind, international peace and security will inevitably deteriorate to the point where the current international laws, in refusing to adapt to the changes in cyber space, destroy their own primary objectives of peace-keeping.

Many of these events have already happened, as various reports cited in Section II suggest; the rest are highly likely to happen absent an event-breaker of international cyber consensus adopting a new paradigm other than "use of force," coercion, or territoriality. For instance, the proliferation of cross-border CDOs will almost certainly ensue given how cheap, accessible, and innocent CDOs are in comparison to conventional military measures. CDOs will also destabilize inter-state relationships, as they are perceived as threatening yet inconspicuous and un-punishable under current international laws. As I have mentioned, Russia claimed that its CDOs into other states were responding to threats from external cyber informational activities, which are themselves a form of CDOs. The U.S. also put forward several presidential executive orders imposing economic sanctions through OFAC in the event of foreign interference,[139] culminating in the most draconian one - Executive Order 13848. This Order, on top of its regular making of SDN, also provides authority to the Secretary of the State and the Secretary of the Treasury to sanction the largest business entities licensed or domiciled in the interfering state in sectors of particular "strategic significance" to that interfering state, regardless of whether or not those business entities have engaged in foreign interference[140].

While we previously questioned the effectiveness of economic sanction regimes, as those regimes targeting individuals and entities might be insufficient to deter CDOs, here we have a different worry of over-punishment when no standard of international law can provide guidance on the proper counter responses. Targeting the largest business entities could significantly undercut the economy of the entire sanctioned state, which might seem to many as an unjustified over-response to CDOs. Executive Order 13848 contains an implicit proportionality requirement as "all recommended sanctions shall be appropriately calibrated to the scope of the foreign interference identified,"[141] but the standard of proportionality is not discussed and seems to be within the discretion of "appropriate agencies." Without clearer international laws on CDOs, such discretionary countermeasures would be one of the only few resorts for a state to seek justice.

Even if the prisoner's dilemma is overstated in the event cascade, and that interstate stability would not be significantly undermined by the lack of international cyber norms, it is still to states' great benefit to start establishing international cyber norms, to provide deterrence against cyber interference, and to commit to information integrity and cost-effective mechanisms to mitigate the externality of cyber chaos.

---

[139] Exec. Order No. 13848, 83 Fed. Reg. 46843 (Sept. 12, 2018).
[140] *Id.*
[141] *Id.* at 46845.

States are starting to demand the construction of international cyber norms. The Tallinn Manual 2.0 project is but one piece of evidence of the demand. Most recently, a cyber group called the Open-Ended Working Group ("OEWG"), was created by the U.N. in 2018 at the initiative of Russia to discuss how international law should stop cyber warfare.[142] A meeting was subsequently held in New York where representatives from Russia warned the audience of a "global cyberwar."[143] As Lukasz Olejnik, a cyber-security researcher that was present at the meeting said, "everyone agrees international law applies to cyberspace; the trick is how it applies."[144] States, like in the case of Resolution 2131 and 2625, agree that some high-level abstract principles of order should regulate cyberspace. However, they diverge significantly on the details of the implementable rules, especially how a new cyber paradigm should dovetail existing international norms and how it will interact with international humanitarian laws.

B. The Principles of Non-Intervention are the Best Available Tools to Combat CDOs

There is a general high-level consensus that international laws and norms apply to cross-border cyber operations,[145] but precise applications of these norms to specific cyber operations are still being debated and the international community has failed to reach a consensus on the detail of applications in most of cyber operations short of Article 2(4)'s "use of force." Similar to what has been shown in the previous "use of force" and norm of non-intervention analysis, different states have conflicted interests with each other, and even conflicted self-interests. As Schmitt summarizes, "a permissive view of international law would afford [states] leeway to conduct their operations abroad but leave them without normative firewalls that will enhance their cyber security."[146] Cyber security is a competitive matter of national security, where states constantly evaluate their own capacities in relation to those of other states and advocate for an international framework that better preserves their competitive advantages over those of others. Meanwhile, the international legal vacuum regarding CDOs is undesirable for

---

[142] The United Nations Doubles Its Workload on Cyber Norms, and Not Everyone Is Pleased, CFR (Nov.15, 2018), https://www.cfr.org/blog/united-nations-doubles-its-workload-cyber-norms-and-not-everyone-pleased.

