2021 13th International Conference on Cyber Conflict
Going Viral
T. Jančárková, L. Lindström, G. Visky, P. Zotz (Eds.)
2021 © NATO CCDCOE Publications, Tallinn

Windmills of the Mind: Higher-Order Forms of Disinformation in International Politics

James Shires
Assistant Professor, Cybersecurity Governance
Institute of Security and Global Affairs
University of Leiden
The Netherlands
j.shires@fgga.leidenuniv.nl

Abstract: Disinformation – the organised and deliberate circulation of verifiably false information – poses a clear danger to democratic processes and crisis response, including the current coronavirus pandemic. This paper argues for a conceptual step forward in disinformation studies, continuing a trend from the identification of specific pieces of disinformation to the investigation of wider influence campaigns and strategic narrative contestation. However, current work does not conceptually separate first-order forms of disinformation from higher-order forms of disinformation: essentially, the difference between disinformation about political or other events, and disinformation *about disinformation itself*.

This paper argues that this distinction is crucial to understanding the extent and consequences (or lack thereof) of disinformation in international politics. The paper first highlights how political disinformation is often sparked by leaks – the release of secret or confidential information into the public domain. It suggests that disinformation and leaks intersect with conventional cybersecurity threats through the increasingly common phenomenon of hack-and-leak operations. The paper then introduces the concept of higher-order disinformation. This discussion is followed by an empirical example: the case of US intelligence assessments of Russian hack-and-leak operations during the US presidential election campaign in 2016. The paper concludes with offensive and defensive policy implications, arguing that the relevance of second, third, and higher orders of disinformation will only increase as more experienced actors draw on the material, successes, and lessons of previous campaigns.

Keywords: *disinformation, hack-and-leak operations, leaks, Russia, US, narrative*

1. INTRODUCTION

Disinformation is an essentially social problem: in one useful definition, it is not simply misleading communication but communication that has *the central function of misleading* a specific audience.[1] However, it would be a mistake to see disinformation simply as a defect in systems of communication, whether written, verbal, or visual. Such an approach, common in both policy and academic literature on disinformation, draws on a simple "transmission" view of communication. It approaches political communities in an almost cybernetic fashion, focusing on the extent to which accurate information is transferred between different parts of the system.[2]

However, disinformation, along with a broader array of misdirection and deception, is not a secondary add-on to or corruption of pure information flows in an ideal body politic but an integral part of that political community. The community itself would not exist without the rumours, lies, and half-truths that circulate within it.[3] In this view, there is no such thing as "pure" – unbiased, not slanted, non-ideologically committed – communication against which to compare clear examples of disinformation. To continue the biological metaphor, just as bacteria are not an external, negative threat to biological organisms but a central part of their inner constitution, the same applies to societies and disinformation. Consequently, although the theme of this conference is "going viral", viruses – especially in the current pandemic times – are a misleading "organizing metaphor" for disinformation: a better one is bacterial.[4]

This is not a new insight, and most approaches to political science and international relations recognise that questions of truth and falsity cannot be answered without considering broader issues around discursive power and silence, and narrative construction and contest.[5] In studies of disinformation more specifically, this insight has encouraged a trend away from the identification of specific pieces of disinformation, to be countered by education and fact-checking, to the investigation of wider "influence" campaigns. Such campaigns are often identified and investigated along the lines of Advanced Persistent Threat (APT) methodologies in cybersecurity, notably demonstrated in high-profile "takedowns" by large platform companies – a point to which I return below.[6]

1 Alexander Lanoszka, "Disinformation in International Politics", *European Journal of International Security* 4, no. 2 (June 2019): 227–248, https://doi.org/10.1017/eis.2019.6.

2 For an example of this approach, see Bruce Schneier and Henry Farrell, "Common-Knowledge Attacks on Democracy" (Berkman Klein Center for Internet and Society, Harvard University, October 2018).

3 Sally Engle Merry, "Rethinking Gossip and Scandal", in *Toward a General Theory of Social Control: Fundamentals*, edited by Donald Black (Orlando and London: Academic Press, 1984), 271–302.

4 Jordan Branch, "What's in a Name? Metaphors and Cybersecurity", *International Organization* 75, no. 1 (2021): 39–70, https://doi.org/10.1017/S002081832000051X.

5 David Campbell, *Writing Security: United States Foreign Policy and the Politics of Identity* (Manchester: Manchester University Press, 1998); Ronald R. Krebs, *Narrative and the Making of US National Security* (Cambridge University Press, 2015).

6 See, e.g., Facebook, "February 2020 Coordinated Inauthentic Behavior Report", *About Facebook* (blog), March 2, 2020, https://about.fb.com/news/2020/03/february-cib-report/.

This trend has reached its most sophisticated and historically aware treatment in the concept of "active measures", the subject of Thomas Rid's recent book of the same name.[7] Active measures are more than disinformation campaigns: they are (usually or mostly covert) bureaucratic efforts to marshal the combination and spread of information (whether true, false, or somewhere in between) for specific strategic ends, persisting – in Rid's treatment of the Russian case – beyond the lifetime of organisations, technological infrastructures, and even entire political regimes. However, even Rid's exemplary work only identifies instances of first-order forms of disinformation but does not conceptually separate first-order forms from higher-order forms: essentially, the difference between disinformation about political or other events, and disinformation *about disinformation itself*.

