

Digital Propaganda, Counterpublics and the Disruption of the Public Sphere: The Finnish Approach to Building Digital Resilience

Corneliu Bjola
University of Oxford

Krysianna Papadakis *University of Oxford*

Abstract:

Drawing on the case of Finland and its model of ‘resilience’ to digital disinformation, the article provides an original framework for understanding the logic of digital propaganda and the conditions that may facilitate an effective response. Building on recent research connecting digital propaganda to the disruption of the public spheres in democratic societies, the study argues that the logic of digital propaganda could be unpacked by examining the mechanisms that facilitate the diffusion of disinformation from the microsphere of the daily individual experience to the macrosphere of political decision-making. It is thus argued that the connection between the two spheres is enabled by the formation and political mobilization of ‘unruly’ counterpublics, that is, of arenas of textual and visual contestation of politically marginalized groups promoting issues aligned with the disinformation agenda. Using this framework, the study shows that Finland’s resilience progress has mostly occurred on the macrosphere level, where effective institutions have been applying transparent and proactive policies grounded in collaboration and research. However, these efforts are at risk of being weakened by the rise of influential counterpublics unless Finland takes further measures to protect its public sphere by reducing divisions in the microsphere.

Keywords: digital propaganda, resilience, public sphere, counterpublics

INTRODUCTION

“The point of modern propaganda - the former world chess champion, Garry Kasparov, and outspoken critic of President Putin, tweeted shortly after the U.S. Presidential elections in 2016 - , “isn't only to misinform or push an agenda. It is to exhaust your critical thinking, to annihilate truth” (Kasparov, 2016). Interestingly, Kasparov’s tweet struck a chord with the online public and went viral, suggesting that the message has captured a qualitatively important feature concerning the goal and supporting mechanisms by which propaganda is performed in the Digital Age.



In the past, propaganda mainly served to (covertly) manipulate the views and attitudes of the target group in a pre-determined direction (Cull, Culbert, & Welch, 2003, p. xix; Lasswell, 1927). The digital version that has been deployed against Western democracies by pro-Kremlin networks in the past few years, particularly since the 2014 annexation of Crimea (Mejias & Vokuev, 2017), works differently. It is primarily interested in the destruction of the epistemological foundation that informs the process of validation and adjudication of claims of political relevance for the constitution of contemporary societies (Pomerantsev, 2015; Pomerantsev & Weiss, 2014; Ramsay & Robertshaw, 2019). In other words, the goal of the ‘new propaganda’, at least in form that has been used by the Russian government, is not primarily to alter people’s views in support of certain policies, but to induce them into a state

of self-defeating and endemic scepticism by undermining the very criteria on the basis of which they develop their cognitive abilities to make sense, interpret and shape social reality.

The fact that Russia has decided to pursue such a course of action could perhaps make sense from a broader strategic perspective, as asymmetrical measures do carry tactical weight for those subscribing to a *real politik* vision of world politics where states engage in (zero-sum) competition for power and resources (Pynnöniemi, 2019; U.S. Senate Committee on Foreign Relations, 2018). To be sure, the Russian strategy of using disinformation as a foreign policy tool is not particularly new (Leighton, 1991). During the Cold War, the Soviet Union perfected the art of compromising, undermining or derailing policies pursued by Western countries via carefully tailored public disinformation campaign (Sorrels, 1983; US Department of State, 1989), including the infamous AIDS libel, “Operation Infektion” (Boghardt, 2009).

The more intriguing question is, of course, why Western societies with their strong record of democratic development may suddenly feel vulnerable to such nefarious forms of digital intrusion? As the disruptive effect of digital disinformation looms large in the minds of many in the context of the recurring controversies surrounding the results of the U.S. presidential elections and of the Brexit referendum in 2016, the message is hard to miss: Western countries face an unprecedented, systematic, and unrelenting disinformation assault on their cyberspace, mainly from Russia, which has been enabled by the very digital technologies that Western countries have created. And, even more troublingly, they have no clear understanding of how to protect themselves against it.

Recent studies suggest that social vulnerability to digital propaganda is closely connected to the disruption of the public spheres in democratic societies. Bennett and Livingston argue, for instance, that the breakdown of trust in democratic institutions of press and politics, along with

educational and civil society institutions in more advanced cases, is a key factor behind the emergence of the current ‘disinformation order’. Declining public confidence in democratic institutions undermines the credibility of official information in the news and thus opens publics to alternative information sources (Bennett & Livingston, 2018, pp. 126–127). Benkler, Faris and Roberts also focus on the transformation of the (American) public sphere and the closure of the right-wing media ecosystem to explain the mechanisms by which this new public sphere has become so “vulnerable to disinformation, propaganda, and just sheer bullshit” (Benkler, Farris, & Roberts, 2018, p. 8). A study commissioned by the European Parliament has also noted that the emergence of social media signals the formation of a new form of content-overladen, hyper-pluralistic and democratic-distorting public sphere in which relativism reigns supreme: every fact and opinion can be found online and so can their counterparts (Bayer et al., 2019, pp. 59–60).

If we take seriously the conclusion reached by these studies, namely that the public sphere in Western democracies is experiencing a disruptive process of transformation, then it stands to reason that certain configurations of the public sphere could be more resilient to the challenge of digital propaganda, while others would face serious vulnerabilities. The question we therefore need to contemplate is quite straightforward, although pregnant with theoretical and policy implications: *what attributes help build the resilience of the public sphere to digital propaganda and why so?* To address this question, the paper brings together three strands of scholarship. The first one relates to tactics and approaches of digital propaganda (Woolley and Howard. 2018; Walker and Ludwig 2017; Bjola and Pamment 2018), the second body of literature strand engages with matters of cognitive and physical resilience (Hansen 2017, Vilmer 2019), while the third deals with the issue of the constitution, functioning and transformation of the public sphere (C. Bjola, 2010; Fraser, 1992; Habermas, 1989).

To address these questions, the paper will proceed in two steps. As the first order of business, we need to unpack the logic of the digital propaganda by better understanding the connection between the online content promoted by disinformation sources and policy decision-making, especially in the area of foreign affairs. Second, we need to make sense of the mechanisms by which this connection could be broken or at least diluted to the point that it becomes sufficiently manageable as a political risk. We examine these issues by drawing on the case of Finland and its model of ‘resilience’ to digital disinformation, which has attracted the attention of foreign policy experts and academics worldwide for its ability to develop a promising degree of immunity to propaganda-induced political pathologies (Standish, 2017; Mackenzie 2018; Mackintosh 2019).

To this end, the paper provides an original framework for understanding the logic of the digital propaganda and the conditions that may facilitate effective resilience responses. Rather than treating the public sphere in which the ‘new propaganda’ operates as a unified space of narrative engagement, the study differentiates between the *macrosphere* of diplomatic and policy decision-making and the *microsphere* of the individual or collective quotidian experience of disinformation by citizens. It will be thus argued that the connection between the two spheres is enabled by the formation and political mobilization of ‘unruly’ counterpublics, that is of arenas of textual and visual contestation of politically marginalized groups promoting issues aligned with the disinformation agenda. Using this framework, four prominent examples of disinformation in Finland are discussed, each illustrating different aspects of the mechanisms and contents of disinformation.

THEORISING PROPAGANDA IN THE DIGITAL AGE

There is a famous remark about knowledge attributed to Aristotle, "Men do not think they know a thing unless they have grasped the 'why' of it" (cited in Kim, 1994, p. 51), which helps frame the discussion about how to approach the question of resilience to digital propaganda. Aristotle's remark reminds us that in order to make reasonable claims about an intriguing social phenomenon, we need to draw a distinction between descriptive and explanatory knowledge, that is between knowing what and knowing why. Applied to our case, this distinction encourages us to approach the question of resilience to digital propaganda in two steps: first, we need to clarify what exactly we know about digital propaganda and resilience in terms of definition, approaches and effects; second, we need to develop an understanding how this descriptive knowledge can help us produce an explanatory model for countering digital propaganda.

What is digital propaganda and how does it work?

Despite the explosion of references in the public discourse, definitions of digital propaganda generally lack conceptual clarity as the term is often proposed as an umbrella label for a range of affiliated notions and ideas, including 'post-truth', 'fake news', or disinformation. To enhance definitional precision, we call attention to the distinction between the more traditional, information-oriented aspect of propaganda and the more recent, digital or computational component of the concept. On the information side, post-truth relates, for instance, to the "legitimation of arguments based on their emotional appeal and symbolic value and subjective rather than impersonal truth" (Kalpokas, 2018, p. 2; Oxford Dictionaries, 2016). Within the broader post-truth universe, terms like 'fake news' cover incidents of incorrect, manufactured or skewed use of information, but they may refer to different types of content, including

political satire, news parodies, press releases, political propaganda, and photo manipulation (Tandoc, Jenkins, & Craft, 2019, p. 674). Recognising the analytical elusiveness of the term, a recent report by the European Commission has even declined to refer to ‘fake news’ and insisted on replacing such terms with ‘disinformation’ (European Commission, 2018, p. 3).