[143] Tim Starks, *UN Debates Cyber Treaty, Norms,* POLITICO (Sep. 16, 2019, 10:00 AM), https://www.politico.com/newsletters/morning-cybersecurity/2019/09/16/un-debates-cyber-treaty-norms-743266.

[144] *Id.*

[145] Michael N. Schmitt, *Peacetime Cyber Responses and Wartime Cyber Operations Under International Law: An Analytical Vade Mecum*, 8 HARV. NAT'L SEC. J. 242 (2017).

[146] *Id.* at 242-43.

many states for the reasons illustrated in the previous sections. Therefore, moving forward, what states will and should be seeking is how to achieve an acceptable "give and take" consensus—even regionally.

To gain a consensus of international cyber norms and principles consistent with existing international laws, two different approaches could be utilized. One can read existing international laws as narrowly tailored towards addressing territorial and forcible inter-state conducts. By doing so, one can frame cyber activities as categorically different from those traditional activities, thereby proposing an entirely new regime for cyberspace that is compatible, yet also independent from the logic of existing international laws. Alternatively, one can read cyberspace as an extension of territorial space, and the goals of cyber activities as analogous to those of forcible measures. Consequently, the spirit and logic of existing international laws would still be the basis for governing cyber activities, but modifications would need to be made to accommodate for the features of cyber activities that deviate from the traditional model. This paper advocates for the second approach, as it has been the direction of international cyber discussions so far. State audiences are less skeptical of a new paradigm built upon existing international laws than a new cyber regime that needs to be created out of thin air.

International discussions of cyber responsibility have largely been prompted and developed in the context of international security in the UN. A draft resolution was first introduced by Russia in 1998 in the First Committee of the General Assembly and later "adopted without a vote… as [a] resolution 53/70." [147] The Resolution highlighted the concern that information technologies "can potentially be used for purposes that are inconsistent with the objective of maintaining international stability and security," and urged member states to develop basic notions related to information security, including the "misuse" of information systems.[148] Subsequently, six Groups of Governmental Experts (GGEs), which are UN-mandated working groups, have been established since 2004 on the basis of "equitable geographical distribution."[149] The GGEs consists of the five permanent members of the Security Council and, at different times, other selective states who need to officially request for a seat on a GGE that is of a particular interest to them[150]. The most recent 2019-2021 GGE is currently ongoing and is composed of

---

[147] U.N. OFFICE FOR DISARMAMENT AFFAIRS, DEVELOPMENTS IN THE FIELD OF INFORMATION AND TELECOMMUNICATIONS IN THE CONTEXT OF INTERNATIONAL SECURITY (2019), https://unoda-web.s3.amazonaws.com/wp-content/uploads/2019/07/Information-Security-Fact-Sheet-July-2019.pdf.
[148] G.A. Res. 53/70, ¶ 6-8 (Dec. 4, 1998).
[149] UN GGE and OEWG, GENEVA INTERNET PLATFORM, https://dig.watch/processes/un-gge (last visited Oct. 9, 2020).
[150] *Id.*

twenty-five states.[151] As mentioned in the previous section, another UN-mandated working group, OEWG 2019-2020, was established in 2018 in parallel with the GGEs, open to all interested states as well as non-state stakeholders.