In this paper, I argue that this distinction helps us to bridge the two approaches to disinformation above: on the one hand, a "transmission" view of communication in which specific pieces of information have verifiably factual or false content, and on the other hand, a recognition that all communication, especially of the political kind, takes place against a backdrop of powerful discursive presuppositions and broader narrative contest. By considering the reflexive quality of individual instances of disinformation, and their references back to and dependence upon prior contested claims, we can progress analytically from the former view to the latter, tracing how fundamental splits in worldview emerge, stacked upon a succession of divergent factual claims as well as different political commitments. Understanding higher orders of disinformation is thus crucial to understanding the extent and consequences (or lack thereof) of disinformation in international politics overall.

The paper is structured as follows. The following section narrows the focus of the paper from disinformation overall to a specific kind of influence operation – hack-and-leak operations – that, due to their intersection with conventional cybersecurity threats, are a key focus of US defence and cyber policy.[8] The third section uses hack-and-leak operations to introduce the concept of higher-order forms of disinformation, meaning that second-order disinformation is leaks about (alleged) hack-and-leak operations, third-order disinformation is leaks about those leaks, and so on. The fourth section applies this largely abstract discussion to the case of US intelligence assessments of Russian influence operations around the 2016 presidential election. The final section concludes, reflecting on both offensive and defensive policy implications of this paper: offensively, the problems in mounting counter-disinformation disinformation operations, and defensively, the limits of relying on content moderation and fact-checking services to police disinformation.

7 Thomas Rid, *Active Measures: The Secret History of Disinformation and Political Warfare* (New York: Profile Books, 2020).

8 Stephen G. Fogarty and Bryan N. Sparling, "Enabling the Army in an Era of Information Warfare", *Cyber Defense Review* 5, no. 2 (Summer 2020). See also US Department of Defense, "Summary: Department of Defense Cyber Strategy", Washington, DC, 2018, p. 1.

2. THE MECHANICS OF HACK-AND-LEAK OPERATIONS

Hack-and-leak operations are, as several scholars have argued, the pinnacle of disinformation operations: they combine a compromise of digital networks to obtain information (hack) with the release of that information for strategic effect (leak).[9] This is not a necessary combination: many hacks occur without compromised information ever coming to light, while many leaks occur through more mundane forms of access – although they are no less dependent on the broader communications ecosystem built around the internet.[10] The paradigm example of a hack-and-leak operation is the release of information gained from the Democratic National Committee (DNC) and related entities and people before the 2016 US election, attributed to Russian intelligence agencies (for the leak, specifically, the military Main Intelligence Directorate [GRU]) by subsequent US government investigations and many independent observers.[11] However, we should not let the impact of this incident on academic and policy research on disinformation leave us unable to see the wood for a single large tree: notable state-sponsored hack-and-leak operations have taken place against international sporting bodies (the World Anti-Doping Agency [WADA], the International Federation of Association Football [FIFA]), private entities in the US (Sony Pictures), and in other national contexts (Macronleaks, the 2019 UK election, and the Saudi cables), as well as in situations of destabilisation and conflict (for example, in many instances in Syria).[12]

Before considering the precise relationship between hack-and-leak operations and disinformation, it is instructive to briefly outline the conceptual mechanics of hack-and-leak operations from an analytical perspective, rather than that of the target or perpetrator. Basically, hack-and-leaks, like leaks more generally, function within larger constructions of privacy and/or secrecy.[13] Words and deeds must first be kept private, or at a national level, classified as secret, to then be leaked. The first conceptual building block of a hack-and-leak is thus the protection and limitation of information. Many analyses omit this element, missing how variations in social expectations of secrecy or technological means for achieving it affect the outcome of

9 Jaclyn Alexandra Kerr and Herbert Lin, "On Cyber-Enabled Information/Influence Warfare and Manipulation", SSRN, March 13, 2017; James Shires, "Hack-and-Leak Operations: Intrusion and Influence in the Gulf", *Journal of Cyber Policy* 4, no. 2 (2019): 235–256.

10 Ronald J. Deibert, *Reset: Reclaiming the Internet for Civil Society* (Toronto: House of Anansi Press, 2020).

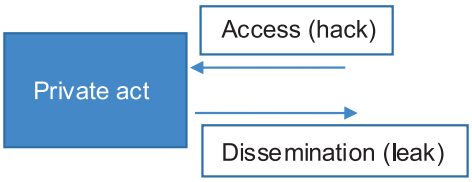
11 Robert S. Mueller, *Report on the Investigation into Russian Interference in the 2016 Presidential Election*, Submitted Pursuant to 28 C.F.R. § 600.8(c) (Washington, DC: US Department of Justice, March 2019).

12 Ben Buchanan, *The Hacker and the State: Cyber Attacks and the New Normal of Geopolitics* (Cambridge, MA: Harvard University Press, 2020); Jean-Baptiste Jeangene Vilmer, "The 'Macron Leaks' Operation: A Post-Mortem" (Atlantic Council and IRSEM, June 2019); James Shires, "Understanding the Tactics behind Hack-and-Leak Operations", *Atlantisch Perspectief* 4 (September 2020); Marie Baezner, "The Use of Cybertools in an Internationalized Civil War Context: Cyber Activities in the Syrian Conflict", CSS Cyber Defense Project (Center for Security Studies, ETH Zurich, October 18, 2017).

13 Sissela Bok, *Secrets: On the Ethics of Concealment and Revelation* (New York and Toronto: Vintage, 1989).

a leak.[14] Given such a private act, the two other key elements of a leak are *access* – an "outsider" gaining access to the private place – and *dissemination* – the spread of that information once obtained (Figure 1). Of course, both are spectrum rather than binary concepts: insider threats highlight the difficulty in access control, while leaked information rarely emerges into the open in a symmetric, equal fashion. Hack-and-leak operations, as a subset of leaks more broadly, can be defined as those involving a particular means of access: offensive cyber capabilities for remote intrusion into digital networks.

