COMBATING DISINFORMATION THROUGH INTERNATIONAL LAW

ALI STRONGWATER

I. INTRODUCTION

While not a new phenomenon, the rapid proliferation of false, misleading, inaccurate, and manipulated information facilitated by contemporary information communication technologies (ICT) has recently wreaked havoc in several states, disrupting political activities and causing significant harm to civilians.[1] False rumors and hate speech incited violence against the Rohingya in Myanmar.[2] Donald Trump's election conspiracies culminated in vicious riots at the Capitol and continue to undermine election integrity in the United States.[3] Inaccurate information about the efficacy and safety of vaccines during the Covid-

1. *See* Center for Information Technology and Society, *The Danger of Fake News in Inflaming or Suppressing Social Conflict*, UC SANTA BARBARA (2022), https://www.cits.ucsb.edu/fake-news/danger-social (detailing how fake news led a man in the United States to shoot up a restaurant falsely linked to a human trafficking ring); Gillian McKay, *Disinformation and Democratic Transition: A Kenyan Case Study*, STIMSON (June 22, 2022), https://www.stimson.org/2022/disinformation-and-democratic-transition-a-kenyan-case-study/ (describing how foreign and domestic actors have used disinformation to interfere with elections in Kenya); Sheera Frenkel, *Lies on Social Media Inflame Israeli-Palestinian Conflict*, N.Y. TIMES (May 14, 2022), https://www.nytimes.com/2021/05/14/technology/israel-palestine-misinformation-lies-social-media.html (showing how false information circulating on social media in Israel and Palestine was inflaming tensions on both sides).

2. Steve Stecklow, *Why Facebook is Losing the War on Hate Speech in Myanmar*, REUTERS (Aug. 15, 2018), https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/.

3. Graeme Massie, *A Timeline to Insurrection: The Trump Tweets that Security Experts Say Led to the Capitol Riots*, INDEPENDENT (Jan. 18, 2021), https://www.independent.co.uk/news/world/americas/us-election-2020/trump-tweets-attacks-capitol-violence-b1786246.html.

19 pandemic depressed vaccination rates and resulted in hundreds of thousands of preventable deaths.[4]

Much of the spread of false information online can be attributed to mere error. Inaccurate or misleading information that is broadcast by actors who do not intend to cause harm is known as "misinformation."[5] However, many actors intentionally spread fabricated information or misleadingly frame genuine facts for a variety of reasons, including for economic or political gain. In this vein, "disinformation" refers to the dissemination of knowingly false, inaccurate, or misleading information with the intent to cause harm.[6] Additionally, some argue that disinformation should include the manipulation and sharing of true information in a way that creates a false impression of reality.[7]

Non-state, state, and state-sponsored actors may engage in targeted influence operations and disinformation campaigns to exploit social fissures, undermine democratic processes, disrupt political activities, and promote preferred narratives. While information influence campaigns can be directed at individuals, organizations, and groups, many states have been targeted by disinformation activities, with elections and democratic processes coming under increasing scrutiny.[8] Disinformation is also a long-favored military tactic and is often deployed in coordination with physical operations in armed conflict. False and inaccurate information has, for example, plagued the war in Ukraine, not only targeting military operations but also threatening humanitarian work in the region.[9] Disinformation during conflicts can expose

---

4. A. Martinez & Allison Aubrey, *How Vaccine Misinformation Made the Covid-19 Death Toll Worse*, NPR (May 16, 2022), https://www.npr.org/2022/05/16/1099070400/how-vaccine-misinformation-made-the-covid-19-death-toll-worse.

5. *The Strengthened Code of Practice on Disinformation*, EUR. COMM'N (June 16, 2022), https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation; *Information Disorder*, COUNCIL OF EUR. (2022), https://www.coe.int/en/web/freedom-expression/information-disorder.

6. *Id.*

7. Tomoko Nagasako, *Global Disinformation Campaigns and Legal Challenges*, 1 Int. Cybersecur. L. Rev. 125, 128 (2020).

8. *See id.* at 129 (detailing a sharp rise in the proportion of elections targeted by cyber threat activities and noting that the majority of cyber activities targeting democratic processes have been strategically deployed to affect the outcome of the process).