The move is not without analytical merits as disinformation implies a broader approach, looking at “systematic disruptions of authoritative information flows due to strategic deceptions that may appear very credible to those consuming them” (Bennett & Livingston, 2018, p. 124). Disinformation is also distinct from misinformation not so much because of the nature of information disseminated but due to the intent behind the dissemination. Disinformation is used for *deliberately* propagated false information, in opposition to ‘misinformation’ which is *unintentionally* propagated false information, and ‘online propaganda’, which refers to information of varying factual accuracy used to discourage opposing viewpoints (Born & Edginton, 2017). Recognising the shifting meaning of the term and the rising influence of the digital medium, Tucker et al. take the view that disinformation covers a broader set of circumstances, ‘the types of information that one could encounter online that could possibly lead to misperceptions about the actual state of the world’ (2018: 3).

Approaching the issue from the digital side of the equation, a second body of literature takes a close look at the transformative effect of the digital medium on the mechanisms by which disinformation is generated and disseminated. Woolley and Howard have coined the term “computational propaganda” to describe “the use of algorithms, automation, and human curation to purposefully manage and distribute misleading information over social media networks” (2018, p. 4). The key issue for them is less the informational component of the propaganda, but rather its technical aspect tasked with the manipulation of the public opinion. This includes the data-driven techniques and tools like automation (bots) and algorithms

(decision-making code) that allow small groups of actors to amplify in a targeted fashion false and abusive information into mainstream online environments (Woolley & Howard., 2018, pp. 6–7). Put differently, without this “computational enhancement”, propaganda would hardly be able to “pierce, penetrate or perforate the political and information environments in the targeted countries” (Walker & Ludwig, 2017).

That being said, the way in which the ‘computation enhancement’ works and the likely impact it has on the effectiveness of disinformation remain a matter of dispute. Dawson and Innes (2019) have found, for instance, that the disinformation campaign run by the St. Petersburg based Internet Research Agency in Europe relied on rather low-key tactics of digital engagement: account buying, ‘follower fishing’, and narrative switching. Similarly, Zannettou et al., concluded that while Russian trolls managed to stay active for long periods of time and to reach a substantial number of Twitter users with their tweets, their ability to spread news content and make it viral, however, was rather minor, with certain exceptions (Zannettou et al., 2019). On the other hand, lack of digital sophistication should not be easily dismissed. Bots have been found to be highly successful in disseminating low-credibility sources and targeting influential users (Shao et al., 2018), trolls have demonstrated good skills at attracting non-troll users into conversations (Szwed, 2016), while cross-media and mix-media strategies have proved effective in framing and disseminating malicious narratives (Agarwal & Bandeli, 2018).

Taking note of the informational and computational dimensions of the concept, we define digital propaganda as the use of digital technologies with the intention to deceive the public through the generation and dissemination of verifiably false or misleading information. By acknowledging the dual interface of the strategies of digital manipulation of information, the proposed definition provides a more insightful understanding of what resilience means in the digital context and how it may work. The emphasis placed on the information vs the

computational component of digital propaganda sets the tone, for instance, to how resilience responses are designed, crafted and implemented.

If the information dimension is considered to be the main vulnerability, then media literacy (Posetti et al., 2018) and strategic communication (Government Communication Service, 2018) emerge as the likely responses to countering digital propaganda. The objective in this case is to develop '*cognitive resilience*' so that disinformation will be prevented from taking root and being internalised by members of the target audience (Hansen, 2017, p. 35). This can be done by drafting and promoting guides of good practice by which disinformation can be recognised, monitored and responded to by journalists, governmental communicators and the public at large.

If, on the other hand, the computational aspect is viewed with greater concern, then containment (Bjola and Pamment 2016), regulation (Renda, 2018) or even cyber tactics (Vilmer, 2019) become the prioritized approaches to building and managing resilience. The objective in this case is to develop '*physical resilience*' by preventing or blocking disinformation from entering the information space in the first place. Disrupting the network of disinformation nodes, working with social media companies to remove malicious content in real-time, and using proactive cyber operations are some of the tactics that may assist this goal.

As benchmarks for assessing the effectiveness of the efforts to counter digital propaganda, the two forms of resilience complement each other well. Cognitive resilience is arguably easier to develop when physical resilience is strong, but it becomes even more important when the latter is difficult to achieve. The way in which they work together resembles the logic behind a viral infection. Digital propaganda could be seen as an outside 'virus' that invades the 'informational body' of the target country and then uses the 'chemical machinery' of the host to keep itself

alive and to replicate itself. Physical resilience plays the role of the antiviral drugs, which like the medical response to the AIDS pandemic, do not destroy the virus but inhibit its development. Cognitive resilience, on the other hand, works as a vaccine and seeks to produce enough antibodies so that further viral infections can be prevented.

The “viral infection” metaphor is useful for understanding how resilience, in its two main components, cognitive and physical, could be developed, but it also comes with certain limitations. It assumes, for instance, that the ‘virus’ will be perceived as an invading entity by the entire body, while in the case of propaganda the reception could be more nuanced, and the message may actually be welcome by a segment of the target population. The metaphor also assumes that, if left unchecked, the virus can have fatal consequences for the body, while in the case of propaganda the results are less dramatic, albeit still significant. That being said, the metaphor helps make a critical point: resilience to viral infection cannot be built without a good understanding of the ‘chemical machinery’ of the body to which the virus attaches itself. In the case of digital propaganda, the body being transgressed is the information space of the target country and the ‘chemical machinery’ that animates it is defined by the communication patterns by which social meaning is generated, negotiated and acted upon. In short, we need to understand how the public sphere is affected by digital propaganda in order to make sense of the possibilities and limitations of building cognitive and physical resilience.

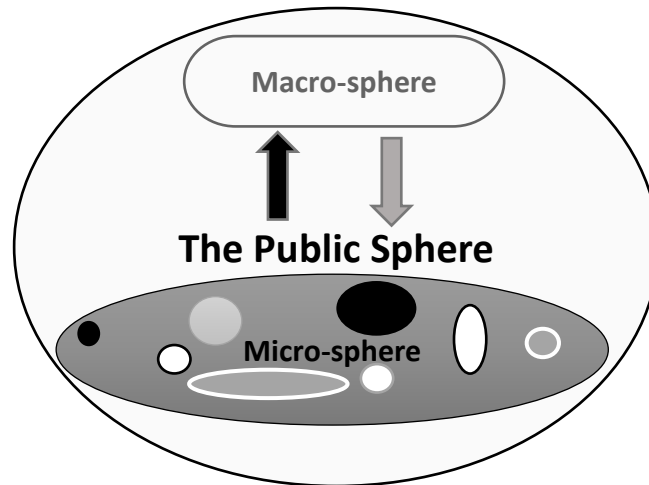
The Information Space: The Macro versus the Microsphere

The concept of the public sphere, as articulated and developed by Jürgen Habermas in his landmark study on *The Structural Transformation of the Public Sphere*, provides a useful analytical angle for exploring the discursive context in which disinformation unfolds. Habermas defined the public sphere as the “the sphere of private people come together as a public [...] to engage [public authorities] in a debate governing relations in the basically privatized but publicly

relevant sphere of commodity exchange and social labor” (Habermas, 1989: 27). For Habermas, the public sphere serves two key functions in (democratic) political systems: to allow decisions of relevance for the governance of the society to be properly articulated (will-formation) and to create suitable conditions for critical engagement and political accountability (opinion-formation). Each function besieges the other but without conquering it, a heuristic principle that helps keep potential social pathologies in check. If opinion-formation is usurped by power relations, then political manipulation of the public agenda turns hardly avoidable. On the other hand, if opinion-formation becomes too dominant then governance effectiveness is undermined (Bjola, 2011: 198).

Building on Habermas’ concept of the public sphere, we argue that digital propaganda operates at two levels: the *macrosphere* of political decision-making and the *microsphere* of the daily individual experience (see Fig 1). In the macrosphere, disinformation exists as a tactical and strategic tool, employed by hostile forces in the growing realm of information warfare. It is treated as a tool of destabilisation during election periods and is connected to broader international concerns such as a country’s diplomatic relationships or security arrangements. The goal of disinformation in the macrosphere is to strategically influence decision-makers to pursue a pre-determined course of action in international affairs by altering their set of policy preferences or, if that’s not possible, by paralysing or neutralising foreign policy decision-making processes. In Habermasian terms, disinformation in the macro-sphere is primarily about influencing or by case undermining the will-formation ability of the target government.

Fig 1: *Digital propaganda and the public sphere*



By contrast, the microsphere is the realm of disinformation within which individuals, as citizens or members of communities, interact with the information sphere by sharing questionable online articles with little or no foundation, visiting Reddit and 4chan sub-threads, and trolling on Facebook comment sections and Twitter. It might be helpful to think of the microsphere as the sphere that contains the lived experience of disinformation. When zooming in to observe disinformation in practice within the microsphere, we can better understand its nuance: it contains both conventional understandings of ‘fake news’, but also individuals and events that make up the greater global information war. In the macrosphere, regular citizens are treated as vulnerable stakeholders (light coloured arrow in Fig 1); they are primarily seen as consumers of information, with the source of the disinformation problem lying outside their purview. Within the microsphere, citizens are both consumers as well as producers and distributors of disinformation. This may happen unwillingly, when citizens fail to detect the false content they share, or sometimes intentionally when the shared message validates a particular political position (dark coloured arrow in Fig 1).