FIGURE 1: CONCEPTUAL MODEL OF HACK-AND-LEAK OPERATIONS



Hack-and-leak operations do not always include disinformation (although they are always "active measures", in Rid's definition of the term). However, it is precisely this expectation of privacy and/or secrecy that makes leaks such powerful vehicles for disinformation through doctoring or altering content. Leaks carry (often erroneous) connotations of franker, more truthful communication, without the many layers of artifice we expect from public political communication. Current scholarship has focused primarily on the *amplifying* relationship between hack-and-leaks and disinformation. Researchers have traced how what François calls "false leaks" spread on social media platforms like Twitter, highlighting how their dissemination through certain hashtags affects their impact.[15] Others have argued that "tainted leaks" of doctored information gained through phishing attacks against journalists and political opponents have been used by the Russian government to "seed mistrust".[16] As Rid demonstrates, these are not new tactics and existed well before the internet.[17] Elsewhere, I have argued that "edge cases" of hack-and-leaks, where almost all the released information is doctored – such as the cyber operation against the Qatar News Agency in 2017 – highlight the shifting boundaries between leaked and manufactured information.[18] Furthermore, the act that is the subject of a leak does not have to be documentary in form: the Shadow Brokers leaks highlight how offensive cyber capabilities can themselves be

14 David Pozen, "The Leaky Leviathan: Why the Government Condemns and Condones Unlawful Disclosures of Information", *Harvard Law Review* 127 (February 25, 2013): 512–635.
15 Presentation by Camille François of Graphika at CyberWarCon, Washington, DC, November 2018.
16 Adam Hulcoop et al., "Tainted Leaks: Disinformation and Phishing With a Russian Nexus", *Citizen Lab*, May 25, 2017.
17 Rid, *Active Measures*.
18 James Shires, "The Cyber Operation against Qatar News Agency", in *The 2017 Gulf Crisis: An Interdisciplinary Approach*, edited by Mahjoob Zweiri, M. Mizanur Rahman, and A. Kamal (Berlin and Heidelberg: Springer Nature, 2020).

the subject of (alleged) hack-and-leaks, introducing an entirely new level of damage from their release.[19]

This growing body of scholarship demonstrates how disinformation – be it doctoring, falsifying, forging, or tainting – changes the mechanics of hack-and-leak operations above, in terms of both access and dissemination. If a claimed hack-and-leak is in fact a disinformation operation, then no access to a private or secret space is required. Of course, successful tainting requires raw material, and successful forgeries are usually based on close knowledge of genuine documents, so good access is likely to increase the impact of a disinformation operation based on a "leak" – but it is not necessary. In terms of dissemination, the problem is no longer how to identify relevant information on the target networks and extract it undetected, but how to muddy the sourcing so it *appears* to the eventual audience that a hack-and-leak was a plausible originating point. A good example of such vague genesis is the appearance of controversial documents about National Health Service (NHS) funding shortly before the 2019 UK general election. They first appeared on Reddit and took a while to catch the attention of the media before ending up in the hands of the opposition leader, Jeremy Corbyn, in a national televised debate.[20]

Overall, hack-and-leak operations can be a potentially effective but highly complex vehicle for disinformation. At their most effective, they act as the "simulation of scandal", combining genuine leaked information with difficult-to-detect nuggets of disinformation to embarrass or discredit a target.[21] Such operations may remain undetected or misdescribed for years, and it is likely that the empirical record of hack-and-leak operations only captures a small percentage of the overall cases. But their complexity means that they have several potential pitfalls, not least the law of diminishing returns: frequent scandals mean that audiences may be inured to later leaks, especially if manipulation is commonplace enough that people no longer give greater credence to leaked material. Furthermore, as I have argued elsewhere, hack-and-leak operations often backfire, because media attention and cyber "hype" mean that hacks are as newsworthy as leaked content, if not more so – especially when state-sponsored.[22] However, despite this complexity of effect, the basic mechanics of hack-and-leak operations – access to and dissemination of a private act – are relatively simple. This, in addition to their inclusion in US and other policy priorities, makes them a good focal point for the introduction of higher orders of disinformation in the following section.

19 Buchanan, *The Hacker and the State*.
20 Ben Nimmo et al., "Secondary Infektion", *Graphika*, June 2020.
21 James Shires, "The Simulation of Scandal: Hack-and-Leak Operations, the Gulf States, and U.S. Politics", *Texas National Security Review*, August 2020.
22 Shires, "The Simulation of Scandal".

3. HIGHER-ORDER FORMS OF DISINFORMATION

The concept of higher orders, in abstract terms, is relatively straightforward. For any x, a higher order x is reflexive: it is the application of x to x itself (second order), or the application of x to the application of x to x (third order), and so on. The concept has been used across philosophy and the social sciences, being deployed in the context of everything from conscious awareness (a mental state about a mental state) to the modelling of rational interaction in economics and political science (from x's beliefs about y, to y's beliefs about x's beliefs about y, and so on). Even in these examples, the power of the concept of higher orders should be apparent: it can account for the transition from a simple, single-level phenomenon, to multi-level, complex phenomena, without invoking more and more different types of entities or concepts: the reflexive repetition of a single concept is sufficient to explain the difference in complexity.