Among the five previous GGE meetings, three have successfully produced reports with a consensus, where the major concerns later associated with CDOs and other cyber meddling activities already started to surface, though at the time those reports were produced, particular forms of non-forcible cyber meddling were not foreseen. For instance, the first GGE Report with a consensus, A/65/201, pointed to the necessity of international law in governing information and communications technologies ("ICTs"), as "uncertainty regarding […] the absence of common understanding regarding acceptable State behavior may create the risk of instability and misperception," which would consequently affect states' crisis management in that "no State is able to address [the] threats alone."[152] The report also recognized that ICTs were developed not only as instruments of warfare, but also as instruments of intelligence and for political purposes. The second GGE report with a consensus, A/68/98*, officially recommended that international law be applicable to the cyber environment, and that norms of state sovereignty and norms derived from sovereignty would govern. On the other hand, human rights and fundamental freedoms "set forth in the Universal Declaration of Human Rights and other international instruments" should also be taken into an account.[153] The third GGE report with a consensus, A/70/174, furthered the discussion of applicable international laws regarding cyber activities by emphasizing principles that states should commit to, including UN Charter responsibilities, a duty to refrain from the "threat or use of force against the territorial integrity or political independence of any State," and "non-intervention in the internal affairs of other States."[154]

The multi-stakeholder OEWG takes on the work left by GGE report A/70/174 to "further develop the rules, norms, and principles of responsible behavior of States,"[155] convening for the first time in 2019 and aiming to report to the General Assembly in 2020. One critical goal of OEWG is to specify "existing and potential threats."[156] Member states and other non-state

---

[151] *Id.*
[152] G.A. Res. A 65/201, at 2 (July 30, 2010).
[153] G.A. Res. A 68/98*, at 8 (June 24, 2013).
[154] G.A. Res. A 70/174, at 12 (July 22, 2015).
[155] G.A. A/AC.290/2019/1, at 1 (May 22, 2019).
[156] *Open-ended working group on developments in the field of information and telecommunications in the context of international security, Provisional Agenda and Annotations*,
U.N. OFFICE FOR DISARMAMENT AFFAIRS, (May 22, 2019), https://documents-ddsny.un.org/doc/UNDOC/GEN/N19/150/48/PDF/N1915048.pdf?OpenElement.

stakeholders have submitted papers outlining their positions on various issues. Among the several papers submitted, Egypt[157] and Switzerland[158] make explicit references to cyber disinformation as threats that OEWG should address in its agendas. The submission from Internet Governance Forum ("IGF"), a UN affiliated multi-stakeholder inter-governmental organization, also refers to disinformation and cyber interference as one of the main threats to online safety and security[159] based on the 2019 Berlin IGF messages. Similarly, the current ongoing GGE is also focusing on enumerating existing and emerging threats that demarcate the scope of discussion of responsible state behaviors in cyberspace. The Chair's Summary from the December 2019 consultative meeting cited a number of delegates' emphasis that "misuses of social media platforms" are a form of emerging threats when they are "leveraged to interfere in or influence the domestic processes of other States, including elections."[160]

As the history of UN-GGE and UN-OEWG suggests, discussions of low-level ICT activities, from general cyber interferences to specific disinformation operations, are gradually entering into international conversations of responsible state behaviors, with a mirroring advancement on the discussion of applicable international laws from Article 2(4)'s use-of-force standard, sub-force but forcible measures covered by the norm of non-intervention, and sub-forcible measures yet waiting for a solution. Such a goal-motivated continuum of international legal development thus recommends the principle of non-intervention as the best available framework to construct a new paradigm combatting CDOs, as Article 2(4)'s standard would be too unnecessary and unrealistic a stretch for addressing CDOs, while the conventional principle of sovereignty would be too territorial, unfit for CDOs that are remote in nature. The extended principle of sovereignty, which could cover remote CDOs (one proposal being the standard of "interference and

---

[157] Delegation of Egypt, *To the Open-Ended Working Group on Developments in the Field of Information and Telecommunications in the Context of International Security* (Working Paper, 2020), https://unoda-web.s3.amazonaws.com/wp-content/uploads/2020/01/Egypt-Working-Paper-OEWG-ICTs1.pdf.