9. *Ukraine: Addressing misinformation about ICRC's activities*, INT'L COMM. OF THE RED CROSS (Mar. 26, 2022), https://www.icrc.org/en/document/ukraine-addressing-misinformation-about-icrcs-activities (correcting false information that had been circulating about the ICRC's activities in Ukraine).

civilians to violence, increase mental suffering, and prevent them from accessing humanitarian aid.[10]

Both the deliberate and unintentional dissemination of false information across state borders raises questions about the role of international law in cyberattacks that do not amount to a use of force. While it is clear that international law applies to activities in cyberspace, it is less clear how specific principles should be applied to cyber activities where the activity's causal connection to harm is ambiguous, such as with information influence campaigns.[11] Clarifying the application of these principles to disinformation is critical to ascertain which kinds of operations constitute a breach of international law, thus implicating state responsibility, and to assess which avenue under international law is the most appropriate to guide state responses.

II. DISINFORMATION UNDER THE PRINCIPLES OF SOVEREIGNTY AND NON-INTERVENTION

States are not generally prohibited from spreading misinformation under international law. International humanitarian law, for example, permits acts to confuse or mislead an enemy as a "ruse of war," insofar as these actions comply with other applicable rules of international law.[12] Misinformation and disinformation are generally understood to fall within the boundaries of these kinds of ruses.[13] Some, however, have criticized the incorporation of disinformation as a ruse

---

10. Elan Katz, *Liar's war: Protecting civilians from disinformation during armed conflict*, ICRC (Dec. 2021), https://international-review.icrc.org/articles/protecting-civilians-from-disinformation-during-armed-conflict-914.

11. *See* Rep. of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security, ¶ 27-28, U.N. Doc. A/70/174 (July 22, 2015) (stating that "State sovereignty and international norms and principles that flow from sovereignty apply to the conduct by States of ICT-related activities and to their jurisdiction over ICT infrastructure within their territory").

12. *Protocol Additions to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts,* Art. 37, June 8 ,1977, *1125 U.N.T.S. 3* (allowing "acts which are intended to mislead an adversary or to induce him to act recklessly but which infringe no rule of international law applicable in armed conflict" and forbidding the use of perfidy to kill, injure, or capture an adversary); TALLINN MANUAL 2.0 ON THE INTERNATIONAL LAW APPLICABLE TO CYBER OPERATIONS 184 (Michael N. Schmitt ed., 2d ed. 2017) [hereinafter TALLINN MANUAL 2.0]; CUSTOMARY INTERNATIONAL HUMANITARIAN LAW, VOL. 1: RULES 203-05 (Rule 57) (Jean-Marie Henckaerts & Louise Doswald-Beck eds. 2005).

13. Elan Katz, *Liar's war: Protecting civilians from disinformation during armed conflict*, ICRC (Dec. 2021), https://international-review.icrc.org/articles/protecting-civilians-from-disinformation-during-armed-conflict-914.

of war, arguing that contemporary information influence campaigns differ dramatically from traditional ruses of war: modern disinformation operations disproportionately target civilian populations and circulate at unprecedented speeds and scales.[14]

In certain forms, transnational disinformation and influence campaigns can violate the principles of sovereignty and non-intervention.[15] The sovereignty principle enshrines the authority of the state to exercise power over its territory to the exclusion of other states. States are thus prohibited from actions that intrude on the internal affairs of another sovereign state or from exercising "state powers within the territory of another state without consent."[16] The non-intervention rule forbids states from using coercive means to intervene in another state's sovereign concerns.[17] Cyber activities that interfere with the internal or external affairs of another state are often seen as violating this principle.[18]

Harriet Moynihan, a fellow in the International Law Programme at Chatham House, identifies two schools of thought regarding the application of these principles to state-sponsored cyber activities.[19] One view recognizes the application of the non-intervention principle "to certain state-sponsored cyber-intrusions," but any activity that does not rise to the level of a violation of this principle fails to qualify as a breach meriting state responsibility, even if "unfriendly."[20] The other view contends that cyber activities that do not violate the non-intervention principle may nevertheless violate a state's sovereignty.[21]

From an absolute perspective, the deliberate deployment of a false information operation can constitute a violation of state sovereignty merely as an "unauthorized exercise of authority by one state in another's state's territory."[22] If one adopts the former view, however, the issue becomes murkier, as the harmful effects of disinformation are

---

14. *Id.*

15. Harriet Moynihan, *The Application of International Law to State Cyberattacks*, CHATHAM HOUSE (Dec. 2, 2019), https://www.chathamhouse.org/2019/12/application-international-law-state-cyberattacks/1-introduction.