Disinformation creates a dilemma that seems – on the face of it – to call for a higher-order conceptual architecture. On the one hand, focusing on the "verifiably false" nature of specific claims leads quickly onto thorny ground.[23] For example, the EU External Action Service (EEAS), an EU agency founded in 2010, has an East StratCom Task Force, established in 2015, which seeks to "increase public awareness and understanding of the Kremlin's disinformation operations".[24] To do this, the EEAS runs a website, *EUvsDisinfo* (euvsdisinfo.eu), with a well-populated "disinfo database" of specific pieces of disinformation archived from media websites in multiple languages, with date, target audience, and other key characteristics. Each piece includes a summary and a "disproof", a body of text that contradicts or debunks the claims made by the disinformation piece. However, in many cases the "disproof" is not exactly that, because the piece of disinformation itself was not precise enough to be debunked. Instead, the "disproof" offers a contrasting narrative, drawing on wider geopolitical statements that, crucially, do not represent a shared ground of agreement (for example, between pro- and anti-Russian government positions). This sustained and careful project, focusing on specific pieces of disinformation, runs aground because it is easily drawn into wider contests over frames and narratives.

On the other hand, as highlighted in the introduction, many recent analyses do not interrogate the "verifiably false" nature of specific claims[25] but instead reorient the debate using terms such as "influence campaigns" or some platform companies' preferred term of "Coordinated Inauthentic Behavior" (CIB). The label of "influence campaigns" echoes the cybersecurity industry's shift away from solely detecting specific cybersecurity incidents or events (analogous to a particular instance of

23 European Commission, "A Multi-Dimensional Approach to Disinformation: Report of the Independent High Level Group on Fake News and Online Disinformation" (Luxembourg: European Commission Directorate-General for Communication Networks, Content and Technology, March 2018).
24 EUvsDisinfo, https://euvsdisinfo.eu/about/
25 European Commission, "A Multi-Dimensional Approach to Disinformation".

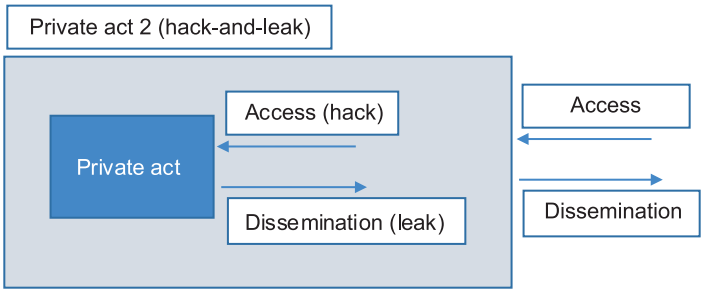
disinformation) to connecting such incidents together as intrusion campaigns according to common tactics, techniques, and procedures (TTPs) and broader strategic objectives. It is more than an echo, in fact, as cybersecurity professional experience, commercial structures, and even specific APT labels can thus be transferred from intrusion campaigns to the problem of disinformation. Rid's concept of active measures represents the apex of this trend, focusing on the strategic and bureaucratic practices and ideologies underpinning a wide range of campaigns. These concepts are each significantly different, but they all operate at a far more sophisticated level than approaches seeking simply to "disprove" disinformation.

The question, then, is: how can we connect these two approaches to disinformation? I suggest that we can understand how disinformation expands into wider differences in frame and narrative using the concept of higher orders introduced above. Such an analysis begins by identifying key informational nodes that fracture audience perspectives, perceived by some as central factual elements of their overall worldview, and as disinformation by others.[26] Such nodes are the basis for further contentious claims, which revolve around the credibility of earlier nodes. These subsequent claims are, for those who disagree with that interpretation of the informational node, a second-order form of disinformation: disinformation *about disinformation*. These claims in turn invite further claims: third-order or higher forms of disinformation.

The case study in the following section applies this approach to a specific case study; before doing so, I illustrate the approach in more detail using the framework of hack-and-leak operations introduced above. A higher-order treatment of hack-and-leak operations would use only the concepts identified in the previous section (a private or secret act, and access to and dissemination of that act). More specifically, to explain how the hack-and-leak itself becomes the subject of media attention, we can see the hack-and-leak as a second private act, encompassing the original private act (the subject of the hack-and-leak) as well as the access to and dissemination of information. Consequently, this second private act (the whole hack-and-leak) can itself be subject to access (discovering the hack) and dissemination (informing the media that the original scandal was the result of a hack-and-leak). This reflexive step is illustrated in Figure 2.

26 This is always a *further* fracturing: there is no single original audience and no cohesive public sphere prior to such disagreements.

FIGURE 2: CONCEPTUAL MODEL OF SECOND-ORDER HACK-AND-LEAK



This reflexive application of the same concepts is useful because it explains more complex cases without resorting to a larger conceptual architecture. In cases where the hack becomes as newsworthy a story as the leak (for example, in the case of Russian intrusion into the DNC in 2016), access to the hack-and-leak operation – through Crowdstrike's technical analysis, the Mueller investigation, and many other means – and its dissemination – the Mueller report, congressional testimony, countless media articles, and many other publications – have turned the hack-and-leak into a private act to be revealed to the public in just the way that the original private act (confidential emails and documents) were revealed to the public by the GRU via Wikileaks. Furthermore, this is only the first step in the application of higher-order concepts: as illustrated below, third- and subsequent-order versions quickly emerge.

4. "RUSSIAN INTERFERENCE" VS "RUSSIA HOAX"

This section examines one aspect of the most high-profile example of an influence campaign in recent history: Russian activities relating to the 2016 US presidential election. As noted above, these activities were reported to include – but were not limited to – a hack-and-leak operation against the DNC and related entities. This hack-and-leak operation acted as a key informational node, morphing into two far broader narratives in US politics. One was an anti-Trump narrative of "Russian interference", taking forensic evidence around the DNC compromise and the subsequent Mueller investigation at face value. The other was a pro-Trump narrative of a "Russia hoax", propagated by President Donald J. Trump himself, his family and close associates, right-wing media outlets, and social media commentators.[27] The "Russia hoax" narrative claims that the DNC hack-and-leak operation and wider claims of links between the Trump campaign and the Russian government were part of a deliberate plan to sabotage the Trump campaign and then the presidency itself.

