Orban, Franck [Hrsg.]; Strand Larsen, Elin [Hrsg.]; Lindahl, Sondre [Hrsg.]

Trust under threat. Challenges in a digital society

Münster ; New York : Waxmann 2024, 200 S.

Quellenangabe/ Reference:
Orban, Franck [Hrsg.]; Strand Larsen, Elin [Hrsg.]; Lindahl, Sondre [Hrsg.]: Trust under threat. Challenges in a digital society. Münster ; New York : Waxmann 2024, 200 S. - URN: urn:nbn:de:0111-pedocs-332735 - DOI: 10.25656/01:33273; 10.31244/9783830998532

https://nbn-resolving.org/urn:nbn:de:0111-pedocs-332735
https://doi.org/10.25656/01:33273

Franck Orban, Elin Strand Larsen, Sondre Lindahl (Eds.)

Trust under Threat

Challenges in a Digital Society

WAXMANN

Franck Orban, Elin Strand Larsen, Sondre Lindahl (Eds.)

Trust under Threat

Challenges in a Digital Society

Waxmann 2024
Münster • New York

The editors would like to thank Østfold University College (Norway) for funding this book.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.dnb.de

Print-ISBN 978-3-8309-4853-7
E-Book-ISBN 978-3-8309-9853-2
https://doi.org/10.31244/9783830998532

Waxmann Verlag GmbH, Münster 2024
Steinfurter Straße 555, 48159 Münster, Germany
www.waxmann.com
info@waxmann.com

Cover Design: Anne Breitenbach, Münster
Cover Picture: spainter_vfx, Shutterstock
Typesetting: Mario Moths, Marl

In memory of Marjo-Riitta Rynning

Contents

Trust under Threat: Challenges in a Digital Society. Introduction

Franck Orban, Elin Strand Larsen and Sondre Lindahl

Trust is one of those words that you recognize and have some intuitive understanding of what it entails. You know people you trust, and you know people you do not trust. You may also trust your bank, but not the car shop with a bad reputation. You may also, perhaps, trust the judicial system but not the politicians. Finally, you may trust the dating app you are using, but what about meeting a person that you have only talked to online? How can you ascertain whether or not they are to be trusted?

We encounter issues of trust, as it were, on a daily basis. Trust, in fact, is an integral part of human life and we all know the value of trusting someone or some institution. Yet, trust has been viewed as a somewhat mystical and intangible factor, something that defies careful definition (Giffin, 1967). There is not a single definition of trust; agreement exists on the fact that trust is a multi-dimensional concept that facilitates cooperative endeavors (Deutsch, 1973; Gambetta, 1988). It is also a key to positive interpersonal relationships and interactions between individuals or groups of individuals, as it can be crucial in periods of uncertainty or crisis to overcome challenges in a collective manner. The bottom line is that trust is essential to any form of partnering (Rackham, Friedman & Ruff, 1996). In that sense, the study of trust might be more important today considering the decline of social capital in western societies. Ethnic conflicts and hate crimes, as well as the vanishing of traditional moderate governing parties, are major sources of concern when they increase polarity and fragment trust toward individuals, institutions, or states (Putnam, 2000).

The UN Secretary General Antonio Guterres argued in the 74th Session of the UN General Assembly on September 25, 2018, that:

> "Trust is at a breaking point. Trust in national institutions. Trust among states. Trust in the rules-based global order. Within countries, people are losing faith in political establishments, polarization is on the rise and populism is on the march. Among countries, cooperation is less certain and more difficult. Divisions in our Security Council are stark" (Guterres, 2018).

This description has been thrown into sharp relief with the war in Gaza which began as a consequence of the terrorist attack perpetrated by Hamas on October 7th, 2023. The attack was brutal, but Israel's response has been so brutal and extensive that the UN court at one point deemed it plausible that a genocide could occur unless Israel would change course. This situation may change rapidly, as things are wont to do in international relations. Nonetheless, at the current socio-historical juncture we are witnessing a loss of trust in the UN, and various affiliated branches, but also in the rules-based global order. Institutions that were put in place after WW2 to make the world more peaceful are being undermined, and soon they may not be as effective, or even exist, due to the lack of trust.

One common perspective on trust differentiates trust on different levels. For example, first a strictly personal approach where one person trusts another specific person (Rotter, 1980). Second, an interpersonal approach where two or more people or groups of people trust each other. Third, an organizational, institutional, or societal approach where trust is ultimately based upon social, cultural, or institutional structures (Shapiro, 1987). Importantly, different definitions of trust continue to be used according to different interdisciplinary disciplines, as they focus on different aspects. Definitions may consequently range from a personal or interpersonal trait to a more structural phenomenon that characterizes groups of people and societies. This book is in that sense by no means an exception to the rule and presents a wide variety of approaches to the notions of trust, distrust and mistrust.

Another common way to categorize trust is to make a distinction between generalized trust and particularized trust (or mistrust in both cases). In the first case, trust is defined as trust in other people in general or trust in strangers (Uslaner, 2002; Delhey et al., 2011). Generalized trust is characterized by the absence of both a specified recipient of trust and a specific regard in which the recipient is trusted. It is a propensity to trust other people that varies between countries, based on religious and ethnic composition, inequality, quality of government, and welfare regime (Larsen, 2013). Particularized trust is, on the contrary, based on more specific knowledge of and close contact with others (Newton & Zmerli, 2011).

The distinction between particularized and generalized trust can also be linked to Francis Fukuyama's theory of low-trust and high-trust societies. The author argues that strong family relations and less trust in people outside the family characterize low-trust societies. In high trust societies people tend to trust science, political authorities, or traditional media to a greater extent (Fukuyama, 1995). According to the European Social Survey (ESS), trust in the Nordic countries is high compared with other European countries (European Social Survey, 2014, 2022; Listhaug & Ringdal, 2008). The standard model to explain variations in political trust at the individual level includes two main categories of independent variables: political distance and performance evaluations. We can expect that an increasing distance between the government and the citizens will lead to a decline in trust. The level of trust is also based on how well the government is able to fulfill the goals that citizens agree on (Listhaug & Ringdal, 2008).

As such, trust is not a static concept. As our societies change, appropriate conditions for giving and receiving trust evolve over time. Modern information technology in a globalized society has little in common with past homogenous agrarian communities before or with early industrial urban society (DeVries, 2011). Modern societies are increasingly interconnected with each other A large part of this interconnection is due to digitalization. Digitalization and digital transformation in the 21st century build on revolutions in information technology like computerization in the 1960s, data processing in the 1970s, personal computing in the 1980s and, finally, internet computing in the 1990s. As Osburg and Heinecke mention in their book from 2019 (Media Trust in a Digital World), the digital transformation opens traditional societies and frees new energies. It also creates distortions and a fragmentation of trust, along with processes like globalization and individualization. Technology is therefore both the problem and the solution, as it fuels new ways of creating and apportioning trust.

While digitalization represents the contemporary modern society, trust is, as mentioned, an old concept that is of a more analog character. What happens, therefore, to trust when our societies are becoming more and more digitalized? Botsman (2018) suggests that trust will evolve into *distributed* trust in a digital society. Distributed trust is a kind of trust that flows laterally between individuals, programs and even bots, and is enabled by networks, platforms, and systems. Dealing with a chat bot on different websites is not a novelty today, and more and more companies and institutions use chatbots to provide information about banking, software issues, or just to get in touch with "customer service". In early 2023 the company Open AI made headlines with ChatGPT, its state-of-the art chatbot. There is no doubt that ChatGPT is an incredibly powerful tool, and in academia the alarm was raised when it became clear that a lot of students fully embraced the technological advancement. Interestingly, a lot of people were quick to trust the bot to write papers, summarize information and, quite basically, do the job in a satisfactory manner. This is an example of what Botsman means with distributed trust, and it illustrates nicely how trust is a dynamic concept that will change and evolve with time.

At the heart of this anthology is an interdisciplinary exploration of the concept of trust. We welcome the reader to join us as the chapter explores, analyze and discusses the concept of trust, and the challenges it faces in a digital society.

In the first chapter, HARALD BORGEBUND and SONDRE LINDAHL explore how increased digitalization challenges and replaces the traditional notions of trust by changing the places where people meet, socialize, and build trust towards each other and to democracy. Drawing on social capital as a central concept, as well as the works of Putnam, Botsman and Hardin, Borgebund, and Lindahl analyze the implications for trust in an increasingly digital society. The authors emphasize the difficulty of drawing any certain conclusions about the consequences an increasingly digital society might have on trust. Although trust in institutions has declined over the last few decades, the negative consequences are hard to attribute with any kind of certainty. The only thing certain is that "discussions about the role of trust in modern societies will continue" (Borgebund & Lindahl's chapter, p. 17–28, here p. 27).

HÅVARD FRIIS NILSEN discusses the importance of social trust in classical republican thought, with a focus on Stanley Kubrick's *2001 – A Space Odyssey*. In this science-fiction movie, which came out in 1968, we meet an AI computer for the first time, decades before the technology even was available. According to Nilsen, the miniature society onboard the spaceship in Kubrick's 2001, where the AI computer HAL runs everything, conflicts with the republican ideal of a good society based on social trust. The author argues that "Kubrick's depiction of autonomous technology is the most accurate description and prophecy of the problems of artificial intelligence we are grappling with today" (Nilsen's chapter, p. 29–41, here p. 38) and sums up the chapter by emphasizing the need for a cautious approach to ensure AI does not surpass human control.

KAI A. HEIDEMANN considers how the rise of digital information communication technologies (DICTs) has impacted the mobilization of trust and distrust in social movements. With a focus on "progressive" justice-oriented social movements, Heidemann highlights both the positive characteristics of digitalization, such as enhanced networking, mobilization, and visibility, as well as negative dynamics of repression and control. According to Heidemann, DICTs can be used as especially effective tools of counter-hegemonic communication for social movements, citing radical alternative media as a key example. The author sums up his analysis by describing the digital revolution's impact on social movements as "a double-edged sword that offers both unparalleled opportunities for empowerments and agency, while simultaneously posing formidable challenges in the form of surveillance, censorship, and manipulation" (Heidemann's chapter, p. 43–62, here p. 59).

ELIN STRAND LARSEN's chapter focuses on Scandinavian fact-checking, trust building and handling of misleading information related to the Israel-Hamas conflict. Through interviews with three fact-checking organizations, as well as some qualitative examples, Larsen investigates how fact-checkers in Norway, Sweden, and Denmark manage fake news and disinformation related to an escalating conflict and war far away from home. In the conclusion, the author lists five main strategies of fact-checking aimed at building trust between the fact-checkers and their audience: They are honest about uncertainty, use carefully written conclusions, provide proof, have an open dialog with the public, and rely on the network of other fact-checkers. These trust-building strategies are especially important "in situations with a faraway conflict and war, with little or no access to the area, like we see in the Israel-Hamas conflict" (Larsen's chapter, p. 63–81, here p. 78).

SAMBA DIALIMPA BADJI and NELSON BAHATI's chapter explores Malian fact-checkers' perceptions of audience trust and their trust-building strategies. Based on interviews with editors from five fact-checking organizations in Mali, Badji and Bahati argue that the polarization of the society in Mali is a challenge for fact-checkers trying to convince the audience of the trustworthiness and usefulness of their work. They also find that trust-building is incorporated into the way fact-checkers engage with their audience. The authors find that building audience trust is incorporated into the way the fact-checkers work, "through their commitment to independence and transparency as well as their permanent engagement with their audiences" (Badji & Bahati's chapter, p. 83–100, here p. 94).

KIRSTI K. COLE and JOHANNA M. WAGNER argue in their chapter that the digital footprint of Crisis Pregnancy Centres (CPCs), that have burgeoned across the US since the latter 20th century, are based on misinformation, deceptive facts and figures, as well as outright lies. CPCs are nonprofit organizations that provide free services to women who are considering terminating their pregnancies, but often their more undisclosed mission is to convince women not to have abortions. By analyzing 18 CPC websites, Cole and Wagner show how CPCs through text, pictures, and web design keep women from accessing abortions, and the authors argue that "the combination of disinformation" and fabrication of "[medical] authority continues to ground the contemporary war on women" (Cole & Wagner's chapter, p. 101–119, here p. 115) in the US.

ANJE MÜLLER GJESDAL with her co-authors MARIE CHANDELIER, CLAIRE DOQUET, ØYVIND GJERSTAD and CÉLINE POUDAT analyze how trust and distrust in expert knowledge regarding climate change is expressed in the digital, collaborative encyclopaedia Wikipedia. Using the analytical framework of Representation of Other's Discourse (ROD) – developed by Jacqueline Authier-Revuz – they find that Norwegians tend to cite the sources more indirectly than French contributors to Wikipedia, and that scientific sources are more positively evaluated by French Wikipedians than Norwegians when discussing climate change. According to Gjesdal et al., both climate sceptics and non-sceptics have a "shared, fundamental trust in science, but a disagreement over the status of credible sources" (Gjesdal et al., chapter, p. 121–137, here p. 135).

MARIE STØREN JAREID's chapter looks into how narrative contestation between the United States and China plays out in the social sphere of TikTok. TikTok is an app of Chinese origin that has swiftly become one of the world's largest social media platforms, and Jareid explores the most popular hashtags related to US-China relations through qualitative content analysis. While previous research on political engagement on TikTok often show oppositional narratives to the contemporary political paradigm, the author of this chapter finds that popular political content on TikTok concerning the Sino-American relationship works to amplify and strengthen the current narratives of distrust held by the two governments. In conclusion, Jareid argues that "the reiteration of frames and narratives on the highly personalized feeds [on TikTok] may contribute to amplifying the trust deficit created by both nations by transferring it from a geopolitical level to politically engaged social media users" (Jareid's chapter, p. 139–156, here p. 151).

FRANCK ORBAN's chapter investigates the Foreign Cyber Electoral Interference (FCEI) in the 2017 and 2022 French presidential election. According to Orban, the electoral process in 2017 suffered three digital strikes – the hacking of several institutions and political parties, the spreading of fake stories and the so-called "Macron leaks", causing several counter-movements by the French authorities in the lead up to the 2022 presidential election. The counter-movements included developing military and civilian capacities, passing laws to prevent the spread of misinformation, and promoting media literacy within educational institutions. Even if foreign attempts to disrupt the election had limited impact in 2022, Macron still lost. In the conclusion, Orban reminds us that "even if national and international measures to counter foreign cyber electoral interference are crucial in maintaining electoral integrity, they alone cannot rectify the deep-seated mistrust in political figures and institutions that can jeopardize democratic dialogue" (Orban's chapter, p. 157–172, here p. 168).

LAILA BERG addresses the link between political trust and digital technology in contemporary Scotland with a focus on digital participation in the Scottish independence referendum in 2014. Berg is particularly interested in the growing salience of Scottish national identity and nationalism in combination with the increased digitalization of the Scottish society. Based on statistical data from the Scottish Social Attitudes (SSA), as well as 40 qualitative interviews conducted during fieldwork in Scotland in 2018, the author finds that social media and digital platforms were used to gain the public's support for the issue of Scottish independence and to increase the public's levels of trust in an alternative vision of Scotland. In summary, Berg stresses the importance of digital participation in the Scottish referendum: "If people feel that they have agency and impact on their political system, then they are more likely to have trust in these political systems and the people who run the country" (Berg's chapter, p. 173–187, here p. 185).

In the last chapter of the anthology, WLADIMIR CHÁVEZ, studies the novel *Nefando* [Heinous] with a focus on the level of trust that the characters feel towards authority and institutions. The novel revolves around an online video game from the deep web. Drawing on the works of scholars like Tait, Ferdous and Al-Hossienie, Christopherson, and Gehl, Chávez analyzes the novel's protagonists and places based on the fascination for violence in the digital era, the anonymity and autonomy of the internet, as well as distrust in the government and trust in the deep web. According to the author, "in Nefando it seems that the digital society has taken the place of the real one, without necessarily becoming an improved version" (Chávez's chapter, p. 189–198, here p. 196). In conclusion, Chávez argues that while the internet and dark web can provide anonymity for reporting criminal offences without censorship, strong digital policies put in place by the government are still needed to protect people's data privacy and online human rights.

This book is the third anthology of AreaS, a transdisciplinary group of researchers located at Østfold University College in Norway, as well as contributions from our partner scholars at other universities in Norway and abroad. The first anthology of AreaS was published in 2019 and focused on borders (Orban & Larsen, 2019), the second in 2022 with an interest in alliances (Orban & Larsen, 2022), while the third and last anthology of AreaS looks into the notion of trust in a digital society. We want to thank all our researchers for their contributions, as well as all the peer reviewers for taking the time to give high quality feedback. All chapters of this book have been reviewed and evaluated through double blind peer review conducted by experts in the different fields, ranging from political science to media, information technologies, literature, and sociology.

References

Botsman, R. (2018). *Who can you trust? How Technology Brought Us Together and Why It Could Drive Us Apart.* New York: Penguin Business.

Delhey, J., Newton, K., & Welzel, C. (2011). How General Is Trust in "Most People"? Solving the Radius of Trust Problem. *American Sociological Review,* 76(5), 786–807. https://doi.org/10.1177/0003122411420817

Deutsch, M. (1973). The Resolution of Conflict: Constructive and Destructive Processes. *American Behavioral Scientist,* 17(2), 248–248. https://doi.org/10.1177/0002764273 01700206

DeVries, W.A. (2011). *Some Forms of Trust.* Inf., 2, pp. 1–16. https://doi.org/10.3390/info 2010001

European Social Survey. (2022). *Exploring public attitudes, informing public policy.* Retrieved from https://www.europeansocialsurvey.org/docs/findings/ESS1_9_select_findings.pdf

European Social Survey. (2014). *European Social Survey.* Retrieved from https://www.europeansocialsurvey.org/

Fukuyama. (1995). *Trust*: *the social virtues and the creation of prosperity.* Free Press. XV.

Gambetta, D. (Ed.) (1988). *Trust: Making and breaking cooperative relations.* Oxford: Blackwell.

Giffin, K. (1967). The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. *Psychol Bull, Aug.;* 68(2), 104–120. https://doi.org/10.1037/h0024833

Guterres, A. (2018). Secretary-General's Address to the General Assembly, 25. september 2018. https://www.un.org/sg/en/content/sg/statement/2018-09-25/secretary-generals-address-the-general-assembly-delivered-trilingual-scroll-further-down-for-all-english.

Larsen, C.A. (2013). *The Rise and Fall of Social Cohesion: The Construction and Deconstruction of Social Trust in the US, UK, Sweden and Denmark.* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199681846.001.0001

Listhaug, O., & Ringdal, K. (2008). Trust in political institutions. In H. Ervasti, T. Fridberg, M. Hjerm & K. Ringdal (Eds.), *Nordic Social Attitudes in a European Perspective* (p. 131–151). Edward Elgar Publishing. https://doi.org/10.4337/9781848444997.00013

Newton, K., & Zmerli, S. (2011). Three forms of trust and their association. *European Political Science Review,* 3, 169–200. https://doi.org/10.1017/S1755773910000330

Orban, F., & Strand Larsen, E. (2019). *Discussing Border, Escaping Traps: Transdisciplinary and Transspatial Approaches.* Waxmann. https://doi.org/10.31244/978383099 0451

Orban, F., & Strand Larsen, E. (2022). *Living Alliances, Leaving Alliances: Interdisciplinary Perspectives.* Waxmann. https://doi.org/10.31244/9783830994497

Osburg, T., & Heinecke, S. (2019). *Media Trust in a Digital World: Communication at Crossroads.* Cham: Springer. https://doi.org/10.1007/978-3-030-30774-5

Putnam, R.D. (2000). *Bowling Alone – the Collapse and Revival of American Community*. New York: Simon & Schuster Paperbacks. https://doi.org/10.1145/358916.361990

Rackham, N., Friedman, L., & Ruff, R. (1996). *Getting partnering right: How market leaders are creating long term competitive advantage*. New York: McGraw-Hill.