There is a reasonable expectation that disinformation in the microsphere would stay contained in the microsphere due to the relatively marginal relevance of most of the issues for policy

makers. However, some of disinformation topics have the potential to infiltrate the macrosphere, and when they do, conditions for the formation of social pathologies are likely to be met. Habermas actually anticipated one important type of social pathology, the ‘feudalisation’ of the public sphere, in which rational-critical discussions of public affairs are replaced with a mood of conformity to public figures, while consensus is manufactured by “sophisticated opinion-molding services under the aegis of a sham public interest” (Habermas, 1989: 195). When this happens, the connection between will- and opinion-formation is broken, and the norm-sustaining function of the public sphere in the society decays or ceases all together. Disinformation infiltration from the micro-to the macro-sphere may operate in a similar fashion and facilitate a ‘feudalisation’ of the public sphere in which critical decisions are taken on the basis of emotional framing and algorithmic-driven engagement with minimal consideration for the rational-critical value of the underlying propositions.

By zooming in on the patterns by which disinformation may grow from isolated online conversations to becoming pressure points on governmental decision-making, the distinction between public macro- and micro-spheres opens the ‘black box’ of digital propaganda and promises to provide an innovative framework for understanding its logic (the way in which it operates to generate political effects). More specifically, the model suggests that certain type of disinformation planted in the micro-sphere (Step 1), manages to recruit adherents and gather momentum (Step 2), to the point that it may prompt a policy change at the governmental level (Step 3).

The logic replicates the Habermasian understanding of the public sphere, but with a critical twist: opinion- and will-formation are not based on a genuine exchange of information between the public and its leaders with the goal to serve the collective good, but on active dissemination of falsehoods with the strategic goal to manufacture and amplify distrust between the public

and its leaders. The macro-micro distinction helps delineate the informational context that digital propaganda seeks to corrupt, but tells us little about how digital propaganda operates and replicates itself and what forms of resilience may prove effective to countering it. By shifting the analytical focus from the structural to the processual dimension of digital propaganda, counterpublic theory helps provide the missing link.

Counterpublic Theory and Resilience

One of the most interesting reactions to Habermas' concept of the public sphere has been the counterpublic theory, which refers to "the parallel discursive arenas where members of subordinated social groups invent and circulate counter discourses to formulate oppositional interpretations of their identities, interests and needs" (Fraser, 1992: 123). The invention of a counterpublic space derives its constitution from the public itself, given that often, the 'discursive styles of some intellectual publics [...] not only help to constitute a community for a marginalised and jeopardised sphere of society but also function to call alternative public into being' (Blaustein, 2004: 178). The role of counterpublics, as their constitution suggests, is derived and determined by their opposition to the mainstream public sphere. To this end, Fraser contends that counterpublics serve two roles, 'on the one hand, they function as spaces of withdrawal and regroupment; on the other hand, they also function as bases and training grounds for agitational activities directed toward wider publics'(Fraser, 1992: 124).

While counterpublic theory has been normatively associated with emancipatory inquiries into the subaltern status of traditionally marginalised social groups (women, workers, ethnic, racial or sexual minorities), its framework offers analytical currency for studying "unruly" publics as well, such as right wing populist movements or radical groups that may reject basic democratic principles (Kaiser & Rauchfleisch, 2019: 4). The key question that deserves attention in this context is whether counterpublic theory can provide useful analytical insights into how digital

disinformation operates, especially with respect to the mechanisms of infiltration between the micro- and the macrosphere. In other words, to what extent disinformation contributes to the formation of counterpublics and how do they, in turn, influence decision-making in the macrosphere.

As Toepfl and Piwoni point out, counterpublics may be distinguished by their inward-outward orientation. Inward-oriented goals are focused on “the invention, elaboration, and formulation of alternative identities, interests, and needs” while outward-oriented goals aim at “breaking up and shifting of consensus structures within dominant publics by engaging with wider audiences and targeting dominant publics” (Toepfl & Piwoni, 2018: 2014). Building on this distinction, Kaiser & Rauchfleisch propose a ‘vicious’ model of counterpublic formation to explain the rise of extreme right groups in Western Europe and the US. More specifically, they argue that above a certain threshold, the counterpublics’ efforts to provoke a response from the system to their demands is no longer recognised. When this happens, a vicious cycle is initiated. Although more extreme activity may lead to more presence in the media and strengthens the counterpublic’s collective identity, it decreases the probability of integration into the mainstream and a response from the political system (Kaiser & Rauchfleisch, 2019: 10).

Applied to the case of digital propaganda, four testable propositions follow from the arguments above about the conditions under which disinformation may lead to the ‘digital feudalisation’ of the public sphere. First, disinformation deforms the public sphere by corrupting the epistemological basis for truth-claim validation. Conditions for establishing epistemic agreement simply vanish when subjectivity is imposed as quasi-standard of objectivity. Second, disinformation assists the constitution of counterpublics by creating a space for discussion in the micro-sphere of topics of high emotional intensity. Emotions drive interests, which in turn facilitate identity connections. Third, disinformation infiltrations from the micro- to the macro-

sphere require the presence of actively outward-oriented counterpublics. Not all counterpublics are the same, and the systematic use of disinformation is one important attribute that sets them apart. Fourth, the more extreme the disinformation attempt to infiltrate the macrosphere, the less likely it is to be accepted by the system and by consequence, the more vocal or radicalised the counterpublic becomes.

The four propositions allow us to complete a second important analytical step necessary for unpacking the logic of digital propaganda. More specifically, they allow us to theorise about a key mechanism by which digital propaganda operates and replicates itself: the formation and promotion of unruly counterpublics. By failing to provide a meaningful space for informational exchange and political dialogue between citizens and authorities, a public sphere infiltrated and controlled by unruly counterpublics has no other recourse than to turn against itself. Without a firm epistemological ground for validating truth-claims, emotions become the primary benchmarks for assessing the authenticity of the statements circulated in the public sphere. Emotional engagement, on the other hand, facilitates identity-affirming conversations and empower constituencies that are primarily interested in excluding rather than inviting others to publicly debate ideas of relevance for the public good. In fact, the very idea of the public good, which is the cornerstone of the public sphere theory, loses its meaning as a collective anchor of policy making in favour of narrowly defined and mostly divergent conceptions of counterpublic interests.

Therein lies the third important contribution of this study: it proposes a conception of resilience, *digital resilience*, that seeks to isolate the connection between digital propaganda and the formation of unruly counterpublics. Digital resilience builds on the idea of cognitive and physical resilience discussed above but with a twist: the object to be protected is the public sphere itself against the threat of dissolution posed by the digital rise of unruly counterpublics.

From a digital resilience perspective, the public sphere requires enhancements that enables it to reliably validate truth-claims based on accurate information and rational engagement, contain the effect of emotional escalation, prevent the radicalisation of unruly counterpublics and reinforce the integrity of the notion of the public good. The case of Finland to be discussed in the next section provides an excellent empirical example for examining the connection between digital propaganda and the formation of unruly counterpublics as well as for discussing the effectiveness and limitations of some of the resilience measures that have been introduced for protecting the Finnish public sphere.

DIGITAL PROPAGANDA IN FINLAND

Finland's reputation for resilience to digital propaganda has attracted wide attention in the global media (Standish, 2017; Mackenzie 2018; Mackintosh 2019) and for good reason. As pointed out by René Nyberg, the former Finnish ambassador to Russia, "Few countries can match Finland's long experience of dealing with Soviet and Russian hybrid warfare—before, during, and after the Cold War—and few countries have had as much success in standing up to it" (Nyberg, 2018). However, a crucial piece of information is still lacking: what exactly are Finns so resilient to and how so? Digital propaganda indeed shares certain similar traits across countries, such as being anti-immigration or favouring radical right-wing or populist media.¹ However, each society is targeted differently with content that is more customised to its own 'weak spots': Russian trolls targeted Hilary Clinton during the US presidential election in 2016 to further ignite the pre-existing rhetoric of 'crooked Hilary'(Griffin, 2016); on the other hand, the 'Lisa Case' in Germany, a Russian-backed false news story claiming a German-Russian woman had been assaulted by Arab migrants, was intended to fan the flames of pre-existing xenophobia and islamophobia (Stefan Meister, 2016).

In the following examples, we examine what digital propaganda has meant in the case of Finland. This research employs a within-case method of qualitative analysis to grasp the complexity of digital propaganda. As Goertz and Mahoney indicate, within-case analysis is a particularly suitable methodology where one seeks to explain why a particular case has a specific outcome. To increase the reliability of the conclusions, the researcher is expected to identify junctures when key events directed the case toward certain outcomes and not others. This way,

¹ For example, see Szabó et al. (2018), who conduct a cross-country comparison of radical right media, focusing on trends in Central and Eastern Europe. Similar trends are noted in the analysis of Finnish radical right media by Ylä-Anttila (2018). Additionally, Marwick & Lewis (2017) show that the American 'post-truth' repertoire consists of similar themes to the former cases (anti-immigration, anti-establishment, often also homophobia and sexism).

the overall explanation will be rich with details about specific events, conjunctures, and contingencies (Goertz & Mahoney, 2012, p. 89). We extract analytical leverage from the Finnish case study by comparing four distinct examples of digital propaganda that have manifested themselves in the recent years in the Finnish public sphere.