27 I use pro- and anti-Trump as the most accurate way of designating US political divisions during the 2016–2020 term, rather than Republican/Democrat or left/right-wing.

This case exemplifies the disinformation dilemma I identified above: the movement from specific document leaks to their embedding in larger narratives and frames. It should be stressed that there are long-term reasons for this split in US political worldviews (not least the decade-long evolution of the right-wing media ecosystem),[28] and the key informational node of the hack-and-leak operation fractured these perspectives *further* rather than beginning the process.[29] Nonetheless, its divisive nature and subsequent policy impact make it a crucial case for the conceptual framework of higher-order forms of disinformation introduced above. To focus this brief account, I centre the following discussion on the declassification of documents relating to Russian activities and the 2016 election that occurred at the end of September 2020, ordered by then-Director of National Intelligence (DNI) John Radcliffe. Radcliffe is a prominent Republican and was a member of Congress until his appointment as DNI in May 2020 by President Trump.

This declassification is important for three reasons. The first was its timing: the declassification occurred at a key point in President Trump's run for re-election, and many commentators claimed it was designed specifically to influence the 2020 campaign. Second, the declassification was itself a leak, insofar as it generated significant controversy within former and current members of the intelligence community about whether it conformed to standard practices of declassification, or even specific regulations.[30] Third, the declassified documents connect individual reports from 2016 to the overall split in narratives above, with both sides claiming that the declassified documents support the "Russian interference" and the "Russia hoax" narrative, respectively.[31]

On 29 September 2020, DNI Radcliffe declassified three US intelligence documents, the first two of which were released by Fox News.[32] The first document contained handwritten notes by then-CIA Director John Brennan from a meeting with President Obama in late July 2016, concerning "alleged approval by Hillary Clinton on 28 July of a proposal from one of her foreign policy advisers to vilify Donald Trump by stirring up a scandal claiming interference by the Russian security service".[33] The second document was a CIA memo to the FBI on 7 September 2016, providing "examples of information the Crossfire Hurricane [investigation into Russia links to the Trump

28 Yochai Benkler, Robert Faris, and Hal Roberts, *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (New York: Oxford University Press, 2018).

29 Kathleen Hall Jamieson, *Cyberwar: How Russian Hackers and Trolls Helped Elect a President: What We Don't, Can't, and Do Know* (New York: OUP USA, 2018).

30 Brian Greer, "John Ratcliffe's Dangerous Declassification Game", *Lawfare*, October 7, 2020, https://www.lawfareblog.com/john-ratcliffes-dangerous-declassification-game.

31 Andrew Desiderio and Daniel Lippman, "Intel Chief Releases Russian Disinfo on Hillary Clinton That Was Rejected by Bipartisan Senate Panel", *Politico*, September 29, 2020, https://www.politico.com/news/2020/09/29/john-ratcliffe-hillary-clinton-russia-423022.

32 Brooke Singman, "DNI Declassifies Brennan Notes, CIA Memo on Hillary Clinton 'Stirring up' Scandal between Trump, Russia", Fox News, October 6, 2020, https://www.foxnews.com/politics/dni-brennan-notes-cia-memo-clinton.

33 Ibid.

campaign] fusion cell has gleaned", including "an exchange [redacted] discussing US presidential candidate Hillary Clinton's approval of a plan concerning US presidential candidate Donald Trump and Russian hackers hampering US elections as a means of distracting the public from her use of a private email server".[34] The third document (not released directly) stated that in late July 2016, US agencies "obtained insight into Russian intelligence analysis alleging" that Clinton "approved a campaign plan to stir up a scandal" against Trump "by tying him to Putin and the Russians' hacking of the DNC".[35]

The media and political response to these documents in the US was extremely polarised, across both traditional and social media. Pro-Trump observers took this declassification in the way it was likely intended by the DNI, seeing it as evidence that the US intelligence community knew of improper practices by the Clinton campaign and yet did not follow them up, thus falling for what these observers saw as the "Russia hoax".[36] By contrast, anti-Trump observers largely focused on the role of Russian intelligence analysis as the source of the alleged Clinton plans in these documents, highlighting the declassified sentence that "the IC [intelligence community] does not know the accuracy of this allegation or the extent to which Russian intelligence analysis may reflect exaggeration or fabrication" as a basis for claiming that the declassification was based on "Russian disinformation".[37] Some anti-Trump commentators went further, claiming not only that Russian disinformation was the source of the documents but that the declassification was therefore itself a form of disinformation, as an inappropriate declassification (i.e., a leak) based on false information and designed to mislead.[38] The declassification event clearly resists a neat analytical interpretation, with almost any treatment likely to lean towards one or the other of the two broader narratives of "Russian interference" or "Russia hoax".[39]

The conceptual model of higher-order disinformation can help analysts trace how these two broader narratives relate to the specific declassified documents and the hack-and-leak operation that is their subject, connecting the two levels of analysis identified earlier. Unlike the second-order model presented in the previous section, this case exhibits at least five orders of reflexivity: (5) the DNI's declassification

34 Ibid.

35 Ibid.

36 Jerry Dunleavy, "Obama Was Briefed on Unverified Russian Report Claiming Clinton Approved Plan to Tie Trump to Putin and DNC Hack", *Washington Examiner*, September 29, 2020, https://www.washingtonexaminer.com/news/obama-was-briefed-on-unverified-russian-report-claiming-clinton-approved-plan-to-tie-trump-to-putin-and-dnc-hack.