[158] FED. DEP'T OF FOREIGN AFFAIRS, POSITION PAPER ON SWITZERLAND'S PARTICIPATION IN THE 2019-2020 UN OEWG 3 (2020), https://unoda-web.s3.amazonaws.com/wp-content/uploads/2020/02/switzerland-position-paper-oewg-gge-final.pdf.

[159] Internet Governance Forum, *Submission to the "Open-Ended Working Group on Developments in the Field of Information and Telecommunications in the Context of International Security"* (Feb. 2020), https://unoda-web.s3.amazonaws.com/wp-content/uploads/2020/02/igf-cybersecurity-oewg-feb2020.pdf.

[160] Group of Governmental Experts [GGE], *Chair's Summary: Informal Consultative Meeting of the Group of GGE on Advancing Responsible State Behavior in Cyberspace in the Context of International Security*, at 3, (Dec. 5-6, 2019), https://www.un.org/disarmament/wp-content/uploads/2019/12/gge-chair-summary-informal-consultative-meeting-5-6-dec-20191.pdf.

usurpation"[161]), is doing nothing more than trying to delineate the line of wrongful remote cyber activities with inter-state agendas. Since a new paradigm would be best constructed dovetailing the norm of non-intervention that is itself derived from the principle of sovereignty, a discussion of the extended principle of sovereignty as a potential solution can be rightly collapsed into proposing a new paradigm built upon the norm of non-intervention, as I shall elaborate in the next section.

C. A New Paradigm of Impermissible Manipulation and Fraud Should Be Established

A new paradigm of impermissible manipulation and fraud should take the place of the ICJ's coercion standard when applying the norm of non-intervention to CDOs. This is to account for the covert and deceptive nature of CDOs, which deteriorates inter-state trust and undermines internet users' senses of security, as well as to deter states that wish to weaponize CDOs for purposes of foreign intervention.

Such a paradigm would declare state-sponsored or state-controlled CDOs that are involved in any fabricated information or fabricated identities (together "disinformation") without disclosures of such fabrications, presumptively wrongful, without the need to show any causation or effect, be it potential or actual. Victim states could get an order of injunction and a public apology from offending states, and upon offending states' re-offense, could carry out countermeasures within the boundaries granted by international legal instruments mandating the paradigm of manipulation and fraud (i.e., legal instruments specified in a treaty, or the ICJ). The alleged states may rebut the presumption of wrongful CDO activities in violation of the norm of non-intervention in several ways, including for example by (1) showing that the original source of fabricated information is not state-sponsored or state-controlled; (2) showing that the number of views of the original disinformation posts promoted by CDOs and any repost combined, is less than a threshold amount determined by relevant treaties or conventions (this rebuttal preserves some room for a scale-of-effect rationale, but the burden is now on the alleged states and not on the victim states to show that no effective harm was done by the CDO); or (3) showing that sufficient disclosures are in place for all disinformation.

Negligence or inadvertence will generally not be a valid defense for generating disinformation, and a single piece of disinformation generated and then promulgated will taint the entire operation unless the wrongfulness of the entire operation can be cleansed by one of those rebuttals. The definition of

---

[161] SCHMITT & VIHUL, *supra* note 83.

fabricated information should be limited to only information that has no factual basis, not including information that is based on selective filtrations of facts. The alleged states will bear the ultimate burden of showing the factual basis for their promoted information.

The definition of fabricated identities should cover all attempts on social media and digital media to promote information in any manner using false identities affirmatively claiming to be citizens, groups, or entities of the victim state, regardless of whether the information promoted is fabricated. Such a definition covers not only trolling, but also certain fraudulent practices of paid advertisements. When coupled with some platforms' advertisement disclosure requirements, this prohibition against fabricated identities adds extra deterrence against foreign paid advertisements with fictitious identities or fictitious disclosures.

Lastly, evidence of bot usage to magnify disinformation in a cyber operation should be perceived as strengthening the presumption of wrongful conduct and undermining the first rebuttal that the original source of fabricated information is not state-sponsored or state-controlled. This rule addresses a possible moral hazard that offending states, instead of producing fabricated information, might just find fabricated information and promulgate them in order to get around the presumption of wrongful conduct under the first rebuttal.