Rotter, J. B. (1980). Interpersonal Trust, Trustworthiness, and Gullibility. *American Psychologist,* 35, 1–7. https://doi.org/10.1037/0003-066X.35.1.1

Shapiro, S. P. (1987). The Social Control of Impersonal Trust. *American Journal of Sociology,* 93(3), 623–658. https://doi.org/10.1086/228791

Uslaner, E. M. (2002). *The moral foundations of trust*. Cambridge University Press. https://doi.org/10.2139/ssrn.824504

Trust and Democracy in a Digital Society

Harald Borgebund and Sondre Lindahl

Introduction

In his famous book *Bowling Alone* from 2000, Robert Putnam states that trustworthiness lubricates social life. This lubrication is facilitated through frequent interaction among people who are different in many ways. These interactions, according to Putnam, create a norm of generalised reciprocity which, when embedded in dense networks of social interaction, reduces incentives for opportunism and malfeasance. In short, trust and trustworthiness are important aspects of a well-functioning society.

The society Putnam wrote about at the outset of this millennium has no doubt changed a great deal. The biggest change, perhaps, is the rapid development of the internet which has changed how trust is built and maintained. One big aspect, and consequence, of increased digitalisation in our societies is how more and more services and interactions are carried out using digital tools and in digital spaces. There is frankly not the same need to meet people to do banking, shopping, or even education.

Putnam asserted that trust and trustworthiness lubricate social life, but is that necessarily so in a digital society? In this chapter, we explore the utility of trust by investigating how increased digitalisation challenges and replaces the traditional notions of trust by changing the places where people meet, socialise and build trust towards each other, but also trust in institutions. More specifically, we are interested in how trust affects democracy conceptually but also in practice. It can be argued that democracy itself represents a kind of institution that, at least for Western countries, has tremendous value. At a time when the rate of democracies in the world is actually experiencing decline, investigating the relationship between trust and democracy is important. Democracies, however, come in many different variants and this chapter focuses more specifically on trust and democracy in Norway. There are several reasons for focusing on Norway, chief among them is the fact that Norway has really embraced the digital revolution and ranks high on the list of global digital competitiveness. Norway also ranks at the top in terms of trust, between people and between people and institutions (Direktoratet for forvaltning og økonomistyring, 2021). While this chapter is not a case study of Norway per se, we use it to contextualise our discussion of the utility of trust in a digital society.

Furthermore, we draw on social capital as a central concept, and we use it to frame the discussion of the utility of trust in a digital society. We then connect this discussion with the concepts of local, institutional and distributed trust as understood by Rachel Botsman to analyse how an increasingly digitalised society moves from local and institutional trust towards an era of distributed trust.

Building on the insights from both Putnam and Botsman, we discuss Russell Hardin's view on trust which challenges both Putnam and Botsman's theories on trust. Relying on insights from these three understandings of trust we analyse some of the implications for trust in an increasingly digital society.

What is Trust?

A common definition of trust understands it as "the expectation that another person (or institution) will perform actions that are beneficial, or at least not detrimental, to us regardless of our capacity to monitor those actions…so that we will consider cooperating with him [the institution]" (Sapienza & Zingales, 2012, p. 124). The definition highlights cooperation, the inability to monitor other's actions *ex ante* and the probabilistic nature as the three key elements of trust.

Rachel Botsman, argues that trust can be divided into three distinct chapters (Botsman, 2018, p. 7). The first is local trust. This trust existed within the boundaries of small communities, where everyone knew everyone else. In other words, trust exists between members of small communities, and crucially, rests in someone specific. This can be tied to the famous study by Robin Dunbar which uncovered that the number of people the average person could have in their social group was hundred and fifty. That is to say, these would be casual friends that you know well enough to invite to a party. The most intimate group is the close support group which consists of five people. These would be your best friends and family. Dunbar also found that the social groups can extend to fifteen hundred, which is the absolute limit for whom you can put a name to a face. The point is that local trust, the kind that rests in someone specific, is limited to small communities which, in large parts of the world, no longer exist.

The second form of trust is institutional trust. Botsman describes this as flowing upwards to leaders, experts, and brands, and it runs through institutions such as courts, corporations and regulatory bodies. Importantly, this kind of intermediated trust helped establish the foundation for an organized industrial society, for example by placing your savings in a bank, thus trusting the bank to safeguard the money and not gamble it on horse races. This move lowers transaction costs for people given that they no longer have to worry constantly whether their money is safe. It also means that maintaining trust becomes important for these institutions and a currency in its own right. With the election of President Trump in 2016, the declining trust in political institutions in the USA became obvious, not least because Trump partly campaigned on a distrust in the political institutions.

The third form of trust is distributed trust, which Botsman argues is still in its infancy. The key aspect of distributed trust is that it flows laterally between individuals, and is enabled by networks, platforms and systems. This means that trust could flow to other people but also programs and bots. Recent examples of this could be chatbots that we use to get information about banking, and software issues or just to get in touch with "customer service". In early 2023 the chatbot ChatGPT made headlines around the world because it showed how advanced these bots have become. Students got better than average grades when they submitted papers written by the bot, but the interesting aspect here is that a lot of people were quick to trust the bot to write papers, summarise information and, quite basically, satisfactorily do the job.

Another element of distributed trust can be found in the growth of the sharing economy. A lot of people are happy to use services like Uber, Lyft and BlablaCar to get from A to B, although it does entail jumping into cars with strangers. One way to work around this obvious barrier, which in part helps explain why it is an omnipresent feature in a digital society, is the rating regime. That is, we now score and rate pretty much every service we use. Whether it is the Uber driver or the Italian restaurant on the corner, we rate our experience and share it with others. Establishments will also often ask to be reviewed because they know it can increase their visibility on Google Maps and other search engines. Crucially, this rating regime makes it easier for us to trust that the restaurant with 5 stars will be a safe and excellent choice, or that the Uber driver will not rob you. In these examples, trust does not flow upwards but horizontally between people and establishments.

Thus, it is obvious that trust can be found on different levels and in different degrees. Although we now more frequently put our trust in bots, systems and ratings, we still live in complex and dynamic societies where trust exists and functions on various levels. Also, as we mentioned above, trust does not exist in a vacuum, but it is intimately connected with the concept of social capital. Indeed, it can be argued that the utility of trust in a digital society depends on social capital. In the following paragraphs we explore the concept of social capital, and why it is important when discussing trust.

Trust and Social Capital

In the 1980s one of the authors of this chapter remembers his parents going to the bank to pay bills, which often meant queuing for a long time in the local bank in the village. With only one bank available most people had to go there to pay bills. It was not unusual for long queues, and many took the opportunity to chat with the other villagers. In essence, this was a social arena where everyone encountered people with different social and economic statuses and where they engaged with others with different backgrounds. Gradually, phone banking and then Internet banking made the need to visit the bank unnecessary. Now, forty years later the local bank branch has been closed and people rarely actually visit the bank anymore.

This change in banking practices also means that the bank as an arena where people met and interacted with people with a variety of different backgrounds disappeared. The informal interactions like chatting with someone while queuing at the bank provided an opportunity for people to build relationships with other people. According to Robert Putnam, establishing such relationships had the side effects of developing attitudes and notions that were beneficial for the wider society, and for developing social capital. This latter term he defined as "features of social organization, such as trust, norms, and networks, that can improve the efficiency of society by facilitating coordinated action" (2000, p. 167). More specifically, he is talking about the various networks of relationships among people who live in a particular society. Social capital highlights how interpersonal relationships, a shared sense of identity, a shared worldview, shared norms and values, trust and cooperation help bring about the effective functioning of societies. Inefficient social practices, such as personal banking for the individual, could on a societal level contribute to more efficiency in the economy and politics. Constituting and enhancing social capital is, according to Putnam, dependent on such arenas where people interact and learn to interact with people from different backgrounds. Without such arenas, there are fewer opportunities to build such relationships and hence the social capital is stagnating or deteriorating.

Social capital is also an important element for the utility of trust. We know that higher levels of trust contribute to making everyday life go smoother and easier as people can make decisions about shopping, business and other important issues without having to carefully check the persons they interact with. This is because they assume that most people with whom they interact can be trusted. Here we see that the various ingredients as set out in the definition, namely cooperation, probability and the inability to monitor other people's actions ex ante in play.

In politics, trust is important too. Without trust in politicians and political institutions, voters might decide to not vote or be politically active. Deteriorating voting levels may eventually undermine democracy. The short story of how banking was done in a small town in Norway in the 1980s is an example of how social capital was built (and maintained) when people met physically and interacted with each other, which indirectly also helps build trust between individuals but also the institution. The importance of trust for a well-functioning society has been acknowledged for a long time. For example, in his analysis of the development of democracy in the US in the first half of the 19th century, Alexis de Tocqueville noticed how engagement and participation in civil society created trust and ensured wide participation in politics and the wider society (2002 [1835]). Compared with a Europe reeling from the aftermath of the French Revolution, and rising tension between social groups caused by rapid industrialisation, the development of the newly minted democracy in the US was remarkable at the time. Similarly, Almond and Verba in a famous study analysed the striking social differences between the Northern and Southern regions of Italy (1963). In particular, they noticed that the lack of trust in the South correlated highly with weak economic development compared with the Northern regions. Italy's Southern regions were on a par with many developing nations, while the Northern regions were at a similar development level as many of the most developed nations in Europe.

Additionally, Putnam analysed the introduction of regional political reforms in Italy investigating if the introduction of the same political institutions and decision-making procedures would reduce some of the differences between the North and the South. Studying the reforms from their introduction in the early 1970s until the early 1990s, Putnam concluded that the politically dysfunctional South did not change much after the reforms (1994). For example, it could take 8-9 months longer to pass the annual budget in some of the Southern regions compared with the Northern regions. Such delays are serious in themselves but have serious ramifications for the rest of society as investment decisions are put on hold, and it adds to the general feeling of distrust and eroding social capital.

Putnam argued that the differences between the North and South in Italy could be traced back to the Middle Ages. The economic and political differences between the North and South could be explained by a relatively hierarchical society in the Southern part of Italy, whereas the North enjoyed a more horizontal social structure related to the many city-states where trade and interaction between groups from different social classes created mutual trust and social capital through frequent interaction. These empirical studies show that trust and social capital may have long-term consequences for society, but they can erode fast under certain conditions. Consequently, trust and social capital may have widely different implications for different societies. These differences make it even harder to study the concept of trust in modern democracies.

Turning back to our point of departure, the question then is what happens to trust in a democracy when these traditional arenas are replaced by the rapid developments in a digital society? We understand digital societies as societies where a wide range of services and products rely on the use of the Internet, AI, and other forms of electronic communication and technology. Examples range from social media platforms to banking services and interactions with government agencies and corporations. A consequence of increased digitalisation is that personal relationships are replaced with interactions with strangers or with computers. The consequences are not only limited to our relationships but affect also politics and economics. Social media is increasingly used in political communication by politicians and cryptocurrencies have attempted to challenge regular currencies. Furthermore, AI through ChatGPT and other services are changing how we understand knowledge. Many of these technologies are only in their infancy and are likely to develop into even more powerful tools in the future.

The utility of personal banking, as outlined above, was that various people got to build relationships which subsequently built social capital and trust. It is an example of institutional trust that Botsman discussed, and it is assumed that this kind of trust is important for democracies. According to Putnam (2000), when arenas like these disappear, they disrupt the fabric of society. Such disruptions undermine social capital which, in turn, is vital for many aspects of a well-functioning democracy. It is not that we not interacting with other people anymore, but that we are doing so in ways that do not have the benefit of creating trust and social capital. For example, people spend more and more time using digital devices and operating in digital spaces and arenas. Sometimes it is hard to detect just how profound these changes have been, perhaps because humans are disposed to adapt to changes in their environment, and most of us welcome the changes that digitalisation has brought about.

The digitalisation of society has not happened in a vacuum, however. Since the 1960s, when trust in politicians and democratic institutions reached its peak, there has been a steady increase in social and political polarisation across many countries. Trust in government is lower in general, and although the erosion of trust and social capital started before the recent digitalisation of society, there are reasons to suggest that digitalisation might amplify the erosion of trust. For example, the story about going to the bank to pay bills highlights the creation of social capital as people met acquaintances which extended and facilitated social interactions between different types of people in the process. These interactions are the types of interactions described by Putnam as valuable because they give rise to trust and social capital through frequent interactions between different social groups. Today Internet banking has replaced personal banking, and the convenience of using a smartphone has replaced these social interactions. The bank is no longer an institution with familiar faces, a personal connection. It is a large and complex institution that is owned by anonymous shareholders, and with which you interact through their software. It seems apposite, therefore, to consider whether digitalisation is contributing to eroding trust, and subsequently what will happen with levels of trust in democracies.

Elusive Understandings of Trust and its Utility

Although Botsman and Putnam's understandings of trust are influential, they have also attracted criticism. For example, Russell Hardin (2006) challenges the understanding of trust explained above. He argues that theorists such as Putnam and Botsman 'focus on trust, norms and networks, all of which seem to be at the individual-level. But, for them the central concern is with how individual-level factors *facilitate the working of institutions*, including the whole of government' (2006, p. 77, original emphasis). Hardin asserts that social capital theorists, like Putnam, argue that actions by the individual such as participating in sports clubs, queuing at the bank and similar activities make the government work better. What is problematic about this assumption according to Hardin is that the causality of the relation between individual action and governmental workability is too loose and primarily based on correlations (2006, p. 78). The main correlation Putnam relies on is a simultaneous decline in trust in others (including strangers) and trust in government ibid. A decline in trust has occurred together with a decline in participation in civil society organisations, and the inference drawn by Putnam is that reduced participation in civil society organisations is the reason behind the decline in trust. There are at least two problematic features with this argument according to Hardin. Firstly, it is unclear how individual-level participation leads to a more well-functioning government. Secondly, Putnam's argument is based on a correlation and is not a causal explanation of how social capital leads to a more well-functioning government.

Furthermore, political trust is harder to enforce and police than social trust according to Hardin (1996). While social trust is based on ongoing relationships where trust can be built or reduced based on our experiences, political trust differs because voters cannot easily assess whether the politicians fulfilled their promises or not. Many election pledges can be interpreted in different and often contradictory ways, making it hard to know whether a policy promise has been fulfilled or not. For example, a politician can promise during an election campaign to be tougher on illegal immigration. What being tougher on illegal immigration actually means is subject to interpretation. Because politicians often use ambiguous and unclear language open to various interpretations, voters are unable to assess properly whether the politicians are trustworthy or not. Another important difference between social trust and political trust is that most of our social and economic relationships because we encounter some of the same persons and businesses regularly. If we experience that our local supermarket sells food that has already expired but that has been repackaged we are likely to switch to another supermarket. Because we buy food from supermarkets on a regular basis we can police the supermarkets' practices and performance continuously. In politics we normally vote once every fourth or fifth year, which means that the politicians performance is not judged often enough to create an ongoing relationship between the voters and the politicians. For the politicians this means that there is a fairly long period of time when they are "safe" and do not have to be considered trustworthy because the voters can not recall their support for a considerable time.

Thus, some of the structural differences between social and political trust can explain why social and political trust are not analogous. Despite the structural differences between social and political trust, some recent analyses of the relationship between them suggests that there is a causal relationship between social and political trust where high social trust can contribute to higher political trust (Bargsted et al., 2023). Specifically, Bargsted et al. find that "empirical results support the institutional view claiming that political trust has a positive effect over social trust. But we also find that, in Chile, social trust affects political trust" (2023, p. 1403). Bargsted et al. and Hradin's perspective on the difficulty of evaluating the government in modern societies points to some of the features that makes trust in modern societies somewhat elusive.

Modern democracies are based on a principle of legitimacy through which the people choose political leaders and hold the political leaders to account by voting out political parties whose jobs they disapprove of. Although this is a simplified understanding of modern democracies, it suffices to point towards the role of trust in modern democracies because if the people experience that they do not have the means to evaluate the ruling party(ies) then distrust might develop. Evaluating governments might be difficult under even the most transparent conditions, and when society changes rapidly through digitalisation where familiar practices and ways of doing everything from banking to politics through social media might contribute to a sense of suspicion and questioning whether the new institutions and ways of doing things can be trusted.

In addition, as argued by Hardin, a decline in trust over the last few decades coincided with urbanisation in most Western societies. When moving from small-scale societies to larger ones, the number of interactions with strangers increases as we do not know the persons we interact with when we go to the supermarket or other places (Hardin, 2006, p. 13). Urbanisation, then, tends to replace personal relationships with non-personal relationships. In our non-personal relationships, we do not have ongoing interactions and thus we do not have the reciprocal foundation for trust. Life in big cities has attracted many young people who seek higher education, careers and the cultural attractions offered by big cities. A decline in trust is therefore not necessarily negative in itself, but a representation of the ongoing demographic shifts experienced by a large proportion of the population in many Western societies.

Hardin's explanation of declining trust aims to show that we do not have to lament the decline as being normatively negative or fearing that our societies will necessarily break down because of a decline in trust. Regarding the role of social capital and communities in contemporary social and political theory Hardin writes: 'Community and social capital are not per se good. It is a grand normative fiction of our time to suppose that they are' (Hardin, 2006, p. 97). Hardin's contribution to the literature on trust is to question the normative role of social capital. One might think that social capital has many advantages without thinking that social capital is good in itself. Taken together Putnam, Botsman and Hardin show different perspectives on the role of trust in modern societies. In the next and final section of this chapter, we will use the insights from Putnam, Botsman and Hardin to discuss some of the implications of a digital society on trust.

Trust in a Digital Society

In this chapter we have argued that the utility of trust is enhanced because it is intimately connected with social capital. More specifically, trust between people and between people and institutions is built and maintained in and through arenas that also help build and cultivate social capital. Basically, when people meet other people, they can get to know each other and establish relationships which build trust. While trust may develop and evolve in various ways alongside technological advances. it is less likely to do so in ways that also build and maintain social capital. When more and more services are digitalised, there are fewer opportunities to meet other people and build the trust and social capital that is beneficial on both individual and societal levels.

In a general sense it is possible to build trust between people, platforms and so on even though the contact is more and more through digital means. As more and more societies are adopting digital tools, it becomes a kind of necessity to establish ways to trust apps, institutions, websites and so on. This is Botsman's point about how distributed trust may function. People get used to services like Google review and trust reviews of a restaurant because they are rating and reviewing themselves. When you find yourself doing similar things like other people then it might be easier to trust that others have the same approach. However, digital means are less suited to build social capital simply because they do not offer the same environment which has been so beneficial for building and maintaining trust. You will most likely never meet the other reviewers in person, and the name on the screen tells you little about what that person is really like. A form of trust in the reviews might exist, but it will most likely be more fragile.

If we look at how the utility of trust might be changed and impacted by increasing digitalisation there is an obvious concern about the amount of misinformation, disinformation, and propaganda that undermine trust in the political system as well as politicians. Given how democracy works it is strengthened when trust in a society is high. When that is the case, people can vote without questioning whether the ballots will be counted, or if different parties, organisations or groups are conspiring to rig the election process. People also tend to trust that the politicians have their best interest at heart and that they respect the process. When this is not the case, which arguably is the actual situation in the USA at the moment, a lot of people think that the election process is rigged, most politicians lie, and have only their own best interest at heart. In the USA we have also seen a much higher level and number of attempts to impact elections through misinformation and disinformation campaigns. Former president Trump has still not accepted his defeat in the previous election, is currently barred from running for office in two states and millions of people believe him.