16. Harriet Moynihan, *The Application of International Law to Cyberspace: Sovereignty and Non-intervention*, JUST SECURITY (Dec. 13, 2019), https://www.justsecurity.org/67723/the-application-of-international-law-to-cyberspace-sovereignty-and-non-intervention/.

17. *Id.*

18. Nagasako, *supra* note 7, at 131.

19. Moynihan, *supra* note 15, at ¶ 20-21.

20. *Id.*

21. *Id.*

22. Moynihan, *supra* note 16.

often indirect and difficult to measure. Disinformation can constitute coercive behavior insofar as it is deployed to pressure a state to achieve a particular outcome. This can include, for example, campaigns that seek to manipulate the views of civilians so as to impede democratic debate or influence governmental functions.

The quandary in applying the principles of sovereignty and non-intervention to disinformation is an issue of measurement and attribution. Both technical and legal attribution of an unlawful act to a state actor—or a nonstate actor under the direction of the state—are a prerequisite for state responsibility.[23] Nevertheless, it can be difficult to identify a coordinated and intentional false information operation and it is often challenging to attribute such activities to an actor, particularly state actors.[24] The sponsors of disinformation actors tend to be obscure, and even when states can make an allegation regarding the state sponsorship of cyber campaigns, they may choose to withhold evidence from the public or other states for fear of disclosing intelligence methods.[25] Thus, absent concrete evidence of wrongdoing, states accused of engaging in disinformation campaigns that breach international law may balk at being held responsible.

Moreover, states are not forbidden from attempting to influence other states. Thus, the question becomes: at what point does disinformation move beyond influence to become an intrusion of sovereignty or coercion? While the dissemination of truly false information may "manipulate [a state's] capacity to reason," "distorted" information—the reframing or manipulation of true facts—is virtually indistinguishable from any ordinary use of "selected information and framing" to rationalize an argument.[26] While the former may constitute coercion, the latter probably does not.[27] The prolific broadcasting of cherry-picked but genuine information to sow division may be disingenuous,

---

23. *See* Lorraine Finlay & Christian Payne, *The Attribution Problem and Cyber Armed Attacks*, 113 Am. J. Int'l L. 202, 204 (2019) (describing how states injured by cyber-attacks must solve both the technical problem of attribution – "how to identify the true origin of a particular attack and the identity of those who carried it out" – and the legal problem of attribution – whether the facts allow "a state to be held responsible for the cyberattack under international law").

24. *Id.*

25. Fan Yang, *The Problem with Ill-Substantiated Public Cyber Attribution: A Legal Perspective*, CARNEGIE ENDOWMENT FOR INT'L PEACE (Mar. 28, 2022).

26. Björnstjern Baade, *Fake News and International Law*, 29 European J. Int'l L. 1357, 1364 (2019).

27. *Id.* (stating that "[m]ere framing and presentation of true facts cannot be held to be coercive in the sense required").

but it is likely legal.[28] There may be little remedy then for cases such as the Russian troll farms that targeted the Women's March, as much of the information promulgated was genuine.[29] Conversely, there is a stronger argument for coercion in cases involving the deliberate publication of false information to interfere in the domestic affairs of another state.

Historically, some states have shown support for formalizing state responsibility to refrain from the use of genuine information to interfere in the affairs of another state. General Assembly Resolution 36/103, for example, articulated a state duty "to abstain from any defamatory campaign, vilification or hostile propaganda for the purpose of intervening or interfering in the internal affairs of other" and forbade states from exploiting and distorting human rights issues "as a means of interference in the internal affairs of States, of exerting pressure on other States or creating distrust and disorder within and among States or groups of State."[30] While many Western states have objected to this resolution, its adoption by the majority of the General Assembly suggests some support for a wider view of coercion with regard to disinformation.[31] In addition, there have been indications more recently that the increased vulnerability of states to foreign cyber operations has shifted Western states' views of what constitutes coercion.[32]

III. DISINFORMATION UNDER INTERNATIONAL HUMAN RIGHTS LAW

States may also rely on international human rights law to regulate misinformation and disinformation. This permits States and individuals to respond to disinformation in a more targeted manner by focusing on its "adverse human rights impacts," rather than the information

28. *Id.* at 1365.