The paradigm of impermissible manipulation and fraud may seem like a big leap from the coercion standard, but such a move is warranted by several factors: the urgency of addressing the challenges of CDOs, the fundamental incompatibility between CDOs and the coercion standard, the potential of the manipulation paradigm to balance the interests of non-intervention to minimalize inter-state conflicts and distrusts, and the interests of one's right to "seek, receive, and impart information and ideas" provided in the Universal Declaration of Human Rights.[162] Existing scholarship on the application of the non-intervention norm to CDOs has been reviewed and evaluated in the previous sections. The remainder of this paper will discuss how the paradigm of impermissible manipulation and fraud arises from the failures of the coercion standard, and how such a paradigm could balance the interests of states' non-intervention principle and the interests of individuals' human rights to receive and impart information freely.

Much work has already been done in applying the norm of non-intervention to CDOs, as shown in Section II.B.2. Scholars like Schmitt and Moynihan, among others, have attempted to play with the standard of coercion, either by extending the concept of coercion out of shape to accommodate for the massive, yet impossible-to-quantify, reach and effect of

---

[162] Art. 19, Universal Declaration of Human Rights, https://www.un.org/en/universal-declaration-human-rights/.

CDOs, or by limiting the coercion standard as narrowly defined for the particular ICJ case where no cyber activities were involved and thereby carving out some space for a new standard more suitable for addressing cyber activities like CDOs. While extensions of the coercion standard fail to account for CDOs' most prominent features of threats, namely CDOs' stealth and manipulation, Moynihan's suggestion of a narrow interpretation of the ICJ's dicta of coercion standard provides a useful ground to reconcile a proposal of a new paradigm with the existing coercion standard, as both standards pertain to the spirit of the non-intervention principle, and are justifiably different because of the disparate features of CDOs and Nicaragua-like measures.

The elasticity of the coercion standard, as previously argued, cannot tolerate CDOs, as the key vocabularies of coercion—pressure, subordination, compulsion—all presume either an affirmative imposition of force or will, or a passive but de facto causation of damage, both contrary to the features of CDOs. CDOs are operated in a way that avoids affirmative impositions of will, and through fraudulent and manipulative tactics to bake the interfering state's agendas into the interfered audience's psychology. Because of the intervening agencies of the audience, the causation between CDOs and the ultimate policy or election results becomes impossible to measure, let alone to prove. CDOs, instead of creating coercions, generate suspicions and insecurities in states and individuals, especially fueled by states' anxieties of their incapability to predict, prevent, or remedy the effect of covert and manipulative CDOs. By the time a CDO reaches its target audience base, the "damage" has already been done—to measure and to undo any psychological mark left by CDOs is not possible. Therefore, a new paradigm of non-intervention, without requiring a proof of causation, is needed. It provides both ex ante deterrence, and expressive signaling effects that condemn the wrongful practices of CDOs and reaffirm the value of informational integrity. The new paradigm proposed by this paper, which establishes a presumption of wrongdoing upon finding of CDOs regardless of their effects, will fulfill these goals.

The paradigm of impermissible manipulation and fraud, although distinct from the coercion standard as no causation of pressure or compulsion needs to be shown but a prima facie establishment of targeted CDOs, embodies the spirit of the principle of non-intervention. To start, the principle of non-intervention could be rooted in non-territorial grounds to deal with non-territorial threats, recalling Lassa Oppenheim's comment that the obligation not to intervene is "the corollary of every State's right to sovereignty, territorial integrity, and political independence."[163] Juxtaposed with "territorial integrity," political independence provides another ground of justification for prohibitions against non-territorial interventions. Granted that

[163] Wood, *supra* note 63.

the definition of political independence often times goes hand-in-hand with the line-drawing of what permissible inter-state activities are, and consequently the discourses of political independence and non-interventions are largely circulatory, the history of the development of the non-intervention principle nevertheless suggests great elasticity of the principle evolving along with changing inter-state threats.