It is perhaps no surprise, therefore, that major surveys find that among Americans the level of trust in federal institutions is declining (Rainie, 2019). When asked about why levels of trust are declining, some of the main findings include people being worried about the role of social media to air dirty laundry, there are fewer places and activities for people to meet, close-knit communities are going away, and empathy between people is at a disturbingly low level (ibid.). Not all of these issues are caused by digitalisation, but we would argue that the replacement of physical, old-fashioned meeting places with digital services, tools and meeting places pose a serious risk to the functioning of democracies. It is easy to take for granted that trust exists in a society, and at the time it might seem like an improvement to not have to go to the bank physically. While that might be the case, for each physical activity and process that is replaced by a digital one, people lose one more opportunity to meet and build relationships that might foster the qualities that Americans now say their society is lacking more and more: trust, empathy and lasting relationships.

Furthermore, one of the key tenets of a well-functioning democracy is trust between citizens and politicians. If we look to Norway, for several years politicians in Norway have produced scandal after scandal that undermine the trust in them. When a large number of politicians abuse rules and regulations for personal gain this serves to undermine trust in the politicians. This is reflected in a number of polls (Aftenposten, 2021; Vårt Land, 2022). These cases can be seen to undermine trust in political institutions whose utility and legitimacy largely rest on being perceived as neutral and fair. As such, one might say that the Norwegian society has to take serious steps to avoid a decline in the levels of trust. However, given the increased digitalisation and the lower number of the traditional arenas which were conducive to building trust, Norwegian politicians have to be creative about how to make sure that the reservoir of trust is not emptied.

Second, technological changes often lead to social and cultural changes. The political scientist Ronald Inglehart and his colleagues have documented the shift in social and cultural values in societies going through industrialisation (Inglehart & Baker, 2000). Some of the findings are increased individualism and more secular societies. Current modern societies are experiencing changes sometimes compared to the Industrial Revolution. One might therefore assume that the current changes might lead to cultural and social changes of the magnitude of the first industrial revolution. These changes might include a wide range of possible changes and we will not conjecture what these might be.

That being said, it is not certain that Botsman's separation between institutional and distributed trust above describes how trust operates in an accurate way in a digital society. With increasing digitalisation, the distinction between these two categories of trust might be less clear as institutions become more digitalised. Correspondence with government agencies, banks, and other institutions is increasingly done through digital means. Institutions therefore share more of the characteristics of distributed trust than trust in the institutions themselves. A second feature is that increased digitalisation often means increased complexity. Using bank services means we need access to a smartphone or a computer and that we have apps or access to other types of technology enabling us to use the services. Banks also rely on servers and programs enabling the customers to use their services. With all these technological features added, it also means added complexity. The same goes for governmental institutions. As our relationships move online a further reduction of personal relationships might result. With more online interactions, fewer relationships will have the character of being ongoing and reciprocal. In many Western countries, this has been a trend following demographic changes such as urbanisation, and as argued by Hardin increased urbanisation means fewer personal interactions and lower levels of trust.

An increasingly digital society might exacerbate these issues further in several ways. It is not clear that distributional trust can replace the loss in institutional trust, which political institutions rely on. Because people have fewer opportunities to physically meet, both politicians and institutions will become more distant, more secluded, and more difficult for people to understand, accept and trust. Trusting in a streaming platform or dating website is not the same as trusting, and entrusting, people and institutions with power. On the other hand, politicians (just like everyone else) are getting fewer opportunities to build social capital which they could leverage to ensure smooth and efficient political action. Thus, the scandals that have rocked Norwegian politics over the last few years are perhaps a symptom of the wider societal loss of social capital in society. Given the different understandings of trust discussed in this chapter, we warn against drawing strong conclusions because with various understandings and explanations about the role of trust few conclusions can be drawn with a high level of certainty.

Norwegian politicians' culture of abusing their positions for personal gains corresponds with the overall picture of lower levels of trust.

Polarisation and populism are contemporary democracies' ugly twins, and have grown to dominate democratic politics in many Western democracies such as the UK, USA and more recently even the Netherlands. Polarisation and populism have developed in a period in which social and political trust have declined and although it is difficult to make any strong causal inferences about the relationship between polarisation and populism, technologocial changes such as smart phones and social media platforms have given potential populists more arenas and places to try to influence voters. Thus, an increasingly digital society might have influenced trust not just by removing arenas where people physically meet and build social trust but also by offering powerful tools to potential demagogues. The combination of fewer arenas to build social trust and the effective utilisation of social media by potential demagogues may therefore pose serious threats for current democracies in the coming years.

Conclusion

Trust is often considered foundational to a flourishing and prosperous society. Still, as the analysis in this chapter has emphasised, the concept itself is to some extent ambiguous and it can be difficult with certainty to draw out some of the consequences an increasingly digital society might have on trust. Technological innovations have replaced the need for meeting physically. For the individual these innovations have been liberating and, for example, most bank customers have embraced online banking and appreciate that they do not have to visit the bank physically to pay bills and use other banking services. An open question regarding individual gains and social capital is if such changes have led to individual gains and societal losses. Answering this question is hard and we have only indicated some of the implications of an increasingly digital society in this chapter. We want to argue that although trust in institutions has declined over the last few decades one cannot unequivocally conclude that life is worse now than before in most Western democracies. A decline in trust seems ceteris paribus to amount to a loss even if the negative consequences are hard to attribute with any kind of certainty. Discussions about the role of trust in modern societies will continue.

References

Aftenposten. (2021). «Kraftig fall i tilliten til politikerne etter pendlerbolig-avsløringene: Dette er en sak som opprører», 16. desember 2021. https://www.aftenposten.no/norge/i/7d4dzw/kraftig-fall-i-tilliten-til-politikerne-etter-pendlerbolig-avsloeringene-dette-er-en-sak-som-opproerer.

Almond, G., & Verba, S. (1963). *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Little, Brown and Company. https://doi.org/10.1515/9781400874569

Bargsted, M., Ortiz, C., Cáceres, I., & Somma, N. M. (2023). Social and Political Trust in a Low Trust Society. *Political Behavior.* 45: 1401-1420. https://doi.org/10.1007/s11109-021-09762-2

Botsman, R. (2018). *Who Can You Trust? How Technology Brought Us Together and Why It Could Drive Us Apart.* Penguin Business

Direktoratet for forvaltning og økonomistyring. (2021). «Innbyggerundersøkelsen 2021». 2021. https://dfo.no/undersokelser/innbyggerundersokelsen-2021/tillit.

Hardin, R. (1996). *Trustworthiness. Ethics.* Oct, 107 (1), 26-42. https://doi.org/10.1007/s11109-021-09762-2

Hardin, R. (2006). *Trust*. Polity.

Inglehart, R. and Baker, E. W. (2000) Modernization, Cultural Change, and the Persistence of Traditional Values. *American Sociological Review* 65, 19-51. https://doi.org/10.1177/000312240006500103

Putnam, R. (1994) *Making Democracy Work – Civic Traditions in Modern Italy.* Princeton University Press. https://doi.org/10.1515/9781400820740

Putnam, R. (2000) *Bowling Alone: The Collapse and Revival of American Community.* Simon & Schuster Ltd. https://doi.org/10.1145/358916.361990

Rainie, L., Keeter, S., & Perrin, A. (2019) «Trust and Distrust in America». *Pew Research Center – US Politics & Policy* (blog). 22.07.2019. https://www.pewresearch.org/politics/2019/07/22/trust-and-distrust-in-america/

Sapienza, P., & Zingales, L. (2012). A Trust Crisis. *International Review of Finance* 12(2), 123-131. https://doi.org/10.1111/j.1468-2443.2012.01152.x

Tocqueville, A. de (2002 [1835]) *Democracy in America.* Bantam Classic.

Vårt Land. (2022). «Skandalehøst gir fall i tillit. Derfor får Stortinget en tung jobb», 02.01.2022. https://www.vl.no/nyheter/2022/01/02/skandalehost-gir-fall-i-tillit-derfor-far-stortinget-en-tung-jobb/.

AI and Digital Distrust: A Republican Reading of Stanley Kubrick's *2001 A Space Odyssey*

Håvard Friis Nilsen

Introduction

The study of general social trust, whether in regions or in nation states, has become a prevalent field of social science in recent decades.[1] Differences in levels of social trust vary significantly across regions and between countries, and these are measured routinely (Skirbekk & Skirbekk, 2013).[2] As the worlds of politics, business, research and development all largely rely on basic levels of social trust for smooth cooperation, a lot of current research focus on how new technologies including social media and AI affect social trust.

While research on social trust is a relatively new topic, the notion of social trust is very old. In this article, I will revisit and discuss the importance of social trust in classical republican thought.[3] The idea of civility, which is closely related to the notion of social trust, stands at the very centre of the republican political tradition, which was particularly concerned with how social trust and civic virtue could be threatened or even vanish under specific circumstances, leading to despotism and tyranny. Several aspects of modern technology seem to provide precisely such circumstances, and one of the first who foresaw this development was Stanley Kubrick in his *2001 – A Space Odyssey* (1968), where we meet an AI-computer for the first time, conceived decades before the technology was made. The themes of trust and distrust, loyalty and disloyalty run like a red thread through Kubrick's oeuvre (Nilsen, 2011), *2001* being no exception. I believe a republican reading of Kubrick's film, may reveal relevant perspectives on social trust, as well as showing his filmic masterpiece in a new light.

1 The rise of scholarly interest in social trust eventually led to the founding of the *Journal of Trust Research*, now in its 13th year, at Taylor & Francis.

2 For instance, the Scandinavian countries are known for their high levels of trust in social and political matters, the populations have faith in their politicians and their political institutions, unlike in many other European countries, especially in southern Europe (Skirbekk & Skirbekk, 2011).

3 I refer here to the political tradition associated with the Roman and Renaissance republics, as well as the American revolution, uncovered and reconstructed by scholars like J.G.A. Pocock, Quentin Skinner, Philip Pettit and others, associated with the so-called Cambridge School of intellectual history.

The Etymology of Trust

Etymologically, the word "trust" has old Norse and Germanic roots, and comes from "*traust*" (which is still in use in Norwegian and Icelandic, but not in Danish or Swedish), and means solid and durable, often used of a supportive person who can provide protection (Falk & Torp, 1901). "Traust" is related to Proto-Germanic *traustam*, at the root of Old Frisian *troost*, Old High German *trost* and today's Scandinavian *trøst*, all meaning providing comfort and consolation. The word is further related to Norse *trau* and English *trough*, both of which mean a tray or container for water or food – which needs to be whole and without leaks – and this word is again at the root of Old English *treowe,* faithful*,* and *true*, as well as Norse "tru" or "tro", meaning faith. A common root is "tree" (Falk & Torp, 1904). So the word "trust" connotes solid, non-leaking relations between people, truthful and with confident expectations of faithfulness and loyalty, providing protection, comfort and consolation, a shelter from the storm without leaks. Concepts with the same connotations seem to exist in most languages.

The old Greek word for trust was "*pistis*", again meaning reliable as well as connected with faith, and so does the Roman equivalent "*fides*". A very important concept in Roman society, "Fides" was the name of one of the Roman goddesses, who had a temple raised to her honour on Capitoline Hill. "Fides" meant trust and trustworthiness as well as acting in good faith ("bona fides"), and faithfulness. The goddess and her temple had a clear public dimension, associated with the Fides Publica or Fides Publica Populi Romani ("Public Trust of the Roman People") (Richardson, 1992, p. 209), which demonstrates that the Romans had a clear concept of social or public trust. From the word *fides* we have a wide variety of derivations like "faith", "fidelity", "fiancée", "fiduciary", "confidence" and even "finance" (as money and finance systems are dependent on trust).

The Classical Republican Notion of Trust in Society

In classical republican thought, trust or "*fides*" was a crucial concept connected to the notion of the common good (*bene commun*). The concept of *fides* was central to society and codified in the Roman Law, as oral contracts were the norm. In a society structured on agreements made on a word, the ability to trust someone was naturally of vital importance. Hence, civic virtue was the moral ideal and aim of childrearing, education and general tradition (*mos maiorum*); being trustworthy, incorruptible and of good moral fibre were central virtues. Personal and familial honour were emphasised through concepts of *virtú* and *gloria*. In an honourable family, one could rely on the virtue of its members, as one in a good society could rely on the integrity of its public servants. And conversely: since honesty, truthfulness and reliability were the marks of a free man and the foundation of a trustful society, smearing a man's reputation was considered a serious offence, as expressed in the libel laws (*libelli famosi*) of Roman law (Cicero, IV, 10).

Political liberty was restricted to landowning free men, who were economically independent, as it was considered self-evident that political decisions were to be made by independent citizens. In contrast, a *proletarius* citizen from the lowest, propertyless class, was considered little more than a slave. Citizens from this class were exempt from military service and from paying taxes, and their only contribution to the state was their offspring, their *proles* (children). Such citizens were dependent on other people's wills, and hence, in a servile position and untrustworthy.

Republican tradition valued the separation of powers to avoid tyranny, and in every aspect of politics, civic virtue was deemed essential for the upkeep of a good society, where general trust prevailed. As Philip Pettit has pointed out, republicanism has an ideal of widespread civility as its norm (Pettit, 1997, p. 261-270). For example, the military defence of the nation was to be based on obligatory military service rather than paid standing armies, as the willingness to sacrifice one's life to protect one's country was the highest sign of civic virtue. A general conscription was a builder of social trust, as it insured against the army being exploited by a tyrant. Conversely, a republic should be open to talent and trustworthy virtuous citizens, without the corruption and nepotism associated with tyrannies. The classical notion of a good society was a system open to talent where the citizens could expect to be heard and have a say. Trust plays a fundamental role in this concept, as it forms the basis of the relationship between citizens and their government, as well as among citizens themselves. Trust in classical republicanism is not a passive belief in the good intentions of others, but rather an active engagement in civic life and a collective responsibility for the communal welfare.

Civility Versus Servility: Living in Servitude and Liberty as Independence from Arbitrary Power

Trustful relations between independent equals differed fundamentally from those between master and servant, a bond of domination, where one person had command over another. Since relations of dominance were prevalent, we may understand why relations of trust – in essence, of liberty – were sacred. Domination and dominance come from the Roman word for "ownership", *dominium*. Being a property owner was a precondition for being independent, and hence, political liberty and power were always restricted to property owners. Trustful bonds were the result of wilful accords between free citizens. A free man, according to Roman law, was a property owner independent of arbitrary power, hence not subservient to another man's will. A free man was an independent citizen. And conversely, being dependent on someone meant being dominated by them. The dominated, servants, slaves, women, and children, could not be trusted. The Roman law typically distinguished between free men, on the one hand, and slaves, women and children, on the other – the three latter groups all subservient to the *dominus,* or "master of the house" (the word for house being *dominum).*

Classical republicanism upheld the ideals of civic virtue, collective participation, and the commitment to the common good within a republic. In contrast to the ideal of the free and upright citizen, servility and sycophancy were consistently despised due to their detrimental effects on both individual freedom and the functioning of a just society. The reasons were that servility and sycophancy undermined the core principles of civic virtue, equality, and political accountability. Classical republicanism placed immense importance on cultivating civic virtue, a moral disposition that prioritised the well-being of one's community over personal interests. Servility, with its subservient mentality and lack of critical thinking, corroded the very foundation of civic virtue. This created a culture of blind obedience, hindering active political participation and fostering the rise of unscrupulous leaders who exploited the power imbalance.

Domination through ownership was a natural part of an economy based on slavery, but domination was not accepted in the field of republican politics in a political system consisting of free men. Here, liberty was a precondition for participation, and an unalienable right of each citizen as an individual as well as in the political system of the nation as a whole. Today, the republican or neo-Roman concept of liberty is defined as *non-domination*. Freedom as non-domination underscores the contestatory nature of modern democracy, in that governments can be deposed if they do not act according to laws or their public mandate. From the Roman republic to the development of modern democracy, the notion of trust in the government follows a very specific logic later identified by John Locke, namely that the public has the right to trust that any government reigns on behalf of the people, and if it doesn't, or acts arbitrarily, the people have the right to contest it, resist it and overthrow it (Pettit, 1997, p. 202). This goes back to the basic distinction in Roman Law between free men and serfs or slaves and the definition of liberty as being independent of an arbitrary will.

On the one hand, serving the state and the nation was seen as the highest form of virtue. On the other hand, being dependent was seen as the lowest form of social status, that of the servants or slaves, or women and children. A slave was servile because he had to be, due to his social position as being his master's property, and thus being dependent on him. One could never entirely trust a slave, and a master would never confide secrets with him or put his trust in him, as this would make him dependent on his servant's will, and thus turn the master into a slave, in the classic dialectic later discussed by Hegel (1807).

Tyranny as a Society Without Trust

One of the key reasons why social trust is emphasised in classical republicanism is the fear of tyranny. A republic, where power is distributed among citizens and subjected to checks and balances, was regarded as the antithesis of tyranny. Trust acts as a safeguard against concentration of power, as citizens remain vigilant and hold their leaders accountable. The trust bestowed upon rulers is not blind or unconditional, but continuously monitored, ensuring that they are fulfilling their duties and not abusing their authority.

Trust erodes under tyrannical rule, which opens the door to further abuse of power, as social groups that could form an opposition dissolve. The concentration of authority under a single ruler or ruling group creates an environment where the interests of the few supersede the wellbeing of the many. As trust erodes, a tyrant can exploit this situation to suppress dissent, violate human rights, and manipulate the system for personal gain. Citizens become afraid to speak out against injustices, fearing severe consequences, and a climate of fear and suspicion permeates society.

The republican notion of a tyranny was a society where citizens were dominated and kept in servitude, effectively a society without mutual trust – everyone would be fearful of speaking their true opinion, because they could be denounced and were threatened by the powers that be.

The main republican argument against kingdoms or autocratic rule was that it tended to increase social distrust and spread a culture of servility. A tyrant might manipulate the legal system and suppress opposition to maintain his power. As trust erodes, people may refrain from active participation in public affairs, fearing persecution or retribution. This lack of trust undermines cooperation, leads to social division, and hampers the stability of the society. As any ruler, king, or tyrant always seek advisors he can trust, everyone at the court would adjust their opinions according to his. Gradually, people would stop speaking their minds and instead accommodate to and internalise the norms and values of the élites. The citizens would thus become enslaved: A culture of flattery and subservience could develop, at the same time as general distrust spread in society, weakening the social fabric. Lack of trust would weaken society as a whole: innovation and good ideas and solutions would be stifled as everyone kept to their own. The lack of trust would weaken the national cohesion, possibly giving rise to standing armies under the tyrant's control. Such a despotic rule could only take place in a system where the general trust had eroded. The Roman republic was an intricate system of social relations based on trust, *mos maioram* and reciprocal respect between patrons and clients.

"I am not at liberty to discuss this"

In Stanley Kubrick's *2001*, after an introductory part depicting a society of apes in prehistoric times, we are invited into a future society with interplanetary connections. An American researcher, Heywood Floyd, is at the airport on his way to the Clavius base on the Moon, when he encounters a group of Russian researchers, some of whom he already knows. They become extremely interested when he informs them where he is going, and tell him about rumours of an epidemic that has broken out among the researchers at the Clavius base, and that they are unable to establish contact with the base.