The four examples are selected here for their representative power. The cases vary as to how they reflect the level of infiltration of digital propaganda from the micro- to the macro-sphere as a result of counterpublic mobilisation. Case 1 is prevalent, for instance, in the macrosphere discussions, and the Russian involvement is obvious. Case 4 exists thoroughly in the microsphere and it is largely shaped by the activities of the counterpublic. Cases 2 and 3 fall somewhere in between Cases 1 and 4. The outcome (“digital feudalisation” of the Finnish public sphere) is thus explained using within-case tracing of how discernible, explicit, and traceable Russian involvement is in each situation, on one hand, and on how relevant counterpublics are in disseminating propaganda, on the other hand. In each example, we empirically probe the four propositions discussed above concerning the epistemic deformation of the public sphere, the role of emotional engagement in the constitution of counterpublics, the presence of actively outward-oriented counterpublics and the conditions for counterpublic radicalisation.

Empirical evidence in each case has been gathered using secondary resources (academic and media reports) and semi-structured interviews with seven Finnish researchers, journalists, and officials (see the complete list in the bibliography). All interviews were conducted in full compliance with academic ethical standards for qualitative research in the social sciences. Interviewees were selected based on their affiliation to major organizations, such as the Finnish Ministry of Foreign Affairs and the Hybrid Center of Excellence, or due to their personal experience or expertise with digital propaganda. Snowball sampling was also used. The role of

the interviews was to provide additional insight into each case so that an in-depth and detailed account could be built for understanding how each case of digital propaganda was received, interpreted and responded to by the Finnish authorities.

Case Study 1: The 'Child Custody' Narrative

In most literature, the Russian annexation of Crimea in 2014 is cited as a critical juncture for the escalation of disinformation campaigns into the information sphere of Western countries (Renz, 2016). This observation was confirmed in an interview with a senior Finnish diplomat, Vesa Häkkinen, who specialises in disinformation in Finland. During and after the annexation of Crimea there was a rise of certain trends in the information sphere, the patterns of which made it obvious that there was a coordinated dissemination of false information. Since 2014, a set of seemingly unrelated events were all followed by disinformation narratives online that shared many similarities.

Specifically, Finnish authorities observed similar narratives across many events involving Russian actors and media: the Russian annexation of Crimea in 2014, the shoot-down of the Malaysian Airlines MH17 flight over the Russian-Ukraine border in 2014, the use of chemical weapons in the conflict in Syria in 2016, as well as the incident involving the poisoning of former Russian military agent Sergei Skripal (Häkkinen, 2018). These instances, Häkkinen claimed, were followed by coordinated disinformation campaigns aiming to distract from the facts of each situation and force artificially inflated topics on the public agenda.

This perspective is shared by many observers in other European countries, who make similar remarks about Russia's involvement in the information sphere regarding these events (EUvsDisinfo, 2018). Therefore, we witness the convergence between disinformation trends across Europe and those in Finland. However, closer analysis reveals that there also exists a

more customised, localised content. One such case is Finland's 'child custody' narrative. Though the story varies, the core of it is rather straightforward: according to Russian media, Finnish social services systematically engaged in unrightfully seizing the children of Russian parents across Europe, placing them in social care or giving them up for adoption to same-sex couples in Western countries. The Russian children's rights ombudsman even accused Finland of perpetrating "juvenile terror" (Pavel Astahov, 2014).

Every interviewee contacted for this research treated this disinformation campaign as common knowledge, suggesting that the story infiltrated the microsphere, even if it did not gain coverage internationally.² These stories were mainly circulated on Russian outlets including *Russia Today*, *Sputnik*, and *Pravda*.³ Within Finland, they were circulated in part by Johan Bäckman, an informal leader of the pro-Kremlin Finnish counterpublic, who positioned himself as a 'human rights activist' advocating for the Russian minority group in Finland (StopFake, 2016).

Saara Jantunen, researcher at the Finnish Defence Research Agency, suggests this disinformation campaign began not long before the Crimean annexation (Jantunen, 2018). The version of the story she is familiar with is that the Finnish government, as well as other European authorities, were kidnapping the children of Russian families and putting them in concentration camps or selling them to gay couples. One Finnish official who has been working as a counsellor on Finnish strategic communication with Russia from 2009 to 2014, claimed this specific disinformation campaign was noticed as early as 2009, citing the child-custody case as the starting point for Finnish public concern around media-influencing (Finnish Diplomat

² The narrative has pervaded some articles on Finland, but is often briefly mentioned without much analysis or contextualization. For examples, see Standish (2017) or Gessen (2015).

³ Examples of such articles in English can be found in Pravdareport (2012), RussiaToday (2012), RT International (2011).

A, 2018). Häkkinen agrees that Finland has been facing these ‘family cases’ since the late 2000s, with 2014 marking a turning point in the intensity of Russian disinformation.

Fig 1: *Visual support for the child-custody narrative in the Russia media*⁴



Russian denied access rights in Finland due to kidnapping scare



Finnish-ed! Russian mum kept from her son

The narrative touches on emotionally intense themes (children-parents separation, same sex relations, Russian minority rights), supported by controversial visuals (see Fig 1) with the intention to provoke a strong public reaction, tarnish the international reputation of the Finnish government and deter the Finnish government from taking a critical position against the Russian policy in Ukraine. The case provides support for three propositions regarding the mode of operation of digital propaganda: the use of false statements for undermining the epistemological basis for truth-claim validation in the public sphere; the importance of emotional engagement in provoking identity disputes; and the role of counterpublic leaders (e.g., J. Bäckman) in facilitating propaganda infiltrations between the micro- and macro-sphere.

However, despite the intense promotion of the Russian media, the ‘child custody’ campaign has failed to become an identity defining issue for the Finnish counterpublic. Finland’s comprehensive security approach to countering disinformation has been given credit for this

⁴ Image (left): (RT International (2010)); Image (right): (RT International 2011)

result (see more in the next section). Russian attempts to fragment the information sphere were countered by a pro-active and unified response involving domestic government and civil society actors. In this particular case, the responsible authority, the Ministry for Social Affairs and Health, coordinated its response with other agencies, including, for instance, the Ministry of Foreign Affairs and the relevant embassies. Additionally, transparency and close cooperation with media served well as tactical instruments for debunking child-custody disinformation almost in real-time (Häkkinen, 2018).

Case Study 2: Finland and the EU Migrant Crisis

This case also directly and explicitly involves Russia as a state actor who is simultaneously deploying disinformation and (anti)diplomatic tactics. Though this case is entirely absent from academic research on disinformation, it was often mentioned by interviewees when asked to cite prominent examples of digital propaganda. The case is a useful example of how actual, ‘true’ events can be orchestrated to have a tactical impact on another country’s information sphere.

In 2015-2016, Finland saw an unprecedented amount of migration across its historically secure and well-functioning Eastern border. At a first glance, this event seems to be a diplomatic incident around border security: the Russian side of the border was being managed improperly, allowing for the ‘sudden appearance’ of over 2000 asylum seekers (of primarily Middle-Eastern origin) to cross into Finland, through a border that had not experienced issues since World War II (WWII). The problem continued until 2016, despite multiple measures from Finnish border control to prevent the phenomenon. Additionally, there were accusations that Russian officials were not doing enough to obstruct human smuggling and were even aiding the illegal crossing (Yle Uutiset, 2016a).

Fig 2: Visual framing of the migrant border crisis in international vs Finnish media⁵



How is this an example of disinformation? According to Häkkinen, “It’s not disinformation exactly, but it’s a message of some kind.” (2018). The complexity of this case is why the term ‘hybrid influencing’ or ‘information influencing’ is often preferred by Finnish officials. There are two levels on which this case sends a message. Here, again, the macrosphere/microsphere framework can be useful. On the broader macrosphere, this is a case of selective manipulation of border security: it is important to note the chronological coincidences surrounding the event. The trajectory of Russian-Finnish border security maps easily onto the timeline of sanctions towards Russia from 2015-2016 (European Council, 2018), with thousands of refugees showing up at the border just one week after Finland supported renewed EU sanctions on Russia in December 2015. The timing also coincided with a period in which public discussions intensified about the possibility of Finland to join North Atlantic Treaty Organisation (NATO). A report on this topic commissioned by the Finnish Ministry of Foreign Affairs (MFA) and published in April 2016, warned, for instance, that Russia would likely take “strong, even harsh” political and economic actions to thwart any move by Finland or Sweden to join NATO (Finnish Ministry of Foreign Affairs, n.d., p. 7).