37 Sonam Sheth, "Trump's Spy Chief Just Released 'Russian Disinformation' against Hillary Clinton that He Acknowledged May Be Fabricated", *Business Insider*, September 30, 2020, https://www.businessinsider.in/politics/world/news/trumps-spy-chief-just-released-russian-disinformation-against-hillary-clinton-that-he-acknowledged-may-be-fabricated/articleshow/78396299.cms.

38 Zachary Cohen and Alex Marquardt, "Former CIA Director Accuses Intel Chief of Selectively Declassifying Documents to Help Trump", CNN, October 7, 2020, https://www.cnn.com/2020/10/06/politics/brennan-ratcliffe-declassifying-intelligence-clinton-russia/index.html.

39 This includes the analysis here, which, one reviewer noted, could be construed as "an attack on right-wing politics".

or dissemination of (4) US intelligence community documents about (3) Russian intelligence analysis about (2) an alleged plan by the Clinton campaign to tie together Trump and Russia in the (1) hack-and-leak of documents from the DNC. These are orders of reflexivity, rather than separate events, because they each revolve around the same claims, piling extra layers of interpretation on at each stage. Crucially, each higher order introduces further potential for disinformation, as the key elements of access and dissemination are questioned at each stage: first, for the alleged plan described by Russian intelligence (which could be exaggerated or fabricated), then the report about the Russian intelligence analysis (which is dependent on US intelligence collection of uncertain reliability), then the original context of the declassified documents (as hard evidence or simply examples of leads), and then the intention and appropriateness of the declassification itself. Unpacking each order of disinformation, their specific means of access and dissemination, and the associated possibilities for falsification and contestation reveals how broader narratives are dependent on the compilation of contested claim upon contested claim, stacking these claims into worldviews that have begun to rupture the US political system.

5. CONCLUSION

Round like a circle in a spiral, like a wheel within a wheel
Never ending or beginning on an ever-spinning reel

Alan and Marilyn Bergman, "The Windmills of Your Mind" (1968)

The relevance of second, third, and higher orders of disinformation will only increase as more experienced actors draw on the material, successes, and lessons of previous campaigns to construct new material. As Rid has demonstrated, Soviet active measures drew extensively on earlier controversies, even to the point of resurrecting previously debunked forgeries decades later.[40] We can expect this dynamic to play out on social media platforms and the internet, as what quickly become historic struggles over the factual record transform into the foundations of future narrative contestation. One of the most striking qualities of these ever-growing chains of disinformation is their reflexive nature – hence the resurrection of a popular 1960s song in the title of this paper and the quotation above. Indeed, the indirect consequences of these chains of higher-order disinformation, fracturing worldviews and exacerbating political polarisation, may themselves be a desired strategic effect of such operations.

More concretely, the specific policy implications of higher-order forms of disinformation can be divided into two kinds: defensive and offensive. Defensively, this approach reinforces scholarship indicating the limited utility of fact-checking services

40 Rid, *Active Measures*.

in countering disinformation. While these services certainly have an important role to play in a turbulent communications ecosystem, many commentators have argued that they are unable to address broader narrative contestation – as in the example of the EEAS website considered earlier. We can now see why: to do so, fact-checking services would have to reverse-engineer multiple orders of disinformation, a time-consuming, resource-intensive process to say the least – and one likely to introduce its own biases. More problematically, a focus on higher orders of disinformation highlights that fact-checking services themselves are an attractive target for disinformation. In Tunisia, an Israeli PR company set up a fake fact-checking service before local elections, while in the UK the Conservative Party renamed its Twitter account "Factcheck UK" in the run-up to the 2019 election.[41] A growing wave of investigative journalism seeks to peel away such layers of misdirection – especially organisations such as Bellingcat, whose use of leaks has itself attracted some controversy – and so further analysis of the exploitation of fact-checking services would be a natural extension of the conceptual approach developed in this paper.

Offensively, the greater the salience of disinformation in international politics, the more all states – and other actors – will employ not just accusations of disinformation but also influence campaigns as a response to unwelcome international attention. Recent events illustrating this trend include China's response to the UK's withdrawal of the media license for a Chinese state-owned channel in February 2021. Chinese statements announced a reciprocal ban against BBC World News. The same statements denounced as "false information" the BBC's investigations of severe human rights violations against Uighurs in Xinjiang province (which in turn relied on leaked documents as well as interviews). This denunciation was backed up by a tightly coordinated influence campaign by government-linked accounts on Twitter.[42] In the same month, Saudi Arabia's furious response to the Biden administration's release of a report on the killing of Jamal Khashoggi not only branded the report itself as a disinformation operation but reinforced this message on Twitter using a network of bots like those Khashoggi worked against before his death.[43]

This response option is not limited to authoritarian states. Many militaries and intelligence agencies – including in the US and other NATO states – are openly considering more active responses to disinformation by adversaries along the lines of

41 Andy Carvin et al., "Operation Carthage: How a Tunisian Company Conducted Influence Operations in African Presidential Elections", *Atlantic Council*, June 5, 2020, https://perma.cc/AEY3-R3XU; Hannah Murphy and Alex Barker, "Conservative Party's 'FactcheckUK' Twitter Stunt Backfires", November 19, 2019, https://www.ft.com/content/0582a0d0-0b1f-11ea-b2d6-9bf4d1957a67.