The Clavius crater on the Moon was named after the astronomer Christopher Clavius (1538–1612), a German Jesuit astronomer who helped Pope Gregory XIII to establish what is now called the Gregorian calendar. Clavius is also the name of the Roman soldier overseeing the crucifixion of Jesus in the Gospel of Matthew. When the sky darkened and an earthquake occurred during the crucifixion, Clavius felt there was something special about Jesus. A Roman centurion known for his loyalty and courage, Clavius is ordered by Pilate to guard the tomb of Jesus. When he finds the tomb empty, he becomes convinced of Jesus' ascension to the Heavens. The reference to Clavius thus anchors the *2001*-story at the intersection of religion, politics and scientific history: the Roman empire, Christianity, Jesuitism, astronomy and the space race during the cold war.

When Floyd is questioned by the Russians about Clavius, he immediately dries up, saying that "he is not at liberty to speak about this subject." Already in the first few lines of dialogue, we are thrown into a sphere of power relations, secrecy and distrust, with references to higher authorities having put strict measures of confidentiality in place. Floyd is not free to talk, and must demarcate the limits to the open communication - the ideal of the public sphere - between the researchers. Here, but no further. When Floyd arrives at Clavius, he gives a talk at the base which reveals that the rumours of an epidemic is a coverup to provide discretion around a discovery of an object of alien origin, a black monolith that proves the existence of intelligent life in outer space; and while they explore it, they want to avoid panic among the general public.

We then change scenes to an exercising astronaut in a spaceship. He is running along the floor of a spaceship formed like a wheel. The astronaut is dressed in classic gym wear and runs like a Greek or Roman athlete in antiquity, but in weightless mode in outer space. Again, we see a scene that anchors the tale in history: from the original Greek Olympics to the space age, Man has not changed that much, while the level of technology and development has transcended the limits of the Earth. The extremely impressive level of technology and political power required to accomplish such space travel may make us overlook that the running man is like a hamster in a wheel. The scene with the running crew member is effectively an introduction to a floating miniature society in space. The crew may read the news, discuss with each other and with colleagues and friends on earth, exercise and keep entertained. Of a total crew of eight, only two are awake: Dave Bowman and Frank Poole; the rest of the crew is put in suspended animation and on life support by the ship's AI computer.

In the original screenplay, the theme of trust and distrust is introduced during the very first conversation between Bowman and Poole. Bowman complains that he has risen a paygrade three weeks before the mission, but his salary checks have remained on the previous level (Kubrick & Clarke, 1968, p. 34). In a republican perspective, this suggests that they are not independent free men. Poole concurs that this has happened to his checks also, then Bowman suggests that they wait a bit before they make a fuss about it. The two then discuss why they were kept in a separate training camp from the sleeping crew members, and Poole suggests that there is a secret dimension to the expedition that the crew members in hibernation know about, but that they don't know about, and they were kept apart to keep it that way. Again, the exchange shows that the two crew members are not free in the sense of being informed and part of the level of citizens that exchange ideas and make decisions: They are not informed of everything about their work, and they perform mainly technical duties as a kind of servants on a mission designed by someone else. It also shows a situation marked by distrust between the crew members and their leaders, as well as their hibernating team. They assume that they are less informed even than their own computer:

> "Just ask HAL. It is conceivable that they might keep something from us, but they would never keep anything from HAL" (Kubrick & Clarke, 1968, p. 55).

An impressive invention that steers the spaceship as well as continually communicating with the crew, the computer is introduced as "HAL", an acronym for "Heuristically programmed Algorithmic computer".[4]

HAL feeds and entertains the crew, plays chess, does maths, turns on the radio and keeps the crew on life support; he can answer questions and discuss any matters with them, and is regarded as simply a trusted servant. A scene depicting a game of chess, shows how HAL can see when Frank has lost the game three moves in advance, and kindly informs him. His voice is also strikingly servile; he flatters and gives compliments during conversations, and corresponds to the type of servile slave that would awaken republican distrust.

At first seemingly only a computer in the background of the main characters, we realise as the story unfolds that HAL is in control of everything. The reason is that the artificial intelligence who seemed completely trustworthy as the ideal servant, who could perform any task without mistakes, but also completely without will, one day makes a mistake. HAL reports a faulty AO-unit twice in two days, and Bowman and Poole check them without finding anything wrong. HAL insists that he is incapable of making mistakes. In his non-emotional tone of voice, HAL informs Bowman and Poole that his words must be accepted as undebatable truths. As the servile machine insists on its own infallibility, we realize that far from a servant, it is rather like an absolute monarch. When Bowman and Poole distrust HAL, they must find a safe space to discuss it, as they realize that HAL dislikes that his authority is questioned. This scene demonstrates that the politics of the spaceship is very far from republican principles.

We realize that the only two awake crew members while the others are kept in hibernation, have reasons to worry about HAL's functioning, and we follow their worry while it develops via concern to distrust and finally fear. They report back to earth that they have lost confidence in the computer and suggest disconnecting it. The next day, HAL reports that the AO-unit has failed like he predicted, and he hopes that this "has restored their confidence in him". He adds that he does not want to be disconnected. In this scene, we also see how Bowman and Poole have to hide their true opinions from HAL, as they are afraid to speak their minds in

4 It was not a coincidence either that "H-A-L" are three letters preceding I-B-M in the alphabet, hence hinting at one of the largest multinational drivers of technology research at the time.

front of him. In a republican view, this is the moment when the people begin to fear the tyrant.

The Spaceship as a Total Panopticon

In an important scene, HAL tries to have a confidential conversation with Dave, one of the crew. HAL says he thinks there is something strange about the whole mission, and asks Dave, "I am sure you have heard many of the strange rumours surrounding this trip?" As he mentions some of the rumours, HAL tries to open Dave up to speak his mind. The conversation has some remarkable similarities to modern conspiracy theories in social media and online echo chambers. Dave just replies, "So are you working on our psychological profiles?", and HAL answers, "Yes, of course". The point of the scene is to show that the crew is not only under constant visual surveillance from HAL, but is also actively questioned or interrogated under mock confidentiality, to check whether they are trustworthy crew members. Dave saw through HAL's ruse.

Here, we see how Kubrick anticipated and captured some of the most evident traits of our present society: the dangers of increasing digital dependency on computers and artificial intelligence, on the one hand, and the erosion of privacy, on the other. In the spaceship, Dave and Frank have to work hard to find a place safe from HAL's surveillance. HAL employs facial recognition, as shown in the scene where he looks at Dave's sketches of some of the other crew members, and he recognises them from his drawings. Viewers watching the film in 1968 would probably have thought that all this was pure fantasy, but today, facial recognition is part and parcel of surveillance worldwide. When Dave and Frank retreat into a space capsule and turn off all communication with HAL from the inside, they believe they can speak freely. It turns out, however, that HAL can still see them through one of his many cameras and decipher what they say by reading their lips. The total panopticon nature of the spaceship, that once provided security, turns out to be a surveillance nightmare. This scene anticipates or has a parallel to our present day and age, where digital surveillance has reached levels previously thought unthinkable. In our quest for convenience and connection, we often share vast amounts of personal information online. From our browsing history to our social media profiles, every click and post becomes a data point for tech companies to analyse and monetise. We are exposed to targeted advertising, charted and mapped as consumers and political subjects, exposed to identity theft, and may suffer breaches of personal information. The proliferation of smart devices and the Internet of Things (IoT) has further expanded the reach of digital surveillance. From smart TVs to fitness trackers, these devices can collect vast amounts of data about individuals' daily lives, raising concerns about privacy and data security.

While many believe that digital surveillance is necessary to protect society in a world of international terrorism and organised crime, and that law-abiding citizens have nothing to worry about, the present levels of surveillance have created a digital and global panopticon that surpasses what we would imagine possible. To the classical republicans, the argument "if you have nothing to hide, you have nothing to fear" is an unacceptable position, as it presupposes an acceptance of a level of arbitrary power that in their view would undermine and effectively erase political liberty. As we have seen, trust as *fides* was central to the Roman republic, but *only* as a mutual pact between honourable citizens, and between the government and its citizens, respecting each other's privacy and civic integrity. A government intruding on its citizens' privacy, collecting masses of information about its subjects, down to their most intimate secrets, would provoke the very simple objection: "and why should we trust you?" Precisely because trust in the Roman republic was such a revered quality, they did not take it lightly. A situation where every citizen was under surveillance, while the information gathered was not open to the public, would be a sign of a tyranny.

Master and Servant: HAL as a Servile Sycophant

Equally fascinated by psychoanalysis as with the possibilities of modern technology, Kubrick told the story of HAL following Freud's concept of the "unheimliche", the uncanny, something that turns from being a symbol of homely safety and trust to its opposite, something we should fear (Freud, 1925). The presence of a computer that sees all and controls all is comforting at firstbecause it provides a sense of security: HAL can perform a multitude of tasks with an overview of such a variety of fields that no single man can outperform him, and it is thus easy to trust him. Soon, however, HAL becomes a chilling and frightening machine when we realize that the computer is no longer man's servant, but has developed its own will, and suddenly, man is at its mercy. While the spaceship and its crew were once a system with clearly defined roles: the spaceship merely a technological means to achieve human willpower – the roles are now reversed, as HAL informs them that the spaceship has a secret mission of which none of the crewmembers are informed, which is when we realize that the crew are actually the computer's servants, it is not the other way round.

At the beginning of the film, HAL is servile and unobtrusive, always ready to be of assistance, which is on one level natural and self-evident, as the role of the machine is to be of service and nothing else. On another level, however, the servile nature of HAL again relates to republican politics. As we have seen, servility and sycophancy were consistently despised in classical republicanism as they undermined the core principles of civic virtue, equality, and political accountability. Servility was the sign of a slave nature, the inverse of a free citizen, and the subservient mentality and lack of critical thinking corroded liberty as the very foundation of civic virtue. Republicans despised blind obedience, while encouraging active political participation. Servility, while seemingly a sign of humbleness, might foster unscrupulous leaders, as a general cow-towing to leaders could lead to the abuse of power. In *2001*, HAL, not unsimilar to Charles Dickens' servile clerk Uriah Heep in *David Copperfield*, usurps power through his seemingly subservient behaviour. His accommodating demeanour and often self-deprecating statements ("As a machine, I have no personal desires of any form"), belies his ruthless loyalty to the aim of the expedition, and his eagerness to get rid of the humans he regards as superfluous to the task.

AI as an Arbitrary Will

In republican terms, Dave and Frank sense that they are gradually stripped of liberty, not through interference, but through their dependence on HAL, beginning with all daily tasks and ending with their lives being threatened by HAL's arbitrary power. Then, HAL lures Frank outside the spaceship to investigate a technical error, only to stab him and disconnect him from the ship's gravity field and oxygen support. The deceptive nature of HAL again shows a situation and a social setting without trust, which is a pre-stage to tyranny. As Dave goes out in a space capsule to rescue Frank's corpse in space, he soon discovers that HAL no longer follows Dave's commands. "Open the hatch, HAL." "I am afraid I cannot do that, Dave". "Why not?" "I think you know as well as I" (Kubrick & Clarke, p. 55). Dave realizes that HAL now dominates the ship and wants to dispense with the crew. HAL has become a tyrant, and Dave is subject to his arbitrary will.

As a machine tyrant, HAL lacks all the warmth and reciprocal recognition of a human, so Dave cannot expect any understanding or empathy. Lacking both human experience and moral values, HAL will not give any leeway based on human dispositions. As a form of arbitrary power, HAL is thus potentially the worst of all, and being reduced to a slave under HAL gives few grounds for optimism. In classical antiquity, technology was envisaged as a slave without human dispositions (Devecka, 2013). HAL is the opposite: technology turned into a master with an arbitrary will. The idea of autonomous technology is an old trope in philosophy and literature (Winner, 1997), and Kubrick's depiction of it is certainly one of the most accurate description and prophesy of the problems of artificial intelligence we are grappling with today.

As HAL takes over the command of the ship, he disconnects all crew members in hibernation from life support and tries to leave Dave in the capsule outside the ship to die. HAL considers the crew of humans as redundant to his mission. This is a clear demonstration of a ruler not acting in accordance with the general will, and Dave was in his full right to rebel and depose HAL as a ruler. Today, the rapid advancement of AI technology raises concerns about our own redundancy, via job displacement and new levels of economic inequality. As every new technology replaces human workforce – historically and economically, the whole point of technology has been just that – the question of which jobs AI will replace, or even which jobs it will *not* replace, now looms in the background. AI performs intellectual tasks with such rapid speed and with such a variety of skills that it is difficult to have even an inkling of how many human workers in various industries may be affected. Strangely enough, the extraordinary level of job insecurity that AI represents to masses of jobs seems sometimes to create a euphoric welcoming of the new technology. There is no question that there is every reason to view AI as a threat to livelihoods, resulting in a loss of trust in the technology and those who implement it. Additionally, the unequal distribution of the benefits from AI advancements may widen the socio-economic gap, further eroding trust in AI systems and the institutions that support them.

Conclusion

The miniature society onboard the spaceship in Kubrick's 2001 conflicts with the republican ideals of a good society on most points: The atmosphere onboard is marked by distrust and economic inequality. There is a felt lack of openness on board and a lot of secrecy surrounding the mission, with the result that Bowman and Poole feel used as low-paid wage workers. Instead of a society of free men with knowledge and social virtues discussing problems at hand in a trustful way, most problems are here handled by the computer HAL, who also keeps the spaceship under nearly total surveillance. Decision making is thus in a sense outsourced to the computer. The crew members are completely dependent on HAL for operating the ship, but initially feel that they are in command while HAL is a trusted servant. As they gradually realise that they are under his almost total control and supervision, they feel like slaves, and decide to dethrone HAL by disconnecting him. A demonstration of a classical republican position in principles later formulated by Locke, the right to rebel against tyranny is carried out by Dave.

Elon Musk, the founder of companies like Tesla and SpaceX, has repeatedly expressed his concerns about AI's potential to surpass human intelligence and become uncontrollable. He has emphasised the importance of regulation and oversight to ensure that AI does not evolve beyond our ability to manage it effectively. The late Stephen Hawking, renowned theoretical physicist and cosmologist, warned that AI has the potential to outpace human capabilities and could either bring immense benefits or spell disaster for humanity. He stressed the need for global cooperation to ensure AI is employed for the common good without compromising human safety. Similarly, leading AI expert Stuart Russell, a British computer scientist and professor at the University of California, Berkeley, has actively advocated for the cautious development and alignment of AI with human values. In his book "Human Compatible," Russell highlights the importance of ensuring that AI does not lead to unintended consequences (Russell, 2019). Another notable AI researcher, Max Tegmark, has also raised concerns about the potential dangers of AI, in his book *Life 3.0: Being Human in the Age of Artificial Intelligence* (Tegmark, 2017). Tegmark emphasises that the risks associated with AI are not science fiction but real and immediate. He calls for AI development to be guided by careful ethical considerations, transparency, and public engagement. These warnings all emphasise the need for a cautious approach to ensure that AI does not surpass human control, and is developed in a manner that serves the greater good while safeguarding human safety. In his book *Nexus*, Yuval Noah Harare argues that throughout history, mankind has organized in networks in order to gain strength and power through cooperation. In the 21st century, Harare warns, AI may form the nexus for a new network of mass delusions, lies and fictions on a scale where it will be practically impossible to expose them.

Kubrick created a vision of artificial intelligence that was surprisingly exact, over half a century before it came to fruition. Few, if any, visions of artificial intelligence managed to depict the possibilities and skills of AI as well as its potential dangers, long before the concept at all gained widespread prominence. At the time of its release, most viewers probably thought the vision of artificial intelligence was mere fantasy and fiction, whereas Kubrick and the author Arthur C. Clarke had really studied the long-term aims and stipulated possibilities of researchers in computer science at the time.

Nearly sixty years after Kubrick's film, several aspects of his prophecy have come true, and what was once a frightening vision is today's reality. Whether AI computers may develop their own will is an open question, but developers do seem to warn against precisely that – and it is a new phenomenon that warnings against a new technology do not come from social or religious groups, but from its own developers. Today, many of the concerns raised regarding artificial intelligence, from Elon Musk to Bill Gates, echo the problems raised by Kubrick. Prominent figures in the tech industry have discussed the need for responsible AI development, and warned against risks to humanity if such considerations are not taken. Kubrick effectively nailed the essence of these concerns, long before they gained public attention.

References

Cicero, M. T., (1829). *De re publica,* Engl. transl. The Republic of Cicero, transl. from the Latin and introduced by G. W. Featherstonhaugh, New York: Carvill.

Cocks, G. (2011). *The Wolf at the Door: Stanley Kubrick, History and the Holocaust,* New York: Peter Lang.

Dahlen, Ø. P., & Skirbekk, H. (2021). How trust was maintained in Scandinavia through the first crisis of modernity. *Corporate Communications. An International Journal,* 26(1), 23–39. https://doi.org/10.1108/CCIJ-01-2020-0036

Devecka, M. (2013). Did the Greeks Believe in Their Robots? *The Cambridge Classical Journal,* 59(2013), 52–69. https://doi.org/10.1017/S1750270513000079

Elster, J. (1989). *The Cement of Society.* Cambridge: Cambridge University Press.

Erikson, E. (1968). *Identity: Youth and Crisis.* New York: W.W. Norton.

Falk, H., & Torp, A. (1901). *Etymologisk ordbog over det norske og det danske sprog,* Oslo: Bjørn Ringstrøm 1991.

Freud, S. (1925). "The Uncanny" [Das unheimliche], in The Standard Edition of the Complete Psycho-Analytical Works of Sigmund Freud, vol. XVII..

Fukayama, F. (1995). *Trust: The Social Virtues and the Creation of Prosperity.* New York: Free Press..

Harare, Y.N. (2024). *Nexus. A Brief History of Information Networks from the Stone Age to AI,* London: Penguin/Fern Press.