The Finnish MFA thus correctly identified this event as a hybrid event: on the macrolevel, it combines traditional tactical threats between countries (border security), but on the micro-level

⁵ Image (left): (Higgins, 2016a); Image (right): (Yle Uutiset, 2016a)

it could severely influence how people *perceive* information about Finland's policies on asylum seekers. Events like this can be seen as "information impulses" because rather than being disinformation as such, they combine disinformation with public actions. Put simply, this event was more than a traditional threat from a larger country to a small country, which may have alarmed some members of the MFA. It was aimed at instilling fear by staging a visual effect (see Fig 2): thousands of refugees (potentially illegally) crossing into Finland, giving the impression that Finland is unprepared and incapable of handling the situation.

Russia correctly identified immigration as Finland's societal vulnerability and tailored its hybrid influencing to target this cleavage. The event thus contributed to amplifying identity disputes in the Finnish public sphere, as suggested by the second proposition. The operation aimed to create a problem for Finnish decision makers with the offer to manage it but without necessarily solving it (Finnish Ministry of Foreign Affairs, n.d., p. 53). The finding is significant as it reveals critical differences in the way in which digital propaganda operates when its main objective is to shape behaviour in the macro- vs the micro-sphere. It implies, for instance, that the use of false information is deemed more important in the micro- rather than the macro-sphere, a fact that can be explained by the different standards by which information is deemed credible in the two spheres. Simply put, the pressure put on governments in the macro-sphere to amend their positions should be highly credible and tangible as deterrence theorists have long argued.

Case Study 3: The Targeting of Journalists and Researchers

Since 2014, Finland has noticed a rise in smear campaigns and 'trolling'⁶ aimed at journalists and researchers who speak out about Russian disinformation tactics. Trolling has become a

⁶ 'Trolling' refers to a specific type of malicious online behavior, intended to disrupt interactions, aggravate interactional partners and lure them into fruitless argumentation (Coles & West, 2016, p. 233).

critical tool of disinformation warfare (Zannettou et al., 2019), not least because it offers those involved the possibility of maximising the online impact of their preferred narratives by silencing potential critics. In so doing, trolling induces a strong deformation of the public sphere that touches on all main components discussed above: forcefully challenging the epistemic basis for validating truth claims, using highly charged emotional language to discredit opponents, and creating a safer environment for the formation of unruly counterpublics. The Finnish case provides a good empirical illustration of the mechanisms by which trolling contributes to the ‘feudalisation’ of the public sphere.

Jessikka Aro, a Finnish investigative journalist, became the public face of this issue after she began working on “a crowdsourced series of stories about pro-Russian trolls, the info warriors distributing social media propaganda and influencing Finnish people” for Yle, Finland’s national broadcasting company (2018). Aro’s experience quickly gained public traction because of the “vicious retaliation” she experienced: abusive messages flooded her social media and email, even from someone pretending to be her deceased father. Smear campaigns labelled her a drug dealer and music videos were released mocking her as a “delusional bimbo” (Higgins, 2016b). Aro refused to let herself become intimidated by the vicious online campaign and turned these trolling activities to her advantage by raising awareness of the issue. She even crowdfunded about \$30,000 (€26,000) for conducting research for a book titled *Vladimir Putin's Troll Empire* (Schultz, 2018).

Doctor of Military Science Saara Jantunen also faced harassment after the release of her book *Infosota* (translation: ‘Infowars’), which details Russian information war tactics such as the use of bots and trolls to flood social media with “anti-Western and pro-Kremlin comments” (Yle Uutiset, 2015). However, despite her research, she acknowledges the “home-grown” nature of the harassment: she views the online users involved in trolling as “cheap labour” for certain

websites and considers figures behind the targeting, like Johan Bäckman, as managers of outsourced harassment.

From a broader perspective, a survey of 875 working journalists in Finland indicates that this kind of harassment has already impacted the media landscape: two-thirds of journalists reported experience with abusive comments or in-person hate speech reactions, while 17% had experienced public defamation. Importantly, journalists were able to identify the topics that trigger harassment: stories focusing on immigration, multiculturalism, sexual minorities, religion, equality issues, healthcare, social security and Russia (Hiltunen, 2018: 18). The Finnish Ambassador for Hybrid Affairs, Mikko Kinnunen, stated that harassment “is part of influencing, [trolls/harassers] try to tire individuals so they switch to silence or make journalists write on something else” (2018). Therefore, disinformation goes beyond the construction of narratives; it can also directly impact the pluralism of voices reaching the public sphere.

Victims of harassment campaigns, such as Jessikka Aro and Saara Jantunen, have launched high-profile court cases to defend themselves against ‘trolling.’ However, their cases face an interesting predicament: while online harassment could grow to the point of overtaking an individual’s life, it is nevertheless difficult to hold individual users accountable, who may have sent just one or two messages each. Aro was unable to use her email or phone, due to the number of messages she received from a myriad of users, some of which may have been bots (Jantunen, 2018). Journalists are accustomed to the occasional critical or aggressive email, but targeting exponentially increases these instances. This leads to the question: how do you deal with the crowdsourced harassment generated by unruly counterpublics? If successful, these court cases could provide a legal approach to prevent similar mobilisation of malicious counterpublics in the microsphere in the future.

While many other Western countries approach online trolling and harassment as a ‘freedom of speech’ debate, Finland has approached these problems from the perspective of freedom of communication *via* an interesting new principle in Finland: *viestintärauba*. Though there is no exact English equivalent, it approximately translates to ‘communication peace’, which can be interpreted in this context as freedom from unsolicited communication (Jantunen, 2018). This concept is critical in the upcoming trials, in which historic principles regarding harassment are being reformulated in the context of online communication.

Another innovative aspect of these trials is their focus on the individuals that instigated the harassment through selective public targeting, as well as the media that facilitated the slander. While certain countries are looking at the role of *platforms*, Finland is focusing on the role of high profile *people*, prosecuting the public figures who they deem ‘responsible’ for the targeting in the first place. One notable defendant is Johan Bäckman, who is charged with harassment and aggravated defamation (Yle Uutiset, 2018a). Another is Ilja Janitskin, the founder of nationalist media including *MV-Lehti*, who is accused of similar offences, and was extradited from Spain to Finland for trial (Yle Uutiset, 2018b). Even though Bäckman may not have personally emailed, called or stalked journalists, he is the one who named and targeted individuals, suspecting that his online followers would respond accordingly (Jantunen, 2018).

In a landmark decision, the Helsinki District Court found in October 2018 that Mr. Janitskin and Mr. Backman had worked together to slander Ms. Aro and had committed “an exceptionally aggravated set of crimes” because their primary motive was to undermine her work investigating “Russian information threats” by destroying her “professional credibility and reputation as a journalist specializing in Russian affairs.” Janitskin was sentenced to 22 months in jail, whereas Bäckman received a one-year, suspended jail sentence for aggravated defamation and stalking (Higgins, 2018). There is thus an expectation that these trials will

improve the legal standards for harassment, encompassing the social harms of violation of ‘communication peace’ in addition to the private psychological harms. Indeed, the role of platforms themselves will become critical in the future of communication peace. Long term resilience to disinformation requires a pluralist information sphere. In turn, this requires finding the appropriate legal and societal framework to ensure rights to peaceful communication.

Case Study 4: Finnish ‘Alternative Media’: MV-Lehti and Hommaforum

There is a growing consensus that disinformation does not just rely on implicit cognitive biases (Lazer et al., 2018: 1095); it also requires active involvement of individuals in the production and dissemination of fraudulent content. At the macrosphere level, disinformation is often dominated by debates about Russian involvement and states’ strategies to countering the dissemination of information deemed false. However, zooming into the microsphere, we see that a framework of ‘true’ and ‘false’ becomes much less valuable. This is because, for many ‘alternative media’ followers, it is the established traditional media or political elites that is seen as being behind false information or propaganda (Ylä-Anttila, 2018).

While there exist many smaller and lesser known sites, *MV-Lehti* and *Hommaforum* are often singled out as being the most active and pervasive news hubs in the Finnish ‘alternative media’ landscape.⁷ *MV-Lehti*, which translates to ‘WTF-media’ resembles the right-wing ‘fake news’ websites that are popular in other countries, such as *Breitbart* in the US. *MV-Lehti* has been compared to British tabloid newspapers like *The Sun*, though it is likely more equivalent to *The Daily Mail* in that it displays a strong partisan stance and stylistic elements of media populism which are not generally regarded as legitimate in the Finnish media landscape (Hiltunen, 2018). In addition to their style (such as sensationalist headlines and images, language and tone), the

⁷ *MV-Lehti* is cited as one of the most popular counter-media in the “Learning” section of Yle (the education part of the site targeting youth) by (Vehkoo, 2016).

substance of these sites shares similar themes with other right-wing media outlets. A close investigation of these cases is useful because it provides insight into the core views of the microsphere and it helps develop an understanding of what these alternative media are advocating or discussing. This is why most of the recent literature chooses to refer to these media and forums as ‘counter-media,’ ‘alternative,’ or ‘populist,’ thus abandoning terms like ‘fake’ or ‘post-truth.’⁸

Information about the opinions, identity and ideology of these two examples, *MV-Lehti* and *Hommaforum*, has been analysed in two recent studies conducted by researchers at the University of Tampere. One quantitative discourse analysis has used topic modelling to analyse large data sets drawn from *MV-Lehti* and *Hommaforum* in order to gauge the most commonly discussed topics and commonly used arguments on the websites (Ylä-Anttila, 2018). The other study is a qualitative investigation of Finnish media users who follow counter-media. The team interviewed 20 counter-media followers, who mostly followed *MV-Lehti* but also populist media sites, whose defining trait is a mistrust of mainstream media (Elina Noppari et al., 2019).