42 Patrick Wintour, "China Bans BBC World News in Retaliation for UK Licence Blow", *Guardian*, February 11, 2021, http://www.theguardian.com/world/2021/feb/11/china-bans-bbc-world-news; Jacob Wallis and Albert Zhang, "Trigger Warning: The CCP's Coordinated Information Effort to Discredit the BBC" (Canberra: Australian Strategic Policy Institute, March 4, 2021), https://www.aspi.org.au/report/trigger-warning.

43 Craig Timberg and Sarah Dadouch, "When U.S. Blamed Saudi Crown Prince for Role in Khashoggi Killing, Fake Twitter Accounts Went to War", *Washington Post*, March 2, 2021, https://www.washingtonpost.com/technology/2021/03/02/saudi-khashoggi-twitter-mbs/.

those developed for other malicious actors in cybersecurity.[44] Such "counter-cyber" responses to disinformation include disrupting technical and social infrastructure, as reportedly occurred for the Internet Research Agency, a Russian "troll farm", before the US 2018 mid-term elections.[45] They also include clandestine social media campaigns, such as the one targeting Russia in the Sahel in December 2020, attributed by Facebook to the French military.[46] But this spectrum of responses also includes leaking adversaries' identities, tactics, and plans and (although this is not publicly stated) potentially including falsified or doctored information in these leaks. The utility of these operations must be evaluated carefully, not just in terms of the operations themselves as second- or third-order forms of disinformation, but also in terms of the potential for blowback – for the operations to be exposed by adversaries and incorporated into even higher order forms of disinformation.

In sum, this paper has argued that narrative contests involving repeated, escalating, and – crucially – *reflexive* accusations of disinformation on both sides are the norm rather the exception in international politics. The concept of higher orders of disinformation helps us to gain analytical purchase on such chains of successive and deeply disputed claims, and so – in a small way – contributes to the accurate diagnosis, and eventual amelioration, of a perennial problem.

44 Max Smeets, "U.S. Cyber Strategy of Persistent Engagement and Defend Forward: Implications for the Alliance and Intelligence Collection", *Intelligence and National Security* (February 15, 2020): 1–10, https://doi.org/10.1080/02684527.2020.1729316.

45 Catalin Cimpanu, "US Wiped Hard Drives at Russia's 'Troll Factory' in Last Year's Hack", *ZDNet*, February 28, 2019, https://perma.cc/763D-CEAY.

46 Nathaniel Gleicher and David Agranovich, "Removing Coordinated Inauthentic Behavior from France and Russia", *About Facebook*, December 15, 2020, https://about.fb.com/news/2020/12/removing-coordinated-inauthentic-behavior-france-russia/.

REFERENCES

Baezner, Marie. "The Use of Cybertools in an Internationalized Civil War Context: Cyber Activities in the Syrian Conflict". CSS Cyber Defense Project. Center for Security Studies, ETH Zurich, October 18, 2017.

Benkler, Yochai, Robert Faris, and Hal Roberts. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. New York: Oxford University Press, 2018.

Bok, Sissela. *Secrets: On the Ethics of Concealment and Revelation*. New York and Toronto: Vintage, 1989.

Branch, Jordan. "What's in a Name? Metaphors and Cybersecurity". *International Organization* 75, no. 1 (2021): 39–70. https://doi.org/10.1017/S002081832000051X.

Buchanan, Ben. *The Hacker and the State: Cyber Attacks and the New Normal of Geopolitics*. Cambridge, MA: Harvard University Press, 2020.

Campbell, David. *Writing Security: United States Foreign Policy and the Politics of Identity*. Manchester: Manchester University Press, 1998.

Carvin, Andy, Luiza Bandeira, Graham Brookie, Iain Robertson, Nika Aleksejeva, Alyssa Kann, Kanishk Karan et al. "Operation Carthage: How a Tunisian Company Conducted Influence Operations in African Presidential Elections". *Atlantic Council*, June 5, 2020. https://perma.cc/AEY3-R3XU.

Cimpanu, Catalin. "US Wiped Hard Drives at Russia's 'Troll Factory' in Last Year's Hack". *ZDNet*, February 28, 2019. https://perma.cc/763D-CEAY.

Cohen, Zachary, and Alex Marquardt. "Former CIA Director Accuses Intel Chief of Selectively Declassifying Documents to Help Trump". CNN, October 7, 2020. https://www.cnn.com/2020/10/06/politics/brennan-ratcliffe-declassifying-intelligence-clinton-russia/index.html.

Deibert, Ronald J. *Reset: Reclaiming the Internet for Civil Society*. Toronto: House of Anansi Press, 2020.

Desiderio, Andrew, and Daniel Lippman. "Intel Chief Releases Russian Disinfo on Hillary Clinton That Was Rejected by Bipartisan Senate Panel". *Politico*, September 29, 2020. https://www.politico.com/news/2020/09/29/john-ratcliffe-hillary-clinton-russia-423022.

Dunleavy, Jerry. "Obama Was Briefed on Unverified Russian Report Claiming Clinton Approved Plan to Tie Trump to Putin and DNC Hack". *Washington Examiner*, September 29, 2020. https://www.washingtonexaminer.com/news/obama-was-briefed-on-unverified-russian-report-claiming-clinton-approved-plan-to-tie-trump-to-putin-and-dnc-hack.

European Commission. "A Multi-Dimensional Approach to Disinformation: Report of the Independent High Level Group on Fake News and Online Disinformation". Luxembourg: European Commission Directorate-General for Communication Networks, Content and Technology, March 2018.

Facebook. "February 2020 Coordinated Inauthentic Behavior Report". *About Facebook*, March 2, 2020. https://about.fb.com/news/2020/03/february-cib-report/.