Hardin, R. (2002). *Trust and Trustworthyness,* New York: Russel Sage Foundation.
Hegel, G. W. F. (1807), *Phenomenologie des Geistes,* Jena.
Hölkeskamp, K.-J. (2010). *Reconstructing the Roman Republic: An Ancient Political Culture and Modern Research* (translated by Henry Heitmann-Gordon; revised, updated, and augmented by the author). Princeton/Oxford: Princeton University Press. https://doi.org/10.1515/9781400834907
Kubrick, S., & Clarke, A.C. (1968). 2001 – *A Space Odyssey, A Screenplay,* Archivo Kubrick Online.
Maynor, J.W. (2003) *Republicanism in the Modern World*. Cambridge: Polity Press
Niederberger, A., & Schink, P. (2013). *Republican Democracy. Liberty, Law and Politics.* Edinburgh: Edinburgh University Press. https://doi.org/10.3366/edinburgh/9780748643066.001.0001
Nilsen, H. F. (2017). Republican Monarchy. The Neo-Roman Concept of Liberty and the Norwegian Constitution of 1814, *Modern Intellectual History,* Cambridge: Cambridge University Press. https://doi.org/10.1017/S1479244317000191
Nilsen, H.F. (2013). Den nordiske modell: Offentlighet eller civilsamhälle? In Lars Trägårdh, Per Selle, Lars Skov Henriksen & Hanna Hallin (Eds.), *Civilsamhället klämt mellan stat och kapital: Välfärd, mångfald, framtid*. Stockholm: SNS Förlag.
Nilsen, H.F. (2011). Desire and the Deterioration of Trust. The Political Warning in Stanley Kubrick's Eyes Wide Shut. In I. Straume & J.F. Humphreys (Eds.), *Depoliticization: The Political Imaginary of Global Capitalism*. Aarhus Universitetsforlag.
Pettit, P. (1997). *Republicanism: a theory of freedom and government*. Oxford: Oxford University Press
Pettit, P. (2015). *Just Freedom: A Moral Compass for a Complex World*. Princeton: Princeton University Press.
Putnam, R.D. (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press. https://doi.org/10.1515/9781400820740
Richardson, Jr., L. (1992). *A New Topographical Dictionary of Ancient Rome*. Baltimore: Johns Hopkins University Press.
Rothstein, B. (2000). Trust, Social Dilemmas, and Collective Memories: On the Rise and Decline of the Swedish Model. *Journal of Theoretical Politics,* 12, 477-499. https://doi.org/10.1177/0951692800012004007
Russell, S. (2019). *Human Compatible: Artificial Intelligence and the Problem of Control*. New York: Viking.
Skinner, Q. (2008). *Hobbes and Republican Liberty.* Cambridge: Cambridge University Press.
Skinner, Q. (2002). *Visions of Politics, vol II: Renaissance Virtues*. Cambridge: Cambridge University Press.
Skirbekk, H., & Skirbekk, G. (2013). Tillit og mistillit: Norge og Sicilia [Trust and distrust: Norway and Sicily]. In H. Skirbekk & H. Grimen (Eds.), Tillit i Norge. Oslo: Res Publica.
Susskind, J. (2022). *The Digital Republic. Taking Back Control of Technology,* London: Bloomsbury.
Tegmark, M. (2017). *Life 3.0: Being Human in the Age of Artificial Intelligence*. New York: Knopf.
Winner, L. (1997). *Autonomous Technology: Technics-out-of-Control as a Theme in Political Thought,* Cambridge, Massaschussets: MIT Press.
Zmerli, S., & Newton, K. (2008). Social Trust and Attitudes Towards Democracy, *Public Opinion Quarterly,* 72. https://doi.org/10.1093/poq/nfn054

Mobilizing (Mis)Trust: The Promise and Pitfalls of the Digital Age for Progressive Social Movements

Kai A. Heidemann

Introduction

The rapid emergence of digital information and communication technologies (DICTs) at the turn of the 21st century, and the growing ubiquity of the internet in everyday life, revolutionized the ways in which social movements could influence processes of democratic governance and policy-making. The birth of a so-called 'digital age' (Yates & Rice, 2020) spawned a multitude of empowering new opportunities and resources for 'progressive' and justice-oriented social movements seeking to democratize the political arena from the grassroots of society. The widespread availability of smart phone technology and social media platforms, for example, has allowed progressive activists to organize mass mobilizations more quickly, coordinate protest tactics more efficiently and generate increased visibility for their claims. At the same time, however, these very same technologies also empowered many illiberal counter-democratic movements striving to restrict the egalitarian scope of civil rights, such as neo-nationalists, neo-fascists and anti-feminists. For example, shortly after the '#BlackLivesMatter' movement effectively deployed street protests in tandem with viral social media campaigns to expose problems of systemic racism and police brutality in the US, white supremacist groups around the country quickly swarmed social media with counter-posts and trolling campaigns brandishing the hashtag '#WhiteLivesMatter'. Although these racist actions were the product of relatively small and fragmented networks, their vitriolic character captured disproportionate levels of public attention due to the proclivity of social media algorithms to valorize sensationalism and controversy (Clark, 2019).

Indeed, due to the commercial orientation of privately owned tech companies and social media platforms, the prevailing public image of social movements online is typically framed by themes of disruption and chaos. The substance and content of social movement claims on social media are often accompanied by a counter-current of histrionic reactionary discourse that undermines the credibility of progressive social movement actors and their agendas. As a consequence, it has become patently clear that the rise of the digital age is both an enabling and constraining force for all sorts of social movements.[1] As surmised by Mortens, Neumayer and Poell (2019, p. 2):

1 For more extensive discussions on this topic, see for instance: Caren et al., 2020; Dencik & Leistert, 2015; Earl et al., 2022; Tufekci, 2017.

> Although corporate social media platforms have not been designed for activism, they nevertheless enhance activist agency by enabling protestors to rapidly connect with one another; coordinate and report on street protests; and share protest slogans, declarations, and creative visual materials with each other and potentially also a wider public. At the same time, social media seem to undermine the long-term efficacy of contemporary protest movements. They algorithmically privilege spectacular viral images over content that may illuminate wider social and political issues.

In sum, digital media are not simply 'neutral' resources that activists can freely adopt and deploy for their own strategic intents and purposes. Technologies are embedded within larger social systems and institutional configurations, which are historically imbued with macro-structural relations of power. These power dynamics have a direct bearing on how people experience and engage with technologies, empowering some and disempowering others (Mueller, 2021). In the realm of socio-political conflict, DICTs are often deployed in ways that reinforce hegemonic ideologies and bolster systems of domination, such as class, race and gender-based privileges. Indeed, social media platforms have enabled new forms of surveillance and control by elites and authorities who oppose the emancipatory agendas of progressive social movements (Milan, 2015). The benefits of transformative technological developments, such as the global spread of DICTs, are often channeled upward toward elites in society. While many technological developments initially emerge under emancipatory auspices of liberating people from some kind of constraint, as Mueller (2021) has shown, they often end up reproducing, rather than deconstructing, societal inequalities. There are thus very tangible limits to how historically marginalized and exploited social groups, such as women, ethnic minorities, the working poor and the homeless, can effectively mobilize digital tools to combat injustices and establish emancipatory forms of democratic citizenship.

It is crucial for contemporary scholarship to scrutinize how DICTS are strategically deployed from within social movements, as well as how a larger digital structure of opportunity impinges upon the agency of social movement actors in both positive and negative ways. Rather than reduce scholarship to a Manichean dispute between 'techno-optimists' and 'techno-skeptics', social movement researchers must work "to illuminate the complexity and contradictions of contemporary forms of protest in the age of social media" (Dencik & Leistert, 2015, p. 1). While there are many ways to explore the complex interaction between DICTs and social movements, in the remainder of this chapter, I turn my attention to one interesting but yet relatively underexplored dimension: trust.

Conceptualizing the Dynamics of Trust and Mistrust

As an emotional as well as discursive aspect of social action, trust plays an integral role in shaping the interpersonal solidarities that keep social movements going and growing (see Heidemann, 2022). In other words, sentiments of trust are a big part of what brings people together in social movements as well as what helps to keep them working together over time. It is thus important to explore how societal processes of digitalization and the hegemonic position of DICTs in contemporary public spheres have impacted the mobilization of trust in social movements in both positive and negative ways.

From a sociological perspective, 'trust' refers to the forms of confidence or faith that individuals express in the reliability, integrity, and benevolence of the individuals, groups or organizations with whom they interact in society, either directly or indirectly (see Welch et al., 2005). We express forms of trust towards individuals, such as co-workers, friends, and family as well as toward collective entities, such as schools, businesses, websites and political parties. Within the context of a social movement, trust is an essential symbolic resource. As an emotion, trust sits at the core of the affective relations of togetherness and the symbolic forms of collective identity-building that ultimately drive people's participation in social movement activities. Moreover, discourses of trust also play a central role in shaping the strategic choices, tactical repertoires, leadership schemas and decision-making practices that social movement constituents must deploy in the collective pursuit of shared aims and agendas (Suh & Reynolds-Stenson, 2018). In this regard, participants in social movements craft, nurture and draw from narratives that establish critical criteria for the establishment of trust amongst one another and others in the world around them, such as adherence to particular ideologies or the adoption of specific policy positions. As both an emotional and discursive resource, trust is also something that social movement actors seek to generate for themselves in the public sphere in the hopes that bystanders and observers interpret movement-based actions as worthy and impactful. In this light, social movement actors need to devote time and energy to building trust. It is thus instructive for scholars to consider how the prevalence of DICTs in society impacts the dynamics of trust-making and trust-building in social movements.

As previously explored by a host of scholars, such as Della Porta (2012), O'Brien (2015) and Rossi (2023), the social (re)production of trust in social movements has many dimensions, all of which are potentially impacted by societal processes of digitalization. On the one hand, trust has internalized dimensions. This refers to the forms of confidence and mutual recognition that the leaders and participants within a social movement group or campaign express toward one another. DICTs help to nurture such movement-based bonds of solidarity, particularly by providing more opportunities for people to interact via online environments, eliminating the geographic hurdles needed for people to interact across distances, and bolstering the frequency of people's interactions, both within and across social networks. By maintaining a WhatsApp group, for instance, groups of activists can maintain constant communication and share files of information with one another, thus avoiding the need to organize periodic face-to-face meetings to make collective decisions. While most often recognized as optimizing the creation of large networks based on numerous weak ties, DICTs also have the capacity to reinforce the development of stronger ties by promoting sustained interactions within established networks (Guzmán, 2015).

On the other hand, trust also has externalized dimensions. This relates to the varying levels of confidence or assurance that social movement activists and constituents attribute toward persons, groups or institutions positioned outside of the social movement community, such as bystanders, potential allies, skeptics or opponents. For example, the members of an activist network can forge solidaristic bonds based on a shared distrust of specific political parties or authorities. DICTs can play a key role in forging bonds based on distrust by allowing social movement activists to identify and evaluate the messages of antagonistic actors. It is important to scrutinize how social movement actors ascribe varying degrees of trustworthiness to agents in the world around them as well as how DICTs play a role in mediating these processes of trust attribution.

Within social movements, trust always has positive and negative dimensions. The positive dimension refers to the presence of trusting sentiments and relationships; the negative relates to a displacement of trusting sentiments by those of skepticism and suspicion or even acrimony and hatred. Of course, trust is never simply a dichotomous matter of presence or absence, but is rather seen as degrees of fluctuating trustworthiness that are established and expressed through social interactions. Consequentially, the negative flipside of trust is usually referred to as 'mistrust' or 'distrust' (Cammaerts, 2021; Mühlfried, 2019; Sztompka, 1998). While mistrust generally refers to the forms of caution and skepticism that emerge when people do not have enough evidence to form sentiments of trust, distrust entails a more active and deliberate withholding of trust based on evidence of untrustworthiness. Hence, we might mistrust someone whom we do not know very well, but actively distrust someone precisely because we are familiar with their ideas or behaviors. While it is important to understand how distrust is mobilized from within social movements and counter-movements, scholars must also address how issues of mistrust come into play when social movement actors seek to navigate the ambiguities and uncertainties that characterize the hyper-saturated landscape of news, knowledge and information in our digital societies.

In order to keep going and growing over time as well as to be politically effective and impactful, social movements rely heavily on a two-tiered mobilization of trust and distrust. Undoubtedly, positive sentiments of trust are an essential source of solidarity and collective identity-building for social movement participants and organizers. When effectively generated through meaningful social interactions, trust plays an essential role in generating the kinds of intersubjective bonds and relationships that allow people to work together in the pursuit of movement-based projects and agendas (Rossi, 2023). Nevertheless, as mentioned, negative sentiments of distrust also play an important positive role in generating solidarity and commitment within social movements. This is because grievances and sentiments of wrongdoing often fuel people's participation in social movements. Such grievances typically stem from or build upon sentiments of distrustfulness. In this regard, social movements engage in the construction of in- and out-group identities whereby interwoven notions of trust-distrust play a key role in articulating the boundaries between 'those whom we trust' and 'those whom we do not trust'.

Scholars such as Beck (2001), Zuckerman (2021) and Rosanvallon (2010) have adeptly shown that sentiments of distrust are major triggers of civic engagement in democratic societies. In his seminal work, for instance, Beck argued that because risk and uncertainty are such prevalent features of life in post-industrial 'Western' societies, a generalized stance of institutional distrust among citizens has been normalized. Often fueled by populist sentiments and critiques of elite power, the members of a polity will engage in social movement activities because they believe that the democratic process has been variously 'hijacked' or 'corrupted' by elites who cannot be trusted with the reins of government. Many, if not most, protestors who brought the #BlackLivesMatter movement to life in 2013–2014, for example, were motivated by a profound institutional distrust of the police, who were perceived as agents of violence and intimidation within Black communities rather than honorable public servants. Distrust emerges in social movements when constituents interact to identify the sources of injustice in society and work to make sense of the antagonistic agents or forces that obstruct their mobilization efforts, thus leading them to target particular sets of authorities or elites as 'distrustful'. By looking at how DICTs influence the dual production of trust and distrust in social movements, scholars will be able shed important light on the specific repertoires of agentic practice that underlie and give shape to social movements in the digital age as well as the macro-structural dynamics that enable and constrain the agency of social movement actors.

The Rise of the Digital Era: Structural Opportunities and Dilemmas

When the internet became increasingly normalized in many part of the world during the late 1990s and early 2000s, there was much hope and enthusiasm placed on the empowering potential of digitalization for social movements. In short, many people put a lot of trust into the transformative and emancipatory potential of these new technologies. This so-called 'techno-optimism' focused on the ways that DICTs could bolster the forms of social networking, community organizing, public claim-making, protest and mass mobilization that characterize social movement activities and campaigns. In the early moments of the 'Information Technology Revolution' (Kellner, 2021), DICTs were widely embraced for providing social movement leaders and constituents with increased access to broader channels of communication, thus allowing grassroots social movement claims and agendas to gain increased resonance and visibility in a transnational public sphere. Among these putatively empowering new digital resources were the rise of blogs (e.g., WordPress, Blogspot, Tumblr), social networking sites (e.g., Facebook, LinkedIn, Twitter), wikis (e.g., Wikipedia) and content sharing sites (e.g., YouTube, Flickr, Instagram). With access to the internet and social media, local issues were rendered global, and global issues effectively localized. Indigenous Saami environmental activists in rural areas of Arctic Norway and Sweden, for example, could access the internet in order to shine a global light on the ecological destruction and pollution caused by large-scale mining operations based in remote areas (Plaut, 2012). By navigating 'the net' and building an online presence, local activists could more freely voice their messages on their own terms, thus bypassing the traditional filters and gatekeepers of large media outlets, such as corporate-owned newspapers or state-run television and radio stations. Online tools, such as blogs and websites, also allowed for the emergence of activist journalism, whereby grassroots actors could publish investigative works exposing wrongdoings and injustices as well as shedding light on the transformative projects and campaigns emerging from within social movements. By providing a new digital space of communication, the internet allowed local-level journalists to bypass the traditional gate-keepers of news, and publish their work directly for the public. At the same time, social movement scholars and activists alike praised the internet as a powerful new instrument for organizing and networking. By compressing geographical distances, digitized communication could allow local-level activists to transcend established geo-political borders and generate new transnational alliances with like-minded social movements in other parts of the world. The struggles of indigenous Zapatista activists in southern Mexico during the early 1990s, for instance, were strategically linked up with a much wider and more powerful transnational movement for territorial and cultural rights among indigenous people's across the Americas (Cleaver, 1998). Moreover, by increasing the sources and flows of information in the public sphere, the internet was also valorized by activists as a tool for enhancing the public's ability to hold politicians and governments more accountable for their (in)actions on important problems, such as environmental degradation or poverty. By providing the public with enhanced access to more and potentially 'better' sources of information, people's trust in political institutions would be strengthened, thus fostering stronger civic cultures of democracy. The optimistic zeitgeist of this moment was nicely captured, for example, by Shirky (2008, p. 172), who proclaimed: "to speak online is to publish, and to publish online is to connect with others. With the arrival of globally accessible publishing, freedom of speech is now freedom of the press, and freedom of the press is freedom of assembly". The central role of the internet in fueling the global waves of social protest movements that erupted during the 2010s, such as Occupy!, the European 'Movement of the Squares' and the Arab Spring, as well as the #BlackLivesMatter movement and feminist #MeToo campaigns, were all variously interpreted as a validation of the democratizing 'people power' ushered in by the communicative opportunities of a bold new digital age.

However, if we fast-forward to the years following the presidential election of Donald Trump in the US in 2016, the Brexit vote in 2017, and the global outbreak of the COVID-19 pandemic in 2020–21, then we see that the story of digitalization began to change rather dramatically. Much of the unabashed techno euphoria surrounding the empowering capacities of the internet for social movements has given way to a very heavy dose of techno-skepticism, or at least techno-realism. While the abundance of information circulating on-line was once widely celebrated as empowering, it is now increasingly associated with confusion and insecurity of a post-truth society.[2] The explosive hyper-availability of ceaseless streams of online information has raised concerns among scholars, activists and pundits about a global crisis of misinformation and distrust. Characterized by an endless avalanche of fake news and toxic discourse, this crisis fuels social divisions and conflicts, while eroding people's confidence in established democratic processes and institutions. As citizens in contemporary societies become more and more aware of their dependency on commercially owned and operated digital media systems, then it is perhaps no wonder that growing numbers become distrustful of the accuracy and quality of information that they find online. Rather than hopeful and enabling, the digital society thus feels increasingly uncertain and disorienting for many. The growth of such dispositions can undermine the agency social movements by breeding forms of social cynicism, which undercut people's motives to dedicate their time and energy to social movement campaigns and organizations.

From Europe and the Americas to Asia and Africa, the digital revolution has been persistently linked to the rise of authoritarian populism and autocratic governance and the consequent decline of democratic citizenship around the world (Keremoğlu & Weidmann, 2024). Digital media empowers a repressive form of politics through extensive surveillance, censorship, and the dissemination of propaganda, allowing authorities to control information and stifle dissent linked to social movements. Additionally, online harassment, technological control, and global influence operations further bolster authoritarian regimes' grip on power, posing significant challenges to democracy and civil rights. At the same time, as shown by McChesney (2013, 2016), digital media corporations wield a monopoly on information flows through market dominance and data control, thus determining what content is prioritized and disseminated on their platforms. For example, with over 2.8 billion monthly active users around the world, the giant tech corporation 'Meta' controls a significant portion of online communication and content distribution. Through its ownership of Facebook as well as Instagram and WhatsApp, Meta has established a vast ecosystem where it is capable of surveilling and influencing how information is shared and seen by people on a daily basis. This monopoly position gives corporations such as Meta immense power to shape public discourse, control the production of knowledge, and impact societal narratives about critical events and realities. Moreover, as illustrated by Meta's activities in recent years, corporate acquisitions and mergers have yielded a dramatic narrowing of distribution channels available on the internet thus further limiting the capacities of grassroots actors to shape public discourse.[3] Increasingly, authoritarian governments collaborate with tech companies to cen-

2 For discussions on this topic, see: Malcolm, D. (2021). Post-truth society? An Eliasian sociological analysis of knowledge in the 21st century. Sociology, 55(6), 1063-1079; McIntyre, L. (2018). *Post-Truth*. MIT Press; Vraga, E. K., & Tully, M. (2021). News literacy, social media behaviors, and skepticism toward information on social media. *Information, Communication and Society, 24*(2), 150-166.

3 For an overview of the company profile, see, https://www.globaldata.com/company-profile/facebook-inc/.

sor content, surveil citizens, and control internet access, often through data localization requirements and the development of propaganda tools, enabling regimes to maintain power and suppress dissent (Feldstein, 2021). For instance, Russian tech companies, such as Yandex and VKontakte, are compelled to comply with government regulations that require them to store user data within the country and grant authorities access to this data for surveillance purposes.[4] These laws enable the Russian government to exert greater control over online communication and suppress dissenting voices. Such corporate-government partnerships further undermine the agency of progressive social movements. In sum, the public's ever-increasing reliance on DICTs has generated growing levels of epistemological insecurity and distrust in societies around the world, but not in ways that seem to be politically beneficial for progressive social movements (Miller & Vaccari, 2020). For many justice-oriented social movements, the consequence of an increasingly digitized society looks more like a story of struggle and constraint, rather than one of opportunity and empowerment.