The norm of non-intervention is essentially problem-oriented, and the problem—as the first UN General Assembly resolution that introduced the expanded notion of intervention, Resolution 2131, has put—is the "increasing threat to universal peace due to […] other direct or indirect forms of interference threatening […] the political independence of States."[164] The norm of non-intervention, therefore, can be perceived as a mechanism to identify any "threat to universal peace" to form or reinforce the boundaries of political independence for the purpose of "nations liv[ing] together in peace with one another," which is at risk of the identified threats, and in defining or redefining the scope of political independence to prescribe norms of permissible and impermissible activities.[165] Why CDOs are existing and potential global threats to universal peace is amply presented and argued in previous sections. To reiterate, such reasons include a CDO's technical easiness of application, its massive scope of actions, its insidious purposes of sowing division and discord, and its inconspicuous modes of operations, culminating in an overhanging inter-state paranoia and public suspicions of foreign interferences, which are detrimental regardless of the actual effects of CDOs.[166] The threat of CDOs, being different from conventional threats, therefore should instruct the norm of non-intervention to adopt appropriate new standards defining what wrongful CDO behaviors are, such that the spirit of the non-intervention norm, namely the preservation of universal peace, is manifested.

It is perhaps useful to differentiate between CDOs, traditional influence campaigns (propaganda and conventional peacetime PsyOps), public diplomacy, and cyber espionage to explain why CDOs are more threatening than those other operations to warrant protections extended by the norm of non-intervention. The particular danger of CDOs comes from three main factors: numerosity of potential influence partially stoked by the accessibility of CDO capacities and the systematic insufficiency of counter-CDO capacities, covertness of operations that fuels inter-state suspicion, and deception used as a manner of operations, which also undermines inter-state

---

[164] G.A. Res. 2131 (XX), *supra* note 69.
[165] G.A. Res. 2625, *supra* note 54.
[166] Matthew Rosenberg, *'Chaos Is the Point': Russian Hackers and Trolls Grow Stealthier in 2020*, N.Y. TIMES (Jan. 10, 2020), https://www.nytimes.com/2020/01/10/us/politics/russia-hacking-disinformation-election.html ("You don't actually have to breach an election system in order to create the public impression that you have").

trust. Using these three factors as the metric of threats, we can draw lines between CDOs and many other inter-state activities that look facially similar.

The key difference between traditional influence campaigns using disinformation and CDOs lies in the varying scope of potential influences. Although causations between CDOs and changes in policy or election results are extremely hard to establish given the intervening agency of the public, CDOs have a much larger scope of potential influence compared to that of traditional influence campaigns. Disinformation promoted and augmented through cyber means can reach a larger audience in a shorter amount of time with an inexpensive budget compared to that of a traditional influence campaign. Traditional influence campaigns are also likely to be more transparent compared to CDOs, as PsyOps do not always hide or fake the sources of its operations. Therefore, the damage done by PsyOps to inter-state trust and peacekeeping is likely to be much smaller compared to the damage done by CDOs. The importance of the scope of potential influence also explains why we would want to have a rebuttal of presumption from the allegedly offending state when the number of views of disinformation is relatively small and inconsequential.

CDOs differ from public diplomacy mainly on the second and the third factor of the threat metric: the covert and manipulative nature of operations. Public diplomacies are generally "government-sponsored efforts aimed at communicating directly with foreign publics."[167] Those operations, often times comprised of cultural communications and political advocacy, are documented in public records with no intentions to hide. Subsequently, public diplomacies do not cause concerns of secret subversions and therefore afford targeted states a transparent way of resolution, namely the chance to respond in public, without having to resort to countermeasures or pre-emptive measures for fear of asymmetrical risks.