Fogarty, Stephen G., and Bryan N. Sparling. "Enabling the Army in an Era of Information Warfare". *Cyber Defense Review* 5, no. 2 (Summer 2020): 17–26.

Gleicher, Nathaniel, and David Agranovich. "Removing Coordinated Inauthentic Behavior from France and Russia". *About Facebook*, December 15, 2020. https://about.fb.com/news/2020/12/removing-coordinated-inauthentic-behavior-france-russia/.

Greer, Brian. "John Ratcliffe's Dangerous Declassification Game". *Lawfare*, October 7, 2020. https://www.lawfareblog.com/john-ratcliffes-dangerous-declassification-game.

Hulcoop, Adam, John Scott-Railton, Peter Tanchak, Matt Brooks, and Ronald J. Deibert. "Tainted Leaks: Disinformation and Phishing With a Russian Nexus". *Citizen Lab*, May 25, 2017.

Jamieson, Kathleen Hall. *Cyberwar: How Russian Hackers and Trolls Helped Elect a President – What We Don't, Can't, and Do Know*. New York: OUP USA, 2018.

Kerr, Jaclyn Alexandra, and Herbert Lin. "On Cyber-Enabled Information/Influence Warfare and Manipulation". *SSRN*, March 13, 2017.

Krebs, Ronald R. *Narrative and the Making of US National Security*. Cambridge University Press, 2015.

Lanoszka, Alexander. "Disinformation in International Politics". *European Journal of International Security* 4, no. 2 (2019): 227–248. https://doi.org/10.1017/eis.2019.6.

Merry, Sally Engle. "Rethinking Gossip and Scandal". In *Toward a General Theory of Social Control: Fundamentals*, edited by Donald Black, 271–302. Orlando and London: Academic Press, 1984.

Mueller, Robert S. *Report on the Investigation into Russian Interference in the 2016 Presidential Election*. Submitted Pursuant to 28 C.F.R. § 600.8(c). Washington, DC: US Department of Justice, March 2019.

Murphy, Hannah, and Alex Barker. "Conservative Party's 'FactcheckUK' Twitter Stunt Backfires", November 19, 2019. https://www.ft.com/content/0582a0d0-0b1f-11ea-b2d6-9bf4d1957a67.

Nimmo, Ben, Camille Francois, C. Shawn Eib, Lea Ronzaud, Rodrigo Ferreira, Chris Hernon, and Tim Kostelancik. "Secondary Infektion". *Graphika*, June 2020.

Pozen, David. "The Leaky Leviathan: Why the Government Condemns and Condones Unlawful Disclosures of Information". *Harvard Law Review* 127 (2013): 512–635.

Rid, Thomas. *Active Measures: The Secret History of Disinformation and Political Warfare*. New York: Profile Books, 2020.

Schneier, Bruce, and Henry Farrell. "Common-Knowledge Attacks on Democracy". Berkman Klein Center for Internet and Society, Harvard University, October 2018.

Sheth, Sonam. "Trump's Spy Chief Just Released 'Russian Disinformation' against Hillary Clinton that He Acknowledged May Be Fabricated". *Business Insider*, September 30, 2020. https://www.businessinsider.in/politics/world/news/trumps-spy-chief-just-released-russian-disinformation-against-hillary-clinton-that-he-acknowledged-may-be-fabricated/articleshow/78396299.cms.

Shires, James. "Hack-and-Leak Operations: Intrusion and Influence in the Gulf". *Journal of Cyber Policy* 4, no. 2 (2019): 235–256.

———. "The Cyber Operation against Qatar News Agency". In *The 2017 Gulf Crisis: An Interdisciplinary Approach*, edited by Mahjoob Zweiri, M. Mizanur Rahman, and A. Kamal. Berlin and Heidelberg: Springer Nature, 2020.

———. "The Simulation of Scandal: Hack-and-Leak Operations, the Gulf States, and U.S. Politics". *Texas National Security Review*, August 2020.

———. "Understanding the Tactics behind Hack-and-Leak Operations". *Atlantisch Perspectief* 4 (September 2020).

Singman, Brooke. "DNI Declassifies Brennan Notes, CIA Memo on Hillary Clinton 'Stirring up' Scandal between Trump, Russia". Fox News, October 6, 2020. https://www.foxnews.com/politics/dni-brennan-notes-cia-memo-clinton.

Smeets, Max. "U.S. Cyber Strategy of Persistent Engagement and Defend Forward: Implications for the Alliance and Intelligence Collection". *Intelligence and National Security* (February 15, 2020): 1–10. https://doi.org/10.1080/02684527.2020.1729316.

Timberg, Craig, and Sarah Dadouch. “When U.S. Blamed Saudi Crown Prince for Role in Khashoggi Killing, Fake Twitter Accounts Went to War”. *Washington Post*, March 2, 2021. https://www.washingtonpost.com/technology/2021/03/02/saudi-khashoggi-twitter-mbs/.

US Department of Defense. “Summary: Department of Defense Cyber Strategy”. Washington, DC, 2018.

Vilmer, Jean-Baptiste Jeangene. “The ‘Macron Leaks’ Operation: A Post-Mortem”. Atlantic Council and IRSEM, June 2019.

Wallis, Jacob, and Albert Zhang. “Trigger Warning: The CCP’s Coordinated Information Effort to Discredit the BBC”. Canberra: Australian Strategic Policy Institute, March 4, 2021. https://www.aspi.org.au/report/trigger-warning.

Wintour, Patrick. “China Bans BBC World News in Retaliation for UK Licence Blow”. *Guardian*, February 11, 2021. http://www.theguardian.com/world/2021/feb/11/china-bans-bbc-world-news.