The core values of these ‘countermedia’ communities are rather easy to identify. They contain a base layer of typical ‘alternative,’ socially conservative views that mirror many trends in the American ‘alt-right’; for example, they spread pseudoscientific views about how “Muslims and Somalis are prone to commit rape and avoid paid work” and blame women, religious minorities and mainstream media and elites for social decay (Ylä-Anttila, 2018). However, the research on the followers of these sites reveals some interesting novel traits. Rather than being directly fascist, sexist, racist or anti-Semitic, these counter-media communities are making an *epistemic* argument: they seek out and promote alternative media because they believe mainstream media

⁸ The point is made in (Elina Noppari et al., 2019), (Ylä-Anttila, 2018) and echoes in interviews by Hiltunen (2018) and Tampere (2018).

is biased in favour of groups they dislike. This provides support for the third proposition regarding the role of actively outward-oriented counterpublics in enabling disinformation infiltrations from the micro- to the macro-sphere. The way they do it is by seeking out and promoting alternative media with a kind of ideological urgency: they see themselves as becoming activists out of a sense of injustice, and claim to be exposing the ‘corrupt’ political environment in a way that mainstream media does not (Hiltunen, 2018).

Rather than being ‘far-right’ or ‘far-left,’ the followers of counter media are united by a dismay with conventional party politics and a suspicion of institutions like the established media. The groups are actively engaging in populist knowledge because they deem it more legitimate, and hence more ‘true.’ Therefore, these online communities use grassroots reporting to construct a kind of counter-knowledge. Like the news examples we saw in previous cases, these communities share similar values: anti-EU, anti-globalist, anti-elite and anti-liberal. They dislike multiculturalism and feminism, arguing that these trends have blinded the public and its institutions to the problems of Islam, immigration, and other factors they see as a threat to Finnish society.

Viewing these information communities as counterpublics is valuable, as this term captures their isolation from the public sphere. However, rather than merely existing in the shadows of the digital microsphere, counterpublics seek to challenge the dominant public sphere (Downey and Fenton, 2003). Warner notes that counterpublics may not always exist in subordination to the public but can potentially acquire agency (2002). These trends can be detected in the Finnish counter-media follows and forums: they aim to establish their own counter-knowledge that actively seeks to delegitimise current institutions, on the one hand, and to elevate populist counter-claims to the realm of public debate on the other hand. As we saw in the case of the

targeting of journalists, member of these counterpublics harness their capacity to disrupt reporting and obstruct the flow of information in the public sphere on certain ‘trigger’ topics.

The usefulness of the framework of counterpublics becomes even clearer when we look at the anti-democratic nature of Finland’s counter-media: they often challenge the status quo of liberal democracy and its party-politics, as well as the multiculturalism and diversity of modern Finland (Elina Noppari et al., 2019). However, these groups resist easy labelling as ‘pro-Russian trolls.’ These alternative media followers play an indirect but significant role in disinformation, but not necessarily because they are motivated by Russia. Instead, these Finnish counterpublics are rather a target of disinformation since they already feel alienated from Finnish society. Interviews with followers of counter-media indicate that most of these websites are operated by disillusioned ‘amateurs’ (Hiltunen, 2018). This is not to underestimate the role of Russian disinformation campaigns in spreading, say, anti-immigrant or anti-EU arguments. Instead, the cases of *MV-Lehti* and *Hommaforum* give us a more nuanced understanding, showing that the relationship between Russia and these media is less direct than may appear. Rather than being agents of disinformation per se, the followers of alternative media are a topic of interest for scholars studying digital disinformation because they comprise the most fertile ground for social divisions. **DISCUSSION: WHAT CAN WE LEARN FROM FINLAND?**

The Finnish experience of coping with disinformation tells us that disinformation can be very damaging for the fabric of the society, but, at the same time, it can be resisted and by case defeated. More specifically, as suggested by the four propositions mentioned above, disinformation may undermine the epistemological basis for truth-claim validation, may enable the formation of ‘unruly’ counterpublics, may force the removal of critical filters between the micro-sphere and the macro-sphere, and may help generate a self-sustainable process of societal radicalisation. However, as the four case studies illustrate, there is no inexorable train of the

logic of digital propaganda as long as the conditions that enable disinformation to thrive are understood and contained, if not dismantled.

If digital disinformation makes it so easy to destabilise truth as the first proposition indicates, then what needs to be done to restore confidence in our ability to make sense of social reality? As seen in case study 1, a well-coordinated response involving domestic government and civil society actors could be highly efficient in reducing the potential for falsehoods to ‘stick’ to the public agenda and go viral. One should also bear in mind that debunking comes with a major limitation: it is *reactive* rather than *preemptive*. The quick reaction of Finnish officials to the “child custody” narrative proved successful in containing its toxic impact, but that may become more difficult in the future given the pace of technological advancements and the rise of ‘deep fakes’. As Jantunen pointed out: “when you can fake a video-stream in real-time, how are you going to fact-check that? If you are only checking facts, you are not looking forward.” (Jantunen, 2018). The solution may rest with providing people with relevant information from sources that share their point of view, introducing facts via well-crafted visuals, and offering an alternate narrative rather than a simple refutation of the disputed story. Päivi Tampere, the Head of Communications at the Hybrid Center of Excellence, supports this conclusion as her goal is to develop a strong, consistent narrative to disinformation rather than being reactive (Tampere, 2018).

The formation of ‘unruly’ counterpublics, some of them with a clear anti-democratic profile, is a tangible result of disinformation as case studies 3 and 4 have shown. This is an important conceptual contribution to the original Habermasian theory of the transformation of the public sphere as well as to its more critical follow-ups. Counterpublics are not only about subordinated social groups seeking to call attention to progressive issues in an effort to expand the agenda of the public sphere. As our study shows, counterpublics are also about groups constituted

online and empowered by digital platforms, seeking to use themes and topics, often in alignment with the digital propaganda of a foreign government, to undermine or even block the functioning of the public sphere.

The core beliefs of certain dominant Finnish counterpublics could be summarised in two traits. Firstly, they have a set of beliefs about issues concerning Finnish society and consequentially identify as anti-immigrant, anti-Islam, anti-feminist. Moreover, they often call for a return to conservative Christian family values. Secondly, they have a set of views about Finnish institutions, which they claim to be corrupt, including the ‘corrupt mainstream media’ which refuses to provide ‘fair’ reporting. Some of these narratives may serve broader geopolitical interests, and hence fit in well with the macrosphere perspective on disinformation by fuelling Russia’s anti-EU and anti-NATO narratives. However, these counterpublic narratives in the microsphere also occur organically and find deeper roots in the domestic political landscape. Horsti shows how the anti-immigration movement, for example, has been developing in the Finnish media landscape as early as 2003, and was shaped by a series of political events, including the success of the nationalist Finns Party in 2013 (Horsti, 2015).

Resilience to the formation of ‘unruly’ counterpublics, then, is about overcoming cleavages in Finnish society. This has been emphasised in recent scholarship on media literacy. Rather than being a politically neutral argument about creating more educated citizens, it has been argued that media literacy should also deal directly with civic issues, emphasising certain values along with more technical skills. ‘Civic media literacy’ should encompass both critical thinking skills and a pedagogy of online culture as a realm of civic engagement (Mihailidis, 2018). Indeed, the problem of online ‘echo-chambers’ goes beyond a lack of exposure to facts or information. Countering disinformation spheres is also about undermining beliefs “that persevere despite encounters with counter-evidence” (Garrett, 2017: 370).

As case studies 2 and 4 have shown, a key mechanism by which disinformation operates is by increasing the fluidity of the boundary between the micro-sphere and the macro-sphere. This is an important finding that adds new insight to the literature concerning the tactics of digital propaganda. Politically divisive issues in micro-sphere on migration and elitism are thus linked to macro-sphere questions on border security and European geopolitics. While current studies have generally focused on how digital propaganda seeks to destabilise the informational environment in the target countries, there has been little discussion about the mechanisms by which such forms of ‘feudalisation’ of the public sphere may take place. As demonstrated by the Finnish case, it is the link between the micro- and the macro-sphere that can ensure the success or failure of the propaganda campaign.

How, then, can decision-makers, who exist in the macrosphere, facilitate resilience building into a ‘feudalised’ microsphere containing a collection of ‘unruly’ counterpublics? The answer, in the Finnish case, rests with the issue of social trust. Finland’s strong public education system emphasises digital and media literacy, teaching citizens to identify bias or skewed narratives in their information sphere and to critically engage with new technological platforms like social media. For example, the Finnish Ministry for Education and Culture is cooperating with the National Audiovisual Institute to promote children’s media skills (KAVI 2018). Finland also has a long-term track record of media freedom and journalistic integrity. Though it recently dropped to the fourth place of the Press Freedom Index, it still maintains a leading position on a global scale (Yle Uutiset, 2018c).