29. Ellen Barry, *How Russian Trolls Helped Keep the Women's March Out of Lock Step*, N.Y. TIMES (Sept. 18, 2022).

30. *Declaration on the Inadmissibility of Intervention and Interference in the Internal Affairs of States*, GA Res 36/103, UN Doc A/RES/36/103 (Dec. 9, 1981).

31. Hitoshi Nasu, *The 'Infodemic': Is International Law Ready to Combat Fake News in the Age of Information Disorder?*, AUSTRALIAN YEAR BOOK INT'L L. ONLINE (Dec. 9, 2021).

32. *See id.* (discussing comments from the former Attorney-General of the United Kingdom and the U.S. General Counsel of the Department of Defense that point to cyber operations that manipulate or interfere with another country's elections as "prohibited interventions").

itself.[33] Disinformation threatens several human rights articulated by the International Covenant on Civil and Political Rights (ICCPR) and the International Covenant on Economic, Social and Cultural Rights (ICESCR).[34] Disinformation can prevent voters from accessing accurate information about candidates and issues during elections, violating Art. 25 of ICCPR, the right to free and fair elections.[35] As an example, Donald Trump's false claims of election fraud impeded the free exercise of elections in the United States.[36]

The Covid-19 pandemic was rife with false information campaigns promoting inaccurate and potentially hazardous information about healthcare. These falsities led many civilians to avoid important preventative and curative treatments, instead driving them to pursue dangerous and ineffective alternative cures and violating Art. 12 of ICESCR, the right to health.[37] Disinformation operations can also foment hate and incite violence and discrimination towards individuals and groups, violating Art. 17 of ICCPR, the right to freedom from unlawful attacks upon one's honor and reputation, and Art. 26 of ICCPR, the right to non-discrimination.[38] Russian disinformation activities regarding the conflict in Ukraine, for instance, have relied heavily on antisemitic propaganda.[39]

Human rights obligations under international law may therefore provide another avenue for imposing legal responsibility for disinformation campaigns, particularly for information operations that do not target other states but focus on civilians instead.

It is also important to note that regulatory measures undertaken by states to combat disinformation may threaten the right of their citizens to freely engage in speech. States that are party to the ICCPR may restrict expression for "respect of the rights or reputations of others" or "for the protection of national security or of public order, or of

---

33. Richard Wingfield, *A Human Right-Based Approach to Disinformation*, GLOBAL PARTNERS DIGITAL (2019), https://www.gp-digital.org/a-human-rights-based-approach-to-disinformation/.

34. *International Covenant on Civil and Political Rights*, Dec. 16, 1966, 999 U.N.T.S. 171 [hereinafter *ICCPR*]; International Covenant of Economic, Social and Cultural Rights, Dec. 16, 1966, 999 U.N.T.S. 3 [hereinafter *ICESCR*].

35. *ICCPR, supra* note 34, at art. 25

36. Richard L. Hasen, Identifying and Minimizing the Risk of Election Subversion and Stolen Elections in the Contemporary United States, 135 Harv. L. Rev. 265 (2022).

37. *ICESCR*, *supra* note 34, at art. 12; Wingfield, *supra* note 33.

38. *ICCPR*, *supra* note 34, at arts. 17, 26.