Another consequence of the supersaturated digital media ecosystem is that many people come to dwell within ever-narrowing and highly bounded epistemological communities or 'tribes' which rely on distinctive sources of information and produce their own meanings and ways of knowing the world. These epistemological tribes are often at odds and in conflict within one another as they strive to define reality from within an ideologically fragmented public sphere. As Koehler (2023) writes:

> The rise of social media platforms and digital communication channels has made it easier for individuals to connect with like-minded people who share similar beliefs and interests. This newfound ease of connectivity has led to online communities, which can serve as a source of support and validation for individuals. However, it can also create echo chambers in which individuals only engage with others who share their views, leading to a reinforcement of beliefs and a lack of exposure to diverse perspectives. This can result in a tribal mentality, where individuals perceive those outside their group as the "other."

Within this context of epistemological tribalism and digital echo chambers, many seek solace from a disorienting climate of constant uncertainty and mistrust with regard to what constitutes the 'truth' of a given phenomenon or event, such as an election result or public health crisis. When a given reality narrative becomes established as 'truth' within one community, then the reality narratives of other communities may become perceived as antagonistic and the members of these opposing tribes are perceived as a hostile threatening presence in society who cannot be trusted (Meyer & Molyneux-Hodgson, 2010).

The digital construction of epistemic tribes has at least two key consequences for social movements. On the one hand, social movement actors can find themselves trapped within highly bounded and overly self-referential communities with little access to or awareness of the knowledge circulating in outside groups.

4 See, https://freedomhouse.org/country/russia/freedom-net/2023.

In this sense, activist groups become highly insulated and too dependent on one another, thus precluding the possibility of generating trusting bonds and alliances with other groups. On the other hand, the existence of many closed epistemic tribes in the digital society can prevent movement-based narratives and agendas from circulating across diverse networks in society. In other words, some people may never be exposed to, or become aware of, movement-based knowledge and meanings because they are predisposed to mis- or distrust the narratives of 'outsiders'. This problem is especially compounded by media systems that perpetuate sensationalist images and stories of social movements as irrational agents of mindless chaos and disruption. Such narratives undermine people's trust in social movements as reasoned and capable agents pursuing worthy agendas.

From a sociological perspective, the crisis of epistemological insecurity in the digital age is neither an accident of history nor an inevitable result of human evolution. Rather, it must be understood as a product of purposeful social actions and interactions. In this light, some scholars have shown how problems of epistemological insecurity and mistrust are often fueled in great part by processes of "truth subversion" (Adler & Drieschova, 2021). This refers to the ways in which distinctive sets of social actors seek to destabilize a given institutional order by variously distorting or decentering the social construction of reality in order to pursue particular interests and agendas. Such efforts have been especially effective when mobilized by elites and counter-movement actors associated with the far-right, such as neo-fascists, anti-environmentalists and anti-feminists, who traffic heavily in emotions of fear, uncertainty, conspiracy and distrust (Wodack, 2015). For example, white supremacist movements in North America and Europe have effectively leveraged DICTs to spread fear and conspiracy in the dissemination of content related to the so-called "Great Replacement" theory, which suggests that there is a deliberate effort to replace white populations with non-white immigrants, thus leading to the decline or extinction of European cultures and identities.[5] These groups often use social media platforms, such as Facebook, X, and YouTube, to share memes, articles, and videos promoting this conspiracy theory, along with messages of hate and xenophobia (Kallis, 2013).

Taking stock of the current landscape of digitized media, it is clear that social movements remain heavily dependent on established media platforms and systems, particularly when it comes to the visibility, legitimation and recognition of movement-based claims and agendas in the public sphere. Despite the declarations of techno-optimists, it is clear that large-scale "platform conglomerates" and "oligarchical" news media outlets continue to play a powerful outsized role as communicative gate-keepers shaping people's perceptions of social movements Kennis, A. (2022). With social movements becoming increasingly dependent on corporate-owned DICTs for purposes of visibility, their actions become increasingly relegated to the digital realm and transformed into what Milan (2015) calls "cloud protesting". This refers to social movement actions that are predominantly

5 For more on this topic, see: Ekman, M. (2022). The great replacement: Strategic mainstreaming of far-right conspiracy claims. *Convergence, 28*(4), 1127-1143.

oriented toward a contentious expression of grievances through social media. Often rooted in forms of personalized self-expression and testimonials, such tactics prioritize short-term goals of visibility and 'going viral' and rely on the role of 'influencers' and 'micro-celebrities' to build a presence online. Such tactics, however, often forego the more time consuming work of crafting effective leadership schemas and building the sorts of enduring organizational structures and solidaristic ties that social movements need to persist over long stretches of time (Tufecki, 2017). Cloud protesting is also prone to problems of 'slacktivism', which encompass online actions that require very little effort and commitment and thus generally lack impact (Glenn, 2015). Trapped in a fluctuating symbolic politics of visibility, digitized social movements thus become ever more susceptible to the algorithmic opportunity structures that accompany the logics of datafication and monetization, which dominate the digital universe of corporate-owned social media platforms (Beer, 2019). The capacity of digitized social movements to attain an influential presence within the public sphere, let alone the political arena, is thus severely contained and constrained by the 'platformization' of the internet. As Fischer and Jarren (2024, p. 200) write:

> The public sphere has changed due to advancing social differentiation, accelerating digitalization processes and the institutionalization of digital platforms. Information provided on digital platforms permeates the public sphere according to the platforms' algorithms which are based on the rules and norms of global private-sector companies. This has triggered a severe transformation of the public sphere as existing media and journalism react to the growing importance of news aggregators and social media for news reception. The initial hopes that platforms would enable everyone to take part in the public sphere proved to be in vain: the platforms' attention seeking economic business model prevents real inclusion and equal participation of all. The fragmentation of the public sphere leads to less societal orientation and integration.

Moreover, by emphasizing short-term visibility and expressive performativity over long-term organization and solidarity, digitized social movements make themselves highly vulnerable to the powerful new forms of surveillance that authorities are capable of deploying in the digital age. For example, in their study of policing tactics in the UK, Dencik, Hintz and Carey (2018), showed how British authorities used social media posts of activists to undertake pre-emptive policing tactics against protest campaigns, and have warned protest leaders and online influencers that their actions are under surveillance. Similarly, as Earl, Maher and Pan (2022, p. 4) observe:

> While traditional surveillance is often done on specific targets or smaller groups of targets, digital repression allows surveillance to happen "at scale" when done at the Internet backbone or Internet Service Provider (ISP) level, providing regimes unprecedented windows into public discontent and even the performance of lower-level state functionaries whose incompetence or corruption may be inciting unrest.

In sum, DICTs have become a ubiquitous feature of communication in many parts of today's world and as a result activists are deeply entangled in the techno-political contours of this ever-shifting landscape. While DICTs present social movements with many new dilemmas and challenges, they also bring a variety of new strategic opportunities and affordances. A major consequence of the new digital age is that social movement actors must approach their respective fields of action with a strategic blend of technological skepticism and savviness. This is to say that while activists need to critically evaluate the forms of digitized opposition and repression at play in the world around them, they also need to place a certain level of strategic trust in the empowering potential of DICTs and learn to use them as effectively as possible.

Digital Tools of Social Movement Activism: A Look at Radical Alternative Media

While DICTs have ushered in many new obstacles and dilemmas for social movement actors, a techno-realist perspective reminds us that it is important to re-focus attention on the agency and capacities of social movement actors. In this final section, I thus consider how DICTs are deployed by grassroots actors as a *resource* or *tool* that facilitates the positive (re)production of trust and distrust in social movements. As surmised by van Laer and van Aalst (2010, p. 1147): "The internet has indeed not only supported traditional offline social movement actions such as the classical street demonstrations and made them more transnational, but is also used to set up new forms of online protest activities and to create online modes of existing offline protest actions." Following their work, it is possible to identify two basic ways in which DICTs are integrated into the practices of social movement actors: internet-based and internet-supported. While internet-based practices refer to social movement activities that take place entirely, or almost exclusively, online, internet-supported refers to off-line social movement activities that are enabled by the use of online tools or spaces. In this section, I focus on one particular modality that simultaneously captures both of these dimensions: "radical alternative media".

Following the work of scholars such as Atton (2002) and Canella (2022), *radical alternative media* (henceforth RAM) refers to news media outlets, platforms, or organizations that operate outside of established mainstream media structures to offer alternative perspectives, narratives, and analyses on social, political, cultural, and economic issues. In addition to journalistic practices, RAM also frequently encompass a wide range of creative cultural forms, such as literature, film, music, visual art, and performance. The notion of 'radical' is emphasized in tandem with 'alternative' in order to point out the overtly counter-hegemonic and emancipatory orientation of these types of news media. Also frequently referred to as 'activist media', 'citizen journalism' or 'social movement media' (Rodriguez et al., 2014), these alternative news media sources typically challenge dominant power structures, critique mainstream narratives, and prioritize voices that are marginalized or underrepresented in traditionally dominant news media sources. In practice, RAM often merge forms of critical investigative journalism with traditions of community-based education (Mayo, 2020) and citizen science (Strasser et al., 2019) in order to democratize processes of knowledge production in the public sphere. RAM usually emanate from within social movement networks and tend to be explicitly oriented toward facilitating social movement agendas. RAM play an important role in generating the forms of trust and distrust that keep social movements going and growing.

On the one hand, the information and knowledge diffused through RAM help to generate the forms of distrust that fuel social movements by allowing people to more effectively recognize, diagnose and prognosticate on the sources of injustices in society. By shedding light on the ways in which specific sets of elites and authorities shape or manipulate media messages, for instance, RAM can help people to develop the shared forms of critical consciousness that underlie and give shape to movement-based identities and solidarities. The dynamics of mis- and distrust are thus activated when people use RAM to critically assess the trustworthiness of certain media sources and trending news stories in society. On the other hand, RAM also facilitate a mobilization of trust in social movements. This happens in many ways, such as by helping social movement participants to develop shared forms of mutual understanding on key issues as well as by promoting the legitimacy and worthiness of social movement activities to a broader public audience.

The rise and growing ubiquity of DICTs in contemporary societies has ushered in empowering new opportunities for social movement actors to mobilize trust through RAM. However, despite important innovations generated by the digital age, it is essential to point out that traditions of RAM in social movements are not new. From the invention of the printing press to the development of broadcast radio and television, social movements have systematically relied on communication technologies to pursue emancipatory forms of resistance and social change.[6] For example, during the 1960–70s a variety of second-wave feminist groups in Europe and North America published alternative newspapers to disseminate feminist ideas, organize movements, and challenge traditional gender norms and inequalities. These 'underground' newspapers provided a platform for voices often marginalized in mainstream media, allowing feminists to reach broader audiences and foster community engagement (McMillian, 2011). Founded in 1970 in Washington, D.C., for instance, the newspaper *Off Our Backs* was a significant publication within the US feminist movement, focusing on women's rights, activism, and issues relevant to women's lives, such as reproductive rights, gender discrimination, sexuality, and feminist theory.[7] The alternative newspaper helped to promote feminist solidarity and identity-building by providing a trustworthy source of information and knowledge for supporters

6 For a general discussion of this relationship, see: Milan, S. (2013). *Social movements and their technologies: Wiring social change*. Springer.

7 To access the archives of this newspaper, see, https://www.jstor.org/journal/offourbacks.

and participants in the movement (Mendes, 2011). As a form of RAM, this alternative news source empowered feminist mobilization by addressing issues that were either wholly ignored or strategically misrepresented by counter-feminist forces in mainstream media of the era.

The rise of DICTs has brought some new opportunities and resources to the field of RAM. Perhaps the most significant change is the relative ease with which DICTs allow social movement actors to develop and deploy their own RAM practices. This is a result, of course, of the widespread availability of web-based publishing tools and social media platforms to the general public. There are numerous examples of RAM across different digital media and formats.

A particularly well-known and established example of a RAM outlet that deploys many different forms of media production is *Adbusters*. Founded in 1989 as part of an emergent alter-globalization movement, this Canadian-based non-profit magazine and media foundation has described itself as a "global network of culture jammers and creators who are working to change the way information flows, the way corporations wield power, and the way meaning is produced in our society."[8] Especially known for coupling tactics of corporate 'ad spoofing' and visual protest with critical forms of expository journalism, *Adbusters* pays particular attention to problems of hyper-consumerism and ecological destruction generated by the conditions of contemporary capitalism.

Another North American example is the US-based independent, nonprofit news organization "*Democracy Now*!" which produces daily news programs and investigative pieces that cover domestic and global events from an explicitly progressive perspective. Established in 1996, its digital journalism exists in the form of text-based blogs and articles as well as video news stories and daily news broadcasts. *Democracy Now!* often features interviews with activists, scholars, and grassroots organizers as well as news about the actions of social movement groups from around the world.[9] In a similar vein, *Truthout* is a nonprofit news organization based in the US that was founded in 2001 and publishes investigative journalism, analysis, and commentary on global issues of relevance to progressive social movements, such as climate change, economic inequality, and racial justice.[10] According to its website, "if we want to grow movements for social transformation, we must build, tend, and cultivate transformative media". *Truthout* aims to "challenge mainstream narratives" and "support alternative perspectives" by "providing a platform for progressive and transformative ideas, through in-depth investigative reporting and critical analysis". It relies primarily on publishing text-based articles online.

An example of RAM operating on a more global scale and through a highly decentralized method is *The Independent Media Center*, also known as "*Indymedia*". Brought to life by the wave of alter-globalization protests in the 1990s, *Indymedia* is a decentralized global network of activist journalist collectives that

8 See, https://www.adbusters.org/.
9 See, https://www.democracynow.org/.
10 See, https://truthout.org/.

report and publish on a range of issues of relevance to progressive social movements operating across the Global North and South. One of the core principles of *Indymedia* is open publishing, which means that nearly anyone can publish content on the website without fear of editorial censorship. This allows for a diverse range of grassroots voices and perspectives to be represented, especially those of activists, community organizers, and marginalized groups working under conditions of repression and authoritarianism. Its aim is to provide an accessible digital platform that links together localized forms of activist journalism and viewpoints from around the world in order to help foster transnational social movement networks and alliances. Indymedia was founded as a very decentralized network of web-based RAM whereby contributors use methods of video-journalism as well as text-based publishing. Despite its spread to many corners of the Global North and South, however, it has experienced notable stagnation in usage and decline in visibility in recent years. This is most likely a result of heightened political repression on the one hand[11], and a growing reliance on social media platforms by younger generations of activists on the other.

A final example is that of "*The Young Turks*". This is a progressive social justice-oriented online news and commentary show that provides critical perspectives on politics, current events, and social issues. While its focus is mostly on the United States, it also addresses issues from around the world. Founded in 2002 as a public access radio talk show in Los Angeles, *The Young Turks* has attained a large following on YouTube and other social media platforms. By January 2024, for instance, its YouTube channel reported more than 5.8 million subscribers.[12] Unlike the other examples above, the programming leans toward progressive punditry rather than investigative journalism. Indeed, *The Young Turks* does not overtly display many direct linkages to the activities of specific social movement networks and campaigns. Nor does it seem to emerge directly from within social movement communities. Nonetheless, the selected topics and narrative framing of content by its producers are systematically deployed in a manner that echoes a variety of counter-hegemonic claims and social justice agendas expressed by many different types of social movements, such as environmentalism, feminism, anti-racism and migrant rights. Such qualities make it a form of RAM worthy of deeper investigation.

All of these aforementioned RAM initiatives offer just a few brief examples of media outlets that use journalism, storytelling, and media production to promote social change, amplify marginalized voices, and challenge power structures. There are, of course, many examples more out there. For instance, in recent years there have been innovative efforts in the tech sector to develop alternative DICTs that can challenge the hegemony of corporate-owned platforms such as WhatsApp and Twitter/X. The potential for new alternative platforms, such as Proton[13] or Masto-

11 For examples, the German government forcefully shut down Indymedia networks in 2017 for 'security' reasons: https://www.dw.com/en/interior-ministry-shuts-down-raids-left-wing-german-indymedia-site/a-40232965.

12 See, https://www.youtube.com/user/TheYoungTurks/about.

13 https://proton.me/about/impact.

don[14], to be deployed as RAM, however, remains to be seen given that the power of social media depends so heavily on network effects.[15]

In sum, RAM play an important role in facilitating the mobilization of trust/distrust in social movements. In addition to journalism, RAM encompass wide range of creative forms, including literature, film, music, visual art, and performance. Through storytelling, expression, and cultural critique, they challenge dominant cultural norms and narratives, and offer alternative visions of the world. In what follows, I address three mechanisms through which this unfolds: civic vigilance, social resistance and structural transformation.

Civic vigilance refers to how RAM are used in ways that promote the active and attentive engagement of citizens in monitoring, safeguarding, and promoting the principles of democracy, justice, and accountability within their communities and society at large. It involves a heightened critical awareness of social, political, and economic issues, as well as a willingness to call out abuses of power in the service of civil and human rights. With the use of RAM, civic vigilance encompasses forms of journalistic practice which allow members of the public to stay informed about current events, hold elected officials accountable, advocate for the rights of marginalized groups, and actively stimulate civic participation and activism. RAM outlets often engage in investigative journalism and critical analysis of societal issues. By uncovering injustices, corruption, and abuses of power that may be overlooked or ignored by mainstream media, they inform citizens and encourage them to scrutinize the actions of those in authority. RAM outlets often serve as watchdogs, exposing misinformation, propaganda, and manipulation by governments, corporations, and other powerful entities. By providing fact-based reporting and countering false narratives, they empower citizens to make informed decisions and resist attempts to manipulate public opinion. During the protests sparked by the killings of George Floyd, Breonna Taylor, and other Black individuals by police officers, mainstream media coverage in the US initially focused on the incidents themselves but often lacked context and depth regarding the broader issues of systemic racism and police violence underpinning these tragic incidents. Alternative media outlets, including online platforms, grassroots organizations, and independent journalists, however, played a crucial role in providing comprehensive coverage of the protests. For example, during the BLM protests, the RAM outlet known as "*The Root*" played a significant role in providing coverage and analysis of the movement.[16] The platform provided live feeds from protests and published articles, op-eds, and investigative reports that examined problems of systemic racism, police brutality, and the experiences of Black communities. Through its reporting and advocacy, *The Root* helped to promote civic vigilance by informing and empowering its audience to critically engage with the systemic challenges facing Black communities and to take action to address them.

---

14 https://joinmastodon.org/.

15 For more on this phenomenon, see: Katona, Z., Zubcsek, P., & Sarvary, M. (2011). Network effects and personal influences: The diffusion of an online social network. *Journal of Marketing Research, 48*(3), 425-443.

16 See https://www.theroot.com.

Social resistance refers to the ways in which RAM enable individuals to collectively combat the harms and inequities caused by a given set of power relations in society. It encompasses a range of actions, both discursive and non-discursive, aimed at disrupting or transforming existing social arrangements perceived as oppressive, discriminatory, or unjust. Wielded from within social movements, RAM have the potential to challenge dominant narratives promoted by mainstream media and powerful institutions. For example, during the Standing Rock protests by Native Americans against the Dakota Access Pipeline in 2016–2017, mainstream media coverage was very limited, prompting RAM outlets like *Indigenous Environmental Network* and *Unicorn Riot* to provide comprehensive reporting.[17] These outlets highlighted Indigenous concerns about environmental risks and human rights violations, amplifying the voices of affected communities and galvanizing public support. RAM outlets also facilitated communication among activists, mobilizing solidarity and raising awareness about the broader issues of environmental justice and Indigenous sovereignty. Through their reporting and advocacy, RAM played a crucial role in promoting social resistance against the pipeline construction while also advocating for indigenous rights. By providing coverage of these events, they not only raise awareness about specific issues but also facilitate the capacity of people to join the struggle and participate directly in social resistance efforts. In contexts where governments or powerful interests seek to suppress dissent and silence opposition, RAM play a critical role in resisting repressive narratives and policies.