Additionally, Finns still have a positive relationship with traditional news sources; even print journalism is still unusually popular despite their extremely high internet-penetration index at 94% (Newman and Fletcher, 2016: 51). Furthermore, Finland traditionally has a “neutral media

space” that is free from the heavily partisan political reporting that is common elsewhere (Hiltunen, 2018). Finland’s Council for Mass Media oversees journalistic practices and examines the reasons for the emergence of fake news in recent years (Koponen, 2018). This point connects to the issue of ‘trust,’ at which Finland is again a global frontrunner. This trust is reinforced by Finland’s history of political transparency; it ranked 3rd in the world on the Transparency International Index for 2017 and is repeatedly listed as one of the least corrupt countries in the world by most metrics (Transparency International, 2018).

Last but not least, the problem of societal radicalisation induced by disinformation cannot be dismissed as case study 3 clearly demonstrates. The legal approach centred on the concept of *viestintärauha* (freedom from unsolicited communication) is a step in the right direction as it promises to reduce the intensity of online harassment. More broadly, however, a ‘containment strategy’ around disinformation based on knowledge from strategic communications (Bjola and Pamment, 2016) might be needed to counter the radicalisation-sustaining narratives. The Prime Minister’s office began training over 100 people throughout government positions, creating a depository of a “public chain of experts in Finland” trained in detecting and countering disinformation (Häkkinen, 2018). This example of a horizontal approach serves both to fight and detect media influencing and to raise awareness of the issue more broadly. Finland’s effective comprehensive security model, based on cooperation between government and civil society actors, has also aided in the fight against disinformation. The Geneva Center for Security Policy, grappling with the early tracing of hybrid influencing on Crimea, has praised Finland’s unique preparedness at the governmental and societal level (Cederberg and Eronen, 2015).

The appointment of the experienced diplomat Mikko Kinnunen as the first Finnish Ambassador for Hybrid Affairs has also contributed to creating a centre of institutional gravity for addressing the issue of disinformation in a more systemic fashion. Looking forward,

Jantunen believes that strategic communication will need to become more proactive: it has to become better at sending messages to fragmented online groups, particularly targeting the ‘margins’ rather than neglecting them (2018). In other words, strategic communications must also take notice of the final frontier of resilience: the margins within the microsphere. In so doing, the Finnish government is taking active steps to developing digital resilience, so that the connection between the digital propaganda promoted by foreign governments and the formation of unruly counterpublics could be diluted or broken. The concept of digital resilience advanced in this study builds on the idea of cognitive and physical resilience discussed in the existing literature but adds an important and novel perspective: it seeks to protect the public sphere against the threat of dissolution posed by the digital rise of ‘unruly’ counterpublics by taking a measured approach against the formation of cognitive and information channels that seek to undermine, deform and feudalize the public sphere.

CONCLUSION

This paper began as an investigation into the causes of instability in the current information sphere more broadly, and in Finland, more specifically. Building on recent research connecting digital propaganda to the disruption of the public spheres in democratic societies, the study argued that the logic of digital propaganda could be unpacked by examining the mechanisms that facilitate the diffusion of disinformation from the *microsphere* of the daily individual experience to the *macrosphere* of political decision-making. It has thus been argued that the connection between the two spheres is enabled by the formation and political mobilization of ‘unruly’ counterpublics, that is of arenas of textual and visual contestation of politically marginalized groups promoting issues aligned with the disinformation agenda. Four testable propositions have been advanced to probe empirically the effects of digital propaganda on the two components of the public sphere. The propositions were applied, using the within-case method of qualitative analysis, to four representative case studies reflecting the level of infiltration of digital propaganda from the micro- to the macro-sphere as a result of counterpublic mobilisation.

The empirical analysis has shown that the Finnish *macrosphere* is, overall, reasonably resilient to the effects of digital propaganda, but these efforts are at risk of being weakened by the rise of influential counterpublics in the *microsphere*. The assumption is that digital propaganda does not just create divides, but that it *requires* divides to succeed. To pre-empt the divides, Finland’s governmental institutions, in conjunction with civil society actors and journalists, have recognised that a society is only as resilient to disinformation as its most vulnerable segments. This paper reinforces the claim that the fissions are brewing outside of the macrosphere, in the fringes of the online world: in aggressively biased and dubiously accurate online magazines and

forums, the followers of whom are increasingly forming counterpublics whose narratives and discourses chip away at the foundations of a democratic, pluralist public sphere.

The study suggests that resilience should embrace a more systemic form that moves beyond narrow configurations of cognitive and physical resilience and seek to isolate the connection between digital propaganda and the formation of unruly counterpublics. From a digital resilience perspective, the public sphere requires enhancements that enables it to reliably validate truth-claims based on accurate information and rational engagement, contain the effect of emotional escalation, prevent the radicalisation of unruly counterpublics, and reinforce the integrity of the notion of the public good. To a certain extent, Finland has already started developing its own version of digital resilience. Rather than fact-checking every single narrative that arises, Finnish officials hope to instil a renewed sense of collective identity, one based not on nationalism or exclusion, but on education, social equality, and political stability.

In the words of Adam Berinsky, who helped Finland craft its own narrative: “you don't just want to say this is wrong and say this is right but you want to give people a reason to accept what is right; they have to accept what the truth is” (Yle Uutiset, 2016b). Carefully crafted communication campaigns now aim to instil a new ‘thickness’ to Finnish collective identity. The impact of this long-term initiative is hard to grasp, but the ideal outcome would be the development of digital resilience by reducing the appeal of disinformation at the micro- level. The Finnish case thus provides a useful and timely example about how Western countries should seek to enhance their resilience to the challenge of digital propaganda.

BIBLIOGRAPHY:

Aro, Jessikka. (2015). “My Year as a Pro-Russia Troll Magnet: International Shaming Campaign and an SMS from Dead Father,” *Yle Kioski* [Online]. Available from: <http://kioski.yle.fi/omat/my-year-as-a-pro-russia-troll-magnet> [Accessed: 29 May 2018].

- Agarwal, N., & Bandeli, K. K. (2018). Examining Strategic Integration of Social Media Platforms in Disinformation Campaign Coordination. *Defence Strategic Communications*. <https://doi.org/10.30966/2018.riga.4.6>.
- Bayer, J., Bitiukova, N., Bárd, P., Szakács, J., Alemanno, A., & Uszkiewicz, E. (2019). *Disinformation and propaganda-impact on the functioning of the rule of law in the EU and its Member States*. Retrieved from [http://www.europarl.europa.eu/RegData/etudes/STUD/2019/608864/IPOL_STU\(2019\)608864_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2019/608864/IPOL_STU(2019)608864_EN.pdf)
- Benkler, Y., Farris, R., & Roberts, H. (2018). *Network Propaganda*. <https://doi.org/10.1093/oso/9780190923624.001.0001>
- Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122–139. <https://doi.org/10.1177/0267323118760317>
- Bjola, C. (2010). The power of the public sphere: (Anti)-diplomacy and crisis management within security communities. In *Arguing Global Governance: Agency, Lifeworld, and Shared Reasoning*. <https://doi.org/10.4324/9780203842577>
- Bjola, Corneliu, & Pamment, J. (2016). Digital containment: Revisiting containment strategy in the digital age. *Global Affairs*, 2(2), 131–142. Retrieved from <http://dx.doi.org/10.1080/23340460.2016.1182244>
- Bjola, Corneliu, & Pamment, J. (2018). *Countering Online Propaganda and Extremism: The Dark Side of Digital Diplomacy*. Abingdon and New York: Routledge.
- Boghardt, T. (2009). Soviet Bloc Intelligence and Its AIDS Disinformation Campaign. *Intelligence*, 53(4), 1–24. Retrieved from www.cia.gov.
- Coles, B. A., & West, M. (2016). Trolling the trolls: Online forum users constructions of the nature and properties of trolling. *Computers in Human Behavior*, 60, 233–244. <https://doi.org/10.1016/j.chb.2016.02.070>
- Cull, N. J., Culbert, D. H., & Welch, D. (2003). *Propaganda and mass persuasion : a historical encyclopedia, 1500 to the present*. Retrieved from https://books.google.co.uk/books/about/Propaganda_and_Mass_Persuasion.html?id=Byzv7rf6gL8C
- Dawson, A., & Innes, M. (2019). How Russia's Internet Research Agency Built its Disinformation Campaign. *Political Quarterly*, 1467-923X.12690. <https://doi.org/10.1111/1467-923X.12690>
- European Commission. (2018). A multi-dimensional approach to disinformation. Retrieved May 30, 2019, from <https://publications.europa.eu/en/publication-detail/-/publication/6ef4df8b-4cea-11e8-be1d-01aa75ed71a1/language-en>
- European Council. (2018). Timeline - EU restrictive measures in response to the crisis in Ukraine - Consilium. Retrieved May 30, 2018, from <https://www.consilium.europa.eu/en/policies/sanctions/ukraine-crisis/history-ukraine-crisis/>
- EUvsDisinfo. (2018). Disinformation on Russian TV: Reported Chemical Attack in Syria and Skripal Poisoning "Western provocation" - EU vs DISINFORMATION. Retrieved May 30, 2019, from <https://euvsdisinfo.eu/disinformation-on-russian-tv-reported-chemical-attack-in-syria-and-skripal-poisoning-western-provocation/>
- Finnish Ministry of Foreign Affairs. (n.d.). *THE EFFECTS OF FINLAND'S POSSIBLE NATO MEMBERSHIP Contents*. Retrieved from https://um.fi/documents/35732/48132/the_effects_of_finlands_possible_nato_membership/c206b3c2-aaaa-5809-c545-7aa67c9bcb2a?t=1525861455616
- Fraser, N. (1992). Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy. In Craig Calhoun (Ed.), *Habermas and the Public Sphere*. (Cambridge, MA: MIT Press.