CDOs differ from cyber espionage mainly on the first and the second factors of the threat metric. Cyber espionage is a costly operation. It requires a much higher technical capacity (i.e., hacking) compared to that of CDOs (i.e., manual trolling), and therefore bars most countries from conducting a highly sophisticated and successful operation. Cyber espionage is also more conspicuous and limited in scope. Once a cyber-security system is breached, the victim states or organizations will likely know of the breach in a short amount of time and remedy damages through response teams like the Computer Emergency Readiness Team. Also, depending on how aggressive offending states use the "stolen" confidential information, cyber espionage can be covered under Article 2(4)'s "use of force" standard or the ICJ's standard of coercion. In contrast, CDOs are hard to uncover in a timely

---

[167] *Public Diplomacy*, ENCYC. BRITANNICA, https://www.britannica.com/topic/public-diplomacy (last visited Oct. 9, 2020).

manner before they spread to significant audience bases, and thereby solutions for CDOs have to rely heavily on ex ante deterrence.

Carving out a space to prohibit wrongful CDOs in the face of human rights, advocated by the Universal Declaration of Human Rights to "seek, receive and impart information and ideas through any media and regardless of frontiers"[168] is not an easy job, but by limiting the prohibition against fabricated information to only that which has no factual basis at all instead of biased or filtered information, it helps to reconcile the paradigm of impermissible manipulation and fraud with individuals' rights of information. No rights are absolute, especially when one's right to impart disinformation impairs another's right to receive authentic and credible information in a cyber environment with a basic level of security and integrity. As the 2019 Berlin IGF Messages put, "the Internet will only achieve its potential as a channel of free speech and an engine for economic growth if it remains a safe place where people feel secure."[169] The interest in creating a credible Internet environment trusted by both states and individuals is often times at tension with the interest of freedom of information. However, a balance must be struck, as freedom of information carried to its extreme might undermine individuals' trust of the Internet and consequently cause a self-defeating reduction of free speech as people divert away from the Internet.

Attempts at, and challenges behind, striking this balance can be seen in, for instance, the Joint Declaration on Freedom of Expression and "Fake News," Disinformation and Propaganda ("Disinformation Declaration") developed by multi-stakeholders including the United Nations Special Rapporteur on Freedom of Opinion and Expression,[170] The Disinformation Declaration put forward, inter alia, two relevant standards: first, State actors should not make, sponsor, encourage or further disseminate disinformation or propaganda; and second, general prohibitions on the dissemination of information based on vague and ambiguous ideas are incompatible with international standards for restrictions on freedom of expression. These standards are based on the realization that threats and worries of disinformation work both ways: some forms of problematic disinformation and propaganda are designed to mislead a population and to impede the public's right to receive credible information, while some other forms of permissible biases and opinions, mischaracterized by public authorities as "disinformation" with hidden political agenda, can get censored against the

---

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

[169] Internet Governance Forum, *Security, Safety, Stability and Resilience* (2019), https://www.intgovforum.org/multilingual/filedepot_download/9212/1804.

[170] U.N. Special Rapporteur on Freedom of Opinion of Expression, *Joint Declaration on Freedom of Expression and "Fake News," Disinformation and Propaganda*, U.N. DOC. FOM.GAL/3/17 (Mar. 3, 2017), https://www.osce.org/fom/302796?download=true.

human rights of freedom of information. The proposed paradigm of impermissible fraud and manipulation honors the concerns of the Disinformation Declaration as it seeks to deter malicious disinformation campaigns but limits its scope of application only to information that has zero factual basis. As for trolling, the proposed paradigm only regulates the manner of imparting information, not the content of information. By doing so, the manipulation standard aims to harmonize those conflicting interests voiced by the Disinformation Declaration.

V. CONCLUSION

In spite of the threat of CDOs and the larger backdrop of cyber risks, the duty of international law has barely been borne out, especially since international laws have largely been shackled to the standards of the past, ill-fitted for modern challenges. In this paper, I have evaluated the elasticity of existing international norms when applied to CDOs and have accordingly proposed a solution embodying a new paradigm. I hope this paper can open up some discussions about CDOs moving forward, and the role of international law in general when dealing with an unprecedented area of challenges.