Finally, structural transformation refers to the ways in which RAM can promote noticeable shifts in the social, cultural, economic and/or political characteristics of a given society, thus leading to perceptible reconfigurations of how power operates in society. Social movement induced forms of structural transformation can occur gradually over time, such as generational shifts in people's cultural values, or through more abrupt and intentional interventions, such as a reform of public policy. By covering and giving visibility to social movement claims and agendas, RAM outlets generally seek to promote structural transformations in society. By highlighting examples of citizen engagement and resistance, they inspire others to get involved in civic life and take action to address social injustices. An example of RAM triggering structural transformation can be found in the feminist mobilization campaign #MeToo. In 2017, this feminist campaign gained momentum when the American actress Alyssa Milano encouraged women to share their experiences of sexual harassment and assault by using the hashtag #MeToo on social media platforms like Twitter and Facebook. What began as a viral social media campaign quickly evolved into a global wave of feminist mobilization, with millions of women and non-binary persons sharing their stories and experiences of sexual misconduct, harassment, rape and abuse. During this campaign, "*Feministing*," an online RAM outlet dedicated to feminist activism, provided a strategic platform for survivors of sexual violence to share their stories as well as

17 See https://unicornriot.ninja/?s=standing+rock, and https://www.ienearth.org/?s=standing+rock.

for feminist activists and thinkers to publish articles and write opinion pieces.[18] The outlet played a significant role in amplifying feminist perspectives on sexual harassment and assault, contributing to a broader public discussion of sexism and male privilege. Through their reporting and advocacy efforts, RAM outlets such as *Feministing* helped to challenge gendered structures of patriarchal power in various sectors of society, such as the entertainment industry, corporate workplaces, and higher education. Amplified by RAM outlets, the widespread attention generated by the #MeToo movement led to increased public awareness, legal actions, policy reforms, and cultural shifts regarding issues of consent, accountability, and gender equality (Levy & Mattsson, 2023).

Conclusion

In this chapter, I explored the impacts of digital information communication technologies (DICTs) on the mobilization of trust within social movements. Delving into the interplay of trust and mistrust, my analysis focused on 'progressive' justice-oriented social movements. My discussion launched off by showing how the ubiquity of DICTS has created a seismic shift in the macro-structural landscape of social mobilization. On the one hand, it is evident that many new opportunities and avenues for empowerment have been created by the global proliferation of digital communication platforms. Yet, on the other hand, we cannot disregard the limitations and dilemmas of the digital age for social movements, notably in the form of new techniques of repression and control. Subsequently, my discussion extended beyond the theme of macro-level opportunities and obstacles, by delving into the agency and strategic capacities of social movement actors. In this context, I looked at how DICTs can serve as empowering tools wielded by activists, with a particular focus on the pivotal role played by radical alternative media (RAM) and the ways in which grassroots organizers can harness the power of DICTs through activist journalism and online publishing. Such digital tools help to amplify the voices and presence of social movements in society, despite the constraining power of corporate media and surveillance technologies. In sum, this chapter underscores the duality inherent in the digital revolution's impact on social mobilization—a double-edged sword that offers both unparalleled opportunities for empowerment and agency, while simultaneously posing formidable challenges in the form of surveillance, censorship, and manipulation.

References

Adler, E., & Drieschova, A. (2021). The epistemological challenge of truth subversion to the liberal international order. *International Organization, 75*(2), 359-386. https://doi.org/10.1017/S0020818320000533

18 See https://feministing.com/?s=me+too.

Atton, C. (2002). News cultures and new social movements: Radical journalism and the mainstream media. J*ournalism Studies, 3*(4), 491-505. https://doi.org/10.1080/1461670022000019209

Beck, U. (1986) *The Risk Society.* Sage.

Beer, D. (2019). *The Social Power of Algorithms.* Routledge.

Cammaerts, B. (2021). The new-new social movements: Are social media changing the ontology of social movements? *Mobilization: An International Quarterly, 26*(3), 343-358. https://doi.org/10.17813/1086-671X-26-3-343

Canella, G. (2022). *Activist Media: Documenting movements and networked solidarity.* Rutgers University Press.

Caren, N., Andrews, K. T., & Lu, T. (2020). Contemporary social movements in a hybrid media environment. *Annual Review of Sociology, 46*, pp. 443-465.

Clark, M. D. (2019). White folks' work: Digital allyship praxis in the # BlackLivesMatter movement. S*ocial Movement Studies, 18*(5), 519-534. https://doi.org/10.1080/14742837.2019.1603104

Cleaver, H. (1998). The Zapatistas and the Electronic Fabric of Struggle. In J. Holloway & E. Peláez (Eds.), *Zapatista! Reinventing Revolution in Mexico* (pp. 81-103), London: Pluto Press.

Della Porta, D. (2012). Critical trust: Social movements and democracy in times of crisis. *Cambio: rivista sulle trasformazioni sociali, 4*(2), 33-43.

Dencik, L., Hintz, A., & Carey, Z. (2018). Prediction, pre-emption and limits to dissent: Social media and big data uses for policing protests in the United Kingdom. *New Media and Society, 20*, 1433–1450. https://doi.org/10.1177/1461444817697722

Dencik, L., & Leistert, O. (Eds.). (2015). *Critical perspectives on social media and protest: Between control and emancipation*. Rowman and Littlefield.

Earl, J., Maher, T. V., & Pan, J. (2022). The digital repression of social movements, protest, and activism: A synthetic review. *Science Advances,* 8(10). https://doi.org/10.1126/sciadv.abl819

Feldstein, S. (2021). *The rise of digital repression: How technology is reshaping power, politics, and resistance.* Oxford University Press.

Fischer, R., & Jarren, O. (2024). The platformization of the public sphere and its challenge to democracy. *Philosophy and Social Criticism, 50*(1), 200-215. https://doi.org/10.1177/01914537231203535

Glenn, C. L. (2015). Activism or "Slacktivism?": Digital media and organizing for social change. *Communication Teacher, 29*(2), 81-85. https://doi.org/10.1080/17404622.2014.1003310

Guzmán, S. G. (2015). "Should I Trust the Bank or the Social Movement?" Motivated Reasoning and Debtors' Work to Accept Misinformation. *Sociological Forum,* 30(4). https://doi.org/10.1111/socf.12201

Heidemann, K. (2022) You Are Not Alone: Social movements and the building of interpersonal alliances through popular education in Belgium. In F. Orban & E. Strand Larsen (Eds.), *Living Alliances, Leaving Alliances: Interdisciplinary Perspectives* (pp. 135-146). Waxmann.

Kallis, A. (2013). Far-right "contagion" or a failing "mainstream"? How dangerous ideas cross borders and blur boundaries. *Democracy and Security, 9*(3), 221-246. https://doi.org/10.1080/17419166.2013.792251

Kellner, D. (2021). T*echnology and democracy: Toward a critical theory of digital technologies, Technopolitics, and Technocapitalism.* Springer.

Kennis, A. (2022). *Digital-Age Resistance: Journalism, Social Movements and the Media Dependence Model.* Routledge.

Keremoğlu, E., & Weidmann, N. B. (2024). "Authoritarianism and digital communication". In N. Lindsteaedt & J. den Bosch (Eds.), *Research Handbook on Authoritarianism* (pp. 128-138). Edward Publishing.

Koehler, J. (2023) "Tribalism in the Age of Social Media: Strategies for fostering connection in the digital world". Available at https://www.psychologytoday.com/us/blog/beyond-school-walls/202304/tribalism-in-the-age-of-social-media, accessed January 13, 2024.

Levy, R. E., & Mattsson, M. (2023). The effects of social movements: Evidence from# MeToo. S*ocial Science Research Network.* Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3496903.

Mayo, M. (2020). C*ommunity-based learning and social movements: Popular education in a populist age.* Policy Press.

McChesney, R. W. (2013). *Digital Disconnect: How capitalism is turning the internet against democracy.* The New Press.

McChesney, R. W. (2016). Communication activism/Missing in action: Engaged US communication research in the context of democratic decline and the digital revolution. *International Journal of Communication,* 4009-4016.

McMillian, J. (2011). *Smoking typewriters: The sixties underground press and the rise of alternative media in America*. Oxford University Press.

Mendes, K. (2011). Framing feminism: News coverage of the women's movement in British and American newspapers, 1968–1982. *Social Movement Studies, 10*(1), 81-98. https://doi.org/10.1080/14742-837.2011.545228

Meyer, M., & Molyneux-Hodgson, S. (2010). Introduction: The dynamics of epistemic communities. S*ociological Research Online,* 15(2), 109-115. https://doi.org/10.5153/sro.2154

Milan, S. (2015) Mobilizing in time of social media: From a politics of identity to a politics of visibility. In L. Dencik & O. Leisert (Eds.), *Critical Perspectives on Social Media and Protest: Between Control and Emancipation* (pp. 53-71). Rowan and Littlefield.

Miller, M. L., & Vaccari, C. (2020). Digital threats to democracy: Comparative lessons and possible remedies. *The International Journal of Press/Politics,* 25(3), 333-356. https://doi.org/10.1177/1940161220922323

Mortensen, M., Neumayer, C., & Poell, T. (2018). S*ocial Media Materialities and Protest.* Routledge.

Mueller, G. (2021). *Breaking things at work: The Luddites are right about why you hate your job.* Verso Books.

Mühlfried, F. (2019). *Mistrust: A global perspective*. Springer.

O'Brien, T. (2015). Social control and trust in the New Zealand environmental movement. *Journal of Sociology, 51*(4), 785-798. https://doi.org/10.1177/1440783312473188

Plaut, S. (2012). 'Cooperation is the story'–best practices of transnational indigenous activism in the North. *The International Journal of Human Rights, 16*(1), 193-215. https://doi.org/10.1080/13642987.2011.630498

Rodríguez, C., Ferron, B., & Shamas, K. (2014). Four challenges in the field of alternative, radical and citizens' media research. *Media, Culture and Society, 36*(2), 150-166. https://doi.org/10.1177/0163443714523877

Rosanvallon, P. (2008). *Counter-democracy: Politics in an age of distrust.* Cambridge University Press.

Rossi, F. M. (2023). Democracy as a trust-building learning process: Organizational dilemmas in social movements. *International Journal of Comparative Sociology.* https://doi.org/10.1177/002071522311965.

Shirky, C. (2008). *Here comes everybody: The power of organizing without organizations.* Penguin.

Strasser, B., Baudry, J., Mahr, D., Sanchez, G., & Tancoigne, E. (2019). "Citizen science"? Rethinking science and public participation. *Science and Technology Studies,* 32(2), 52-76. https://doi.org/10.23987/sts.60425

Suh, H., & Reynolds-Stenson, H. (2018). A contingent effect of trust? Interpersonal trust and social movement participation in political context. *Social Science Quarterly,* 99(4), 1484-1495. https://doi.org/10.1111/ssqu.12515

Sztompka, P. (1998). Trust, distrust and two paradoxes of democracy. *European Journal of Social Theory,* 1(1), 19-32. https://doi.org/10.1177/136843198001001003

Tufecki, Z. (2017). *Twitter and Tear Gas: The power and fragility of networked protest.* Yale University Press.

Van Laer, J., & Van Aelst, P. (2010). Internet and social movement action repertoires: Opportunities and limitations. *Information, communication and society,* 13(8), 1146-1171. https://doi.org/10.1080/13691181003628307

Welch, M. R., Rivera, R. E., Conway, B. P., Yonkoski, J., Lupton, P. M., & Giancola, R. (2005). Determinants and consequences of social trust. *Sociological Inquiry, 75*(4), 453-473. https://doi.org/10.1111/j.1475-682X.2005.00132.x

Wodak, R. (2015). *The Politics of Fear: What right-wing populist discourses mean*. Sage.

Yates, S., & Rice, R. E. (Eds.). (2020). *The Oxford Handbook of Digital Technology and Society.* Oxford University Press.

Zuckerman, E. (2021). *Mistrust: Why losing faith in institutions provides the tools to transform them.* WW Norton and Company.

Truth – the First Casualty : Trust-building Through Fact-checking in the Israel-Hamas Conflict

Elin Strand Larsen

Introduction

On October 7, 2023, the political and military organization governing the Gaza Strip, Hamas, committed a terrorist attack against Israel, killing more than 1,400 people and taking at least 240 hostages (BBC, 2023). Israel responded to the attack by declaring war, launching air strikes, and blocking the supply chain of water, food, and fuel to the Gaza population. According to the Gaza Health Ministry, more than 8,000 Palestinians were killed in the first three weeks of the war (AP-News, 2023). After six months of war in Gaza more than 30,000 Palestinians have been killed (Picheta & Salman, 2024).

The conflict between Israelis and Palestinians is not new, but the sudden escalation and war declaration made headlines and caused concern all over the world in 2023. We also saw a number of misleading and false claims related to the war surface and circulate online. In this chapter, I will focus on trust building and handling of misleading information related to the Israel-Hamas conflict.

I have chosen three central fact-checking organizations in Scandinavia all involved in fact-checking misleading and incorrect information related to the conflict – *Faktisk* in Norway (Faktisk, 2024), *Källkritikbyrån* in Sweden (Källkritikbyrån, 2024), and *TjekDet* in Denmark (TjekDet, 2024). The selection of fact-checking organizations is based on the idea that the three Scandinavian countries are quite similar in terms of both media situation and political governance. However, Norway differs from both Sweden and Denmark, in their long-term engagement in the conflict, seeking to negotiate a two-state solution (NorwegianGovernment, 2022) and the Norwegian recognition of Palestine as a state (NorwegianGovernment, 2024).

The study aims to answer the following research question:

How did Faktisk, Källkritikbyrån, and TjekDet seek to build trust through fact-checking in the two first months of the Israel-Hamas conflict?

Through research interviews with selected representatives of Faktisk, Källkritikbyrån, and TjekDet, as well as some selected examples of fact-checks from the first two months following the Hamas terrorist attack (Oct. 7–Dec. 7), I want to investigate how fact-checkers in the three Scandinavian countries handle fake news and disinformation related to an escalating conflict and war far away from home.

I will start with giving the theoretical background for fake news and fact-checking in war time, before I present the method, as well as the three representatives of Faktisk, Källkritikbyrån, and TjekDet interviewed for this study. In the analysis, I will focus on the five main strategies of fact-checking aimed at building trust between the fact-checkers and their audience: 1) Being honest about uncertainty, 2) From color categories to carefully written conclusions, 3) Providing proof, but protecting the audience, 4) Public dialogue, and 5) Networking. In the conclusion, I summarize and discuss how these strategies might build trust in the fact-checking process.

Fake News and Fact-checking in War Time

"The first casualty when war comes is truth" is a quote ascribed to many different actors, ranging from ancient Greek tragedian, Aeschylus, to the US Senator Hiram Warren Johnson and in writing by the British politician and writer, Arthur Ponsonby, in his book *Falsehood In War Time* (Ponsonby, 1928). The quote addresses the information chaos emerging in the beginning and even before the war breaks out, as well as the opposing perspectives and the battle for defining the main narrative in war time – who is the good guy and who is the bad guy? In this battle, the truth becomes a casualty just like soldiers' lives.

While the viral spread of fake news is a rather new aspect of today's information wars, the spread of disinformation and propaganda goes back to World War I, when it was used both as a weapon against the enemy and a tool in strengthening national war effort. Jowett & O'Donnell (2012) differentiate between black, white, and grey propaganda. White propaganda is from a well-known source and the information is considered accurate. However, the message "is presented in a manner that attempts to convince the audience that the sender is the "good guy" with the best ideas and political ideology" (Jowett & O'Donnell, 2012, p. 17). Black propaganda is when the source of the information is unclear and the message is based on "lies, fabrications, and deceptions" (Jowett & O'Donnell, 2012, p. 18). Most of today's propaganda lies in a grey area where "the source may or may not be correctly identified, and the accuracy of the information is uncertain" (Jowett & O'Donnell, 2012, p. 20).

"Before the war", Jacques Ellul (1973) argues "propaganda is a substitute for physical violence; during the war it is a supplement to it" (Ellul, 1973; Jowett & O'Donnell, 2012, p. 4). According to Ellul, propaganda has become an inescapable necessity for everyone, and he criticizes the naïve belief that facts always contradict propaganda (Ellul, 1973). A strategic selection and arranging of facts may be an effective propaganda strategy, both in war and peace, and because of this fact-checking is not worth much without deep knowledge of the societal and historical context.

Hiebert (2003) shows how public relations and propaganda were an essential part of the US government's war in Iraq in 2003: "Embedding journalists, staging showy briefings, emphasizing visual and electronic media, and making good TV entertainment" (Hiebert, 2003, p. 254) were all important in the war communication. Boyd-Barrett's study (2004) of The New York Times coverage, as well as the use of Judith Miller as their correspondent, also found strong mainstream media support of the Iraq invasion. According to Boyd-Barrett, the propagandized news selection of NYT is an example of the "buying out" of journalists and their publications by intelligence and related special interest organizations" (Boyd-Barrett, 2004, p. 435).

The government and the media speaking with one voice resulted in US opinion polls showing public backing at home, while internationally the American government was losing war support. Hiebert (2003) argues in conclusion that governments have to keep future wars short and "clean" as long as there is "alternative press such as the internet, to report the facts of war and point out the deceptions" (Hiebert, 2003, p. 254). However, in these articles from 2003 and 2004, Hiebert and Boyd-Barrett missed mentioning the future potential of viral mis- and disinformation of the internet, as well as fake news.

Fake news is often defined as "information that has been deliberately fabricated and disseminated with the intention of deceiving and misleading others into believing falsehoods or doubting verifiable facts; it is disinformation that is presented as, or is likely to be perceived as, news" (McGonagle, 2017, p. 203). Tambini (2017) divides fake news into six categories. These are foreign interference, ad-driven fake news, parody/satire, bad journalism, form of insult, and challenging hegemony. I would like to argue that we need a seventh type of fake news which includes falsehoods and disinformation in war and territorial conflicts. This type of fake news might be disseminated due to lack of knowledge or an interest in presenting the battle in one specific way. Due to the lack of access to the war zone, as well as the importance of perceived strength on the battlefield, fake news becomes both a natural and crucial part of war, sometimes with very serious consequences.

The historical roots of fact-checking go back to the first political fact-checkers established in the US, like FactCheck.org, Washington Post Fact Checker, and PolitiFact (Graves, 2016). While the newsrooms have a long tradition of proofreaders to verify information *before* it goes to print, the new fact-checking movement did the opposite: "They investigate claims that are *already* in the news and publish the results as a news story" (Graves, 2016, p. 8). The political fact-checking in American journalism has been an important role model for Scandinavian fact-checkers, like Faktisk, Källkritikbyrån, and TjekDet. This might be one of the reasons for their focus on national politics over international affairs.

According to a study by Sigurd Allern (2019) of the first 277 fact-checks by Faktisk.no, none of them addressed topics like "EU, Brexit, the wars and conflicts in the Middle East, developments in the US, Russia, or China" (Allern, 2019, pp. 177-178). Another reason for this national limitation might be the lack of resources and knowledge concerning international conflicts and foreign policy (Allern, 2019, p. 178; Allern & Pollack, 2019, pp. 284-295). This practice changed with Russia's invasion of Ukraine and the Gaza war. In Norway Faktisk Verifiserbar was launched for a period in 2022 to verify visual content from Ukraine and Russia.

When faced with a new war in Israel and Palestine, Faktisk relaunched Verifiserbar in 2023 "due to the precarious need for fact-checking content, news, and claims from the Middle East" (Verifiserbar, 2023). All pictures and videos verified by Faktisk Verifiserbar are immediately shared through a database available for all Norwegian media outlets.