39. Global Engagement Center, *To Vilify Ukraine, The Kremlin Resorts to Antisemitism*, U.S. DEP'T OF STATE (2022), https://www.state.gov/disarming-disinformation/to-vilify-ukraine-the-kremlin-resorts-to-antisemitism/.

public health or morals."[40] Limitations on expression must meet several conditions: they must not jeopardize the right to expression, they must be "provided by law," they may only be imposed for one of the legitimate exceptions set out in Art. 19, and they must "conform to the strict tests of necessity and proportionality."[41]

Moreover, identifying and removing false information is a Sisyphean task and attempts to corral disinformation may inhibit legitimate expression.[42] For example, several states have established government fact-checking units, but these departments often lack clear operating rules and powers.[43] Absent clear standards and methodologies for fact-checking, these units run the risk of intentionally or inadvertently censoring legitimate expression.[44] Legislation that forbids the publication of false or inaccurate information may also constitute an overly broad restriction of expression.[45] The U.N. Human Rights Committee (UNHRC) has repeatedly noted that regulations criminalizing disinformation, such as laws that permit journalists to be punished for publishing false news solely because the news is false, violate Art. 19 of ICCPR.[46] Requiring intermediaries, such as private corporations, to monitor and remove fake or misleading content poses similar risks.

IV. CONCLUSION

Reviewing the application of international law to disinformation reveals how gaps in the current framework create "legal grey zones" that can be exploited by international actors to engage in disinformation operations.[47] As acts that fall short of "use of force," disinformation campaigns can violate the principles of sovereignty and non-interference. Uncertainty over what constitutes coercion and practical difficulties in attribution and measurement, however, make application

---

40. *ICCPR*, *supra* note 34, at art. 19.

41. H.R.C., Gen. Comment No. 34, CCPR/C/GC/34, at ¶ 22 (Sept. 12, 2011).

42. Wingfield, *supra* note 33.

43. *Submission to UN Special Rapporteur for Freedom of Expression on an Annual Thematic Report on Disinformation*, CENTRE FOR LAW AND DEMOCRACY (Mar. 2021) (highlighting fact-checking units established in Mexico, Italy, India, Pakistan, and Argentina).

44. *Id.*

45. *Id.*

46. Andrey Rikhter, *International Law and Policy on Disinformation in the Context of Freedom of the Media*, OSCE (May 14, 2021) (listing several reports from UNHRC that "made it clear that criminalizing disinformation is inconsistent with the right to freedom of expression").

47. 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. 36 (2018).

of the law an abstract inquiry, rather than a concrete reality. Some have attempted to mitigate this uncertainty by calling for states to generate a comprehensive treaty, analogous to the United National Charter on the Law of the Seas (UNCLOS), governing behavior in cyberspace.[48] While there have been discussions on formulating international law with regards to cyberspace in the U.N., regional, and state bodies, a consensus on cyberspace regulation has yet to emerge.[49]

International human rights law provides another avenue for dealing with disinformation, but questions over the extraterritorial application of these rights, as well as inconsistent rules between nation-states governing the application of rights within states, may be a stumbling block for any practical use of the law. Moreover, many states' regulatory attempts to combat disinformation run afoul of human rights rules on their own merits.

In light of these difficulties, human rights organizations tend to offer solutions grounded in bolstering free expression. Per this understanding, "disinformation tends to thrive where human rights are constrained, where the public information regime is not robust and where media quality, diversity and independence is weak. Conversely, where freedom of opinion and expression is protected, civil society, journalists and others are able to challenge falsehoods and present alternative viewpoints."[50] Combatting disinformation thus requires facilitating the free exchange of information, allowing citizens to debate solutions to social problems, challenge propaganda and dispute false information, and hold fellow civilians and officials accountable for their speech.[51] Promoting transparency and access to data requirements will also allow for independent scrutiny, increasing accountability and trust.

The principles of sovereignty and non-intervention provide an attractive approach to disinformation by implying that there exists an avenue to impose consequences on malicious actors. However, as satisfying as this may appear, the difficulties in applying these principles limit their efficacy in combatting information. Until and if states can reach a consensus on the application of these principles, centering efforts to combat disinformation on facilitating access to legitimate and open expression may be the best, if only, remedy.

---

48. *Id.* at 51.

49. *Id.* at 51-3.

50. Irene Khan (Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression), *Disinformation and freedom of opinion and expression*, U.N. Doc. A/HRC/47/25 (Apr. 13, 2021).

51. *A Human Rights Approach to Tackle Disinformation: Submission to the Office of the High Commissioner for Human Rights*, AMNESTY INT'L (Apr. 14, 2022).