- Goertz, G., & Mahoney, J. (2012). *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. <https://doi.org/10.23943>
- Government Communication Service. (2018). *Counter-disinformation toolkit RESIST*. Retrieved from <https://gcs.civilservice.gov.uk/guidance/resist-counter-disinformation-toolkit/>
- Habermas, J. C. N.-H. . H. 1989 305. (1989). *The structural transformation of the public sphere: an inquiry into a category of bourgeois society*. Cambridge: MIT Press.
- Hansen, F. S. (2017). *Russian Hybrid Warfare: A study of disinformation*. Retrieved from http://pure.diis.dk/ws/files/950041/DIIS_RP_2017_6_web.pdf
- Higgins, A. (2016, April 16). EU Suspects Russian Agenda in Migrants' Shifting Arctic Route. *New York Times*. Retrieved from <https://www.nytimes.com/2016/04/03/world/europe/for-migrants-into-europe-a-road-less-traveled.html>
- Kalpokas, I. (2018). A Political Theory of Post-Truth. In *A Political Theory of Post-Truth*. <https://doi.org/10.1007/978-3-319-97713-3>
- Kim, J. (1994). Explanatory Knowledge and Metaphysical Dependence. *Philosophical Issues*, 5, 51–69. Retrieved from <https://about.jstor.org/terms>
- Lasswell, H. D. (1927). The Theory of Political Propaganda. *American Political Science Review*, 21(03), 627–631.
- Leighton, M. K. (1991). *Soviet propaganda as a foreign policy tool*. Freedom House.
- Mackenzie, W. (2018). What Finland Can Teach the West About Countering Russia's Hybrid Threats. Retrieved July 5, 2019, from World Politics Review website: <https://www.worldpoliticsreview.com/articles/24178/what-finland-can-teach-the-west-about-countering-russia-s-hybrid-threats>
- Mackintosh, E. (2019). Finland is winning the war on fake news . What it ' s learned may be crucial to Western democracy. Retrieved July 5, 2019, from CNN website: <https://edition.cnn.com/interactive/2019/05/europe/finland-fake-news-intl/>
- Nyberg, R. (2018). Hybrid Operations and the Importance of Resilience: Lessons From Recent Finnish History. *Carnegie Endowment for International Peace*, 4. Retrieved from <https://carnegieendowment.org/2018/02/08/hybrid-operations-and-importance-of-resilience-lessons-from-recent-finnish-history-pub-75490>
- Oxford Dictionaries. (2016). Word of the Year 2016 is... Retrieved May 30, 2019, from <https://en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2016>
- Pavel Astahov. (2014). Russian child ombudsman: Finland engaging in “juvenile terror.” Retrieved May 30, 2019, from Yle Uutiset website: https://yle.fi/uutiset/osasto/news/russian_child_ombudsman_finland_engaging_in_juvenile_terror/7631016
- Pomerantsev, P. (2015). Authoritarianism Goes Global (II): The Kremlin's Information War. *Journal of Democracy*, 26(4), 40–50. <https://doi.org/10.1353/jod.2015.0074>
- Pomerantsev, P., & Weiss, M. (2014). *The menace of unreality : how the Kremlin weaponizes information, culture and money*. Retrieved from http://imrussia.org/media/pdf/Research/Michael_Weiss_and_Peter_Pomerantsev__The_Menace_of_Unreality.pdf
- Posetti, J., Ireton, C., Wardle, C., Derakhshan, H., Matthews, A., Abu-Fadil, M., ... Mantzarlis, A. (2018). *Handbook for Journalism Education and Training UNESCO Series on Journalism Education*. Retrieved from <https://en.unesco.org/fightfakenews>
- Pynnöniemi, K. (2019). The Asymmetric Approach in Russian Security Strategy: Implications for the Nordic Countries. *Terrorism and Political Violence*, 31(1), 154–167. <https://doi.org/10.1080/09546553.2018.1555995>
- Ramsay, G., & Robertshaw, S. (2019). *Weaponising news RT, Sputnik and targeted disinformation*. Retrieved from <https://www.kcl.ac.uk/policy-institute/assets/weaponising-news.pdf>
- Renda, A. (2018). *The legal framework to address “fake news”: possible policy actions at the EU level*.

- <https://doi.org/10.2861/468200>
- Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature Communications*, 9(1), 4787. <https://doi.org/10.1038/s41467-018-06930-7>
- Sorrels, C. A. (1983). *Soviet propaganda campaign against NATO* /. Retrieved from <http://hdl.handle.net/2027/mdp.39015010252776>
- Stefan Meister. (2016). The “Lisa case”: Germany as a target of Russian disinformation. Retrieved May 30, 2019, from NATO Review website: <https://www.nato.int/docu/review/2016/also-in-2016/lisa-case-germany-target-russian-disinformation/EN/index.htm>
- StopFake. (2016). The Kremlin’s Voice: Johan Bäckman. Retrieved May 30, 2019, from <https://www.stopfake.org/en/the-kremlin-s-voice-johan-backman-2/>
- Szwed, R. (2016). *Framing of the Ukraine–Russia conflict in online and social media*. Retrieved from <https://www.stratcomcoe.org/framing-ukraine-russia-conflict-online-and-social-media>
- Tandoc, E. C., Jenkins, J., & Craft, S. (2019). Fake News as a Critical Incident in Journalism. *Journalism Practice*, 13(6), 673–689. <https://doi.org/10.1080/17512786.2018.1562958>
- U.S. Senate Committee on Foreign Relations. (2018). *Putin’s asymmetric assault on democracy in Russia and Europe: implications for U.S. national security* (Vol. 2018). Retrieved from <https://www.foreign.senate.gov/imo/media/doc/FinalRR.pdf>
- US Department of State. (1989). *Soviet influence activities :a report on active measures and propaganda, 1987-1988*. Retrieved from [Washington, D.C.?] : website: <http://hdl.handle.net/2027/uiug.30112028340161>
- Vilmer, J.-B. J. (2019). *The “Macron Leaks” Operation: A Post-Mortem*. Retrieved from Atlantic Council website: https://www.atlanticcouncil.org/images/publications/The_Macron_Leaks_Operation-A_Post-Mortem.pdf
- Walker, C., & Ludwig, J. (2017). *The Meaning of Sharp Power* (Vol. 2018). Vol. 2018. Retrieved from <https://www.foreignaffairs.com/articles/china/2017-11-16/meaning-sharp-power?cid=int-fls&pgtype=hpg>
- Woolley, S. C., & Howard., P. N. (2018). *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*. Oxford: Oxford University Press.
- Woolley, S. C., & Howard, P. N. (2018). *Computational Propaganda*. <https://doi.org/10.1093/oso/9780190931407.001.0001>
- Yle Uutiset. (2016). Border Guard: Finland could see 7,500 asylum seekers enter from Russia. Retrieved May 30, 2019, from https://yle.fi/uutiset/osasto/news/border_guard_finland_could_see_7500_asylum_seekers_enter_from_russia/8587431
- Zannettou, S., Caulfield, T., De Cristofaro, E., Sirivianos, M., Stringhini, G., & Blackburn, J. (2019). Disinformation Warfare: Understanding State-Sponsored Trolls on Twitter and Their Influence on the Web. *Companion Proceedings of The 2019 World Wide Web Conference on - WWW ’19*, 218–226. <https://doi.org/10.1145/3308560.3316495>

Interviews

- Häkkinen, V. (2018) Director, Current Affairs Communications, Finnish Ministry of Foreign Affairs, Helsinki, April 6.
- Hiltunen, I. (2018) Researcher in Journalism and Mass Communication, University of Tampere, Helsinki, April 19.

- Finnish Diplomat A (2018) Counselor/Department for Russia, Eastern Europe and Central Asia, Finnish Ministry of Foreign Affairs, Helsinki, April 17.
- Jantunen, S. (2018) Researcher, Finnish Defense Research Agency, Helsinki, April 17.
- Kinnunen, M. (2018) Ambassador for Hybrid Affairs, Finnish Ministry of Foreign Affairs, Helsinki, April 18.
- Standish, R. (2018) Journalist covering northern Europe and Baltics for The Atlantic, POLITICO, Foreign Policy, Helsinki, April 18.
- Tampere, P. (2018) Head of Communications, The European Center of Excellence for Countering Hybrid Threats – Hybrid CoE, Helsinki, April 20.