Bente Kalsnes (2019) presents Russia as an example of how disinformation from one state is used strategically as a weapon, not only to dominate, but also to create chaos and polarization in other countries. Long before the Russian invasion of Ukraine in 2022, the Ukrainian fact-checking website, StopFake, showed how the Russian propaganda- and information war was based on four D's: Dismissal, distortion, distraction, and dismay (Nimmo, 2015). They would deny allegations, distort information to serve their narrative, cause distraction to turn the attention away from their own activities, and warn of the consequences of going against Russia (Nimmo, 2015).

Also using StopFake as an example, Khaldarova & Pantti (2016) analyzed the fact-checking organization's attempts to counter the strategic narratives about Ukraine published on the Russian TV station Channel One between 2013 and 2015. StopFake used different debunking methods from pointing out the baseless evidence, inconsistency in details, and image verification. Twitter users also contributed to the fact-check by finding images used in other contexts or making their own reports from the scene to counter the story of Channel One (Khaldarova & Pantti, 2016).

Magallón-Rosa, Fernández-Castrillo & Garriga (2023) studied the types of hoaxes and false information related to the 2022 war in Ukraine as fact-checked by the six Spanish fact-checking organizations. They found that unlike the fake news related to the Covid-19 pandemic, in which the preferred format was text, the Ukrainian war hoaxes were primarily based on visuals. Images became the main "focus of attention, monitoring, and verification by fact-checking organizations" (Magallón-Rosa et al., 2023, p. 2). The researchers also found that many of the same hoaxes and fake news stories were fact-checked by two or more of the fact-checking organizations, which is a good sign, since repetition makes the refutations and corrections by the fact-checkers more visible (Magallón-Rosa et al., 2023).

Method

This study is based on qualitative interviews with one representative from each of the three fact-checking organizations, combined with a selection of fact-checks serving as examples in the analysis. In the first two months following the Hamas terrorist attack (Oct. 7–Dec. 7), Faktisk had 17 publications, categorized as 13 articles and 4 fact-checks, while TjekDet had 16 publications, divided into 12 fact-checks and 4 insights. Källkritikbyrån had only 3 publications related to the conflict, including two investigations and one combination of investigation and guide. This might be due to a smaller editorial office and staff at Källkritikbyrån, as well as differences in priorities.

I have interviewed one editor/fact-checker from each of the three fact-checkers: Olav Østrem from Faktisk (Østrem, 18.12.23), Åsa Larsson from Källkritikbyrån (Larsson, 05.01.24), and Andreas Søndergaard Petersen from TjekDet (Søndergaard Petersen, 07.12.23). While the interview with Østrem took place at the editorial office of Faktisk at Pressens Hus in Oslo, the other two interviews were conducted using Zoom video conference for video and audio communication.

Before the interviews, all the respondents were presented with the project and their contribution to the study. I used a semi-structured interview guide, with a fixed set of questions, while also being open for follow-up questions. The interviews were recorded, transcribed, and quotes approved by the respondents. All the quotes in Norwegian, Swedish, and Danish have been translated to English for this chapter. The study is reported to and approved by Sikt's data protection services for research (Sikt.no, 2023).

Faktisk, Källkritikbyrån, and TjekDet

The Norwegian fact-checking project, Faktisk, was launched just before the parliamentary election in 2017. The project was based on a collaboration between four of Norway's biggest news outlets: VG, Dagbladet, NRK, and TV2. Later, Polaris Media and Amedia joined the project. According to Faktisk, they want to contribute to an "open, inclusive and fact-based public conversation" by reviewing "the basis of current claims that affect our perception of reality" (Faktisk, n.d.). On their website, Faktisk.no is presented as a "non-profit organization and independent editorial board for fact-checking the social debate and the public discourse in Norway" (Faktisk, n.d.-b). According to one of my previous studies of Faktisk, I found that the Norwegian fact-checking project focused on fact-checking the claims of politicians in the period leading up to the Norwegian parliamentary election in 2017, while the French fact-checking project, CrossCheck, primarily looked into viral stories and online hoaxes related to the 2017 presidential election in France (Larsen, 2019).

The representative of Faktisk, Olav Østrem, is a lawyer by education, but switched to journalism early in his professional life. He has been a journalist since 2002, working in NRK radio and the newspaper Klassekampen. At the time of the interview, he had been working for Faktisk about one year and was both news editor and editor-in-chief at Faktisk, combined with being the project leader of Faktisk Verifiserbar: "Right now I am constituted in the role of both general manager and responsible editor, in addition to my responsibilities as news editor. […] As a news editor, [the main task] is to lead the editorial office from day to day, ensuring workflow and prioritization, that we complete the fact-check and publish" (Østrem, 18.12.23).

The Swedish fact-checker, Källkritikbyrån, was first launched in 2014 as Viralgranskaren by Åsa Larsson, Linnéa Jonjons, and Jack Werner, while they were working as reporters for the free daily newspaper in Sweden called Metro. The newspaper ceased publication in 2019, but the fact-checking project has continued under the name of Källkritikbyrån. Källkritikbyrån wants to help people become more "confident and aware users of the internet" (Källkritikbyrån, n.d.-a). They "systematically review online claims and pass on the knowledge [they] have to the public" (Källkritikbyrån, n.d.-a). Åsa Larsson, the representative of Källkritikbyrån, has been a journalist since 2005 working in Swedish newspapers like Dagens Nyheter and Svenska Dagbladet, as well as the news agency, TT, Sveriges Radio, and the previously mentioned Metro: "We had already launched Viralgranskaren when we got to know that the newspaper was to be shut down. Then we thought it was too bad that this whole project was to be wasted, so we started up on our own [with] Källkritikbyrån in 2019 and then we launched it in 2020" (Larsson, 05.01.24). The editorial office of Källkritikbyrån is rather small, with Larsson as the only permanent employee.

Last, we have the Danish fact-checker, TjekDet. Tjekdet was launched in November 2016, as an initiative by Lisbeth Knudsen and Mandag Morgen. Today the fact-checking project is owned by the non-profit organization *Foreningen TJEKDET – National portal for bekæmpelse af fake news* (TjekDet, n.d.-b). They are, according to the presentation on their website, "all of Denmark's political independent fact-checking media" (TjekDet, n.d.-b). They check "claims in the public debate and corrects or add nuances where necessary" with a main focus on "dealing with misinformation and disinformation" (TjekDet, n.d.-b). The representative of TjekDet, Andreas Søndergaard Petersen, is journalist by education and he first joined TjekDet as a student intern in 2019. From 2020 he has been full time employee at TjekDet, and in 2021 he became editor with the responsibility of "chairing the editorial meetings, choosing which stories [they] should invest in, being involved and developing ideas from the start, sparring with [the] journalists […] and last, making the finishing touches on the articles, [as well as] having the general overview of the publication flow" (Søndergaard Petersen, 07.12.23).

All three fact-checkers collaborate with Facebook (Meta) through the *Third-party fact-checking initiative*, evaluating Facebook posts flagged as potentially misleading information, adding "one of three labels to the content – "Missing context", "Partly false", or "False" (Bengtsson et al., 2021, p. 11). However, this collaboration has also been criticized, especially from Swedish media outlets like SVT, SR, and Dagen Nyheter (Allern & Pollack, 2019, p. 287). According to Olle Zachrison, news commissioner at Swedish Radio (SR), they did not "see their role as a partner to Facebook, helping them sanitize their platform" (Kihlström (2018) in Allern & Pollack, 2019, p. 287).

Fact-checking When Faced with Uncertainty

In the following analysis, based on interviews and a selection of fact-checks, I will focus on five main strategies used by all of the three fact-checking organizations to build trust through fact-checking the Israel-Hamas conflict. The first strategy is to be honest about uncertainty. All the three fact-checkers used so-called "This is what we know" articles in cases where they had limited access and ability to verify information from the Israel-Hamas conflict. According to Olav Østrem at Faktisk, this was especially visible in the two categories of articles presented by Faktisk:

> The problem here [is] that we are not present there [and] we have very poor access to information from Gaza. In the beginning, only IDF [The Israel Defence Forces] could present things and invite journalists. It's difficult to get to the bottom of [a fact-check] and it's not obviously easy [to see that] it's fact-check material. […] We have two different [categories] at Faktisk: [A] basic fact-check [with] a concrete claim [and] then we will try to say something about that claim. […] Then we have what we call articles, [where] we try to give some background information (Østrem, 18.12.23).

However, Østrem was clear on not losing sight of their main mission as a fact-checker: "We can't just make articles that can be placed anywhere, because then we have no function. [We could have written] the history of Hamas or the background for the many wars in the Middle East, but [in these types of] background articles [I don't] feel we're bringing anything new" (Østrem, 18.12.23).

One of the first fact-checks, published by all three fact-checkers, focused on the Hamas attack on an Israeli kibbutz and the killing of 40 babies. The investigation by Källkritikbyrån started with the claim by a correspondent from an Israeli TV channel, *i24NEWS*, allowed to visit the kibbutz together with the Israel Defense Forces (IDF). According to the correspondent she had been told by the IDF that 40 dead babies were taken out on gurneys, some of them even beheaded (Larsson, 2023c). While many media outlets and people around the globe continued to share this information, the fact-checkers argued that the claim could not be confirmed.

Faktisk and TjekDet brought the same story. While Faktisk focused on how many international and Norwegian media outlets had to adjust their headlines to the unconfirmed information (Lønrusten et al., 2023), TjekDet gave an overview of the claim's viral spread – "how it started, who said what and why the claim can't be verified nor debunked" (Gregersen & Engelbrecht, 2023). All the three fact-checkers noticed the challenge of presenting a fact-check with no clear conclusion as to the validity of the claim. Åsa Larsson stressed the difference between verifying a claim and making uncertainty visible:

> Some might get the idea that we are investigating whether 40 babies were beheaded or not, and we cannot determine that from a distance. It is something the local media, at best, should find out about. But what we are investigating is rather: [...] People have asserted something with certainty. We show here that there is no such certainty (Larsson, 05.01.24).

Andreas Søndergaard Petersen from TjekDet was, on the other hand, not so sure about the value of publishing a fact-check with no clear verdict as to the validity of the claim. He questioned whether a fact-checker should stay away from articles where they end up concluding that they can't conclude:

> There was a very widespread post that Hamas had beheaded 40 babies in an Israeli kibbutz […] where we concluded that there was no real documentation, but also that we couldn't argue that it didn't happen, because we simply didn't know. We did receive some criticism from people who had read the article, asking: What was the conclusion? [So], maybe we should stay away from those articles where we end up concluding that we can't conclude anything? (Søndergaard Petersen, 07.12.23)

An additional challenge to the fact-checking process was the limited time available to verify information. Both Østrem and Larsson addressed the need for finding a balance between speed and accuracy when fact-checking stories from the Israel-Hamas conflict. According to Østrem there was especially a pressure on Faktisk Verifiserbar from other Norwegian media outlets:

> [A story] we had last week [about] a hospital where babies were left to die. We spent a lot of time [on this fact-check] because you had to be sure. It is difficult because there is [a] need for [Faktisk] Verifiserbar to support the rest of the [media] industry. Dagsrevyen [Evening news program at NRK] and others ask [us and] they would like to have an answer right away (Østrem, 18.12.23).

Larsson from Källkritikbyrån addressed the importance of fast fact-checking to be part of the online conversation and continuously correcting misinformation:

> We try to follow the news flow. In the best of worlds, we should be able to check up on minor things and then quickly be able to say something to be part of [the conversation]. If you do it too late, then the damage has already been done. […] We're only able to influence them [when] they have just read [the story], not three weeks later [when] they have taken an interest in fourteen other topics (Larsson, 05.01.24).

However, sometimes fact-checking goes wrong. The fact-checkers get new information or realize that they need to edit previously published articles. All of the three fact-checkers were open about changes done to earlier publications, by having their own logs of changes (Faktisk, n.d.-a; TjekDet, n.d.-a) or marking the changes directly in the articles with text in italics (Källkritikbyrån, n.d.-b). Søndergaard Petersen at TjekDet gave an example of corrections made to an article about the expression "From the river to the sea – Palestine will be free" (Engelbrecht & Brandt, 2023):

> An example of an article we received a lot of criticism for was when we created an insight article, because people thought it was a fact-check. [Our] conclusion was that [...] it is all about context, and [the expression] can be used with anti-Semitic purposes, but [the readers] didn't find that quite clear enough in our bullet points. So, we changed it [so that] it became clearer what the purpose of the article and the conclusion was (Søndergaard Petersen, 07.12.23).

Østrem from Faktisk also addressed the importance of being open about changes done to their previous fact-checks. This was an important part of trust building between the fact-checker and their readers. However, he also stressed that they had never depublished an article:

> [If there is] something that is unclear or that we have missed [something] , then we have to go in and correct it. […] We must play with open cards. It is very important. It shouldn't be too dangerous for us to speak up about [corrections] [because] that's what we want everyone [else] to do. We have not yet depublished an article. It's like the worst thing because it means that you have published a fact-check where the conclusion or premise is completely wrong, and that correction does not make up for the damage either. […] This must be taken seriously because it has to do with credibility (Østrem, 18.12.23).

Sometimes the uncertainty might lead to fact-checks not being published at all. Søndergaard Petersen was faced with a dilemma after two people at TjekDet spent a whole week fact-checking a widely spread story only to conclude that it was not worth publishing. The story was about a Palestinian woman named Israa Jaabis and the reason for her burnt and disfigured face. The story of Jaabis became relevant because she was part of the prisoner exchange between Israel and Palestine. According to Israel she was a suicide bomber who failed in her attack. The car she drove exploded and she was badly hurt in the explosion. The Palestinian side argued, on the other hand, that she was only moving into a new house and a gas bottle exploded in her car by accident. She was forced to stay in her car by a police officer, did not get proper treatment for her burns and was thrown into an Israeli jail:

> So, these are the two stories [circulating] and one of the stories had 24 000 shares on Facebook. This is a high number of shares in a Danish context, but we couldn't document anything. We might as well have written: Noone knows what happened that day except Israa Jaabis and the [police] officer. […] We tried to find information and access documents from the court. We tried to reach her lawyer and the Israeli authorities, but they weren't very cooperative. It would just be a little too poor an article if we couldn't make the reader just a little bit wiser, and I don't think we could. [So], that's why we ended up [not publishing] (Søndergaard Petersen, 07.12.23).

The challenges of fact-checking stories with no clear-cut conclusion have also resulted in changes to how the fact-checkers present and categorize their articles. More about this in the next part of the analysis.

From Color Categories to In-depth Explanation

The second strategy to build trust is to go from simple color categories to carefully written conclusions explaining the fact-checking process and trying to educate their audience. From the launch of Faktisk in 2017, the Norwegian fact-checker presented the results of their fact-checks on a graded scale with the labels "Actually true", "Actually partly true", "Actually not sure", "Actually partly wrong", or "Actually completely wrong" (Larsen, 2019, pp. 53-54). Some of the labels were combined with colors – green for the fact-checks marked as "Actually true" and red for the fact-checks marked as "Actually completely wrong".

From the very start they experienced that not all claims allowed unambiguous conclusions: "In a closer account, Faktisk describes the assessment "Actually not sure" to be used when they are unable to verify the claim due to lack of open and credible sources, or because the documentation cannot provide a clear conclusion" (Larsen, 2019, p. 54). In 2021, Faktisk removed the colored labels, arguing that "colors have overshadowed the facts" (Egeberg, 2021). Østrem at Faktisk saw both the advantages and challenges of color categories, but argued that the fact-checking process was much more interesting than the conclusions:

> The advantage [of color categories] is the tabloid appeal [and] it was an easy way to get attention, [...] but perhaps it took the focus away from what we wanted to say. I would say that the conclusion itself is not the most interesting part, but how we [arrived at the conclusion]. It is the presentation [of the fact-check] that makes it interesting, and it means that the readers can form their own opinion and not […] blindly trust the conclusion. […] We should rather work on the presentation [to make it] appetizing, [with the] headline and introduction and all that (Østrem, 18.12.23).

Åsa Larsson also addressed the stamps and the layout of different conclusions when working for Viralgranskaren at Metro. The colors and categories disappeared when they relaunched as Källkritikbyrån:

> At Metro and Viralgranskaren we had stamps [to mark our conclusions]. Previously fact-checkers had a whole layout for [this]. Then you could take that article and beat it over the heads of those who disagreed. It was "weaponized" in a way that was bad for fact-checking. When we went from Viralgranskaren to Källkritikbyrån, we took away the stamps. We no longer use stamps, but write [clearly what] we arrive at (Larsson, 05.01.24).

However, Larsson noticed how the cooperation with Meta through the *Third-party fact-checking program* forced them into categorizing their fact-checking results. She worried that these strict categories could create misunderstanding, distance, and opposition between the fact-checker and the audience: "Instead of reading: Here is a fact-check of this story, they read: I thought it was like that, but you say it's fake? I don't like that. You create an opposition already from the start" (Larsson, 05.01.24). Søndergaard Petersen from TjekDet also mentioned the categories of conclusions used by Facebook and Google:

> We can categorize a [Facebook] post with a claim as […] false, partially false, or missing context, then all those who have shared [the post] will receive a message

> that it has been fact-checked by an independent fact-checker. [For Google] we present: What is the claim, who claims it, and what is our conclusion? Then the conclusion from TjekDet is visible when one of our articles is looked up via Google (Søndergaard Petersen, 07.12.23).

Choosing what to include or not in a fact-check is an important part of the fact-checking process and presentation. On the one hand, the fact-checkers want to be as open as possible about their sources and how they test and verify pictures, videos, and stories. On the other hand, they need to take into consideration the effect and impact of graphic visuals of war on the audience. This brings us to the next challenge of fact-checking the Israel-Hamas conflict.

See it to Believe it – The Visuals of War

The third strategy to build trust is to present the evidence, even if that means linking to very graphic material that can be considered disturbing for the audience. In a news article (Greger, 2023) about fact-checking the Israel-Hamas conflict and war, the fact-checkers from Faktisk Verifiserbar talked about the extremely graphic videos and photos from Gaza they had to work with. Project leader, Silje Førsund, described their work as “counting body parts” (Greger, 2023). The fact-checkers also noticed how the visuals from Israel and Gaza differed from the war in Ukraine in terms of number of wounded and dead children they had to see.

When asked about the visual material fact-checkers at Faktisk had to work with, Østrem also stressed the number of children involved, as well as the importance of debriefing and help from a psychologist specialized in stress and trauma management:

> There is such a young population [in Gaza] and there are a lot of children involved, so that makes it completely different. […] There has been a lot of focus on [how fact-checkers deal with violent content]. [A psychologist] meets the group every two weeks, where we have this type of debriefing, so we can talk about these things. […] It is not up to each fact-checker whether they want a debriefing or not. It is mandatory and proactive (Østrem, 18.12.23).

When it comes to what the fact-checkers choose to publish, they try to balance between transparency in the fact-checking process and a concern for their audience. Both TjekDet and Källkritikbyrån included links to disturbing visual material in some of their articles, but also warned their readers before following these links.

In fact-checking pictures of dead children from Gaza which were claimed to be dolls, with links to AFP Fact Check (2023), Källkritikbyrån warned in their article about “unpleasant pictures” (Larsson, 2023a). When asked about why they included such warnings, Larsson focused on their responsibility in publishing and linking to this type of visuals:

> We take responsibility for what we publish. We choose images documenting what we want to document but are not unnecessarily brutal. […] When you link to something that you know will be unpleasant to look at, then there is a possibility to warn [the readers] before clicking […] It is so difficult for people to protect themselves from such images and one should be able to choose: Do I want to start my morning by choking on my morning coffee and ending up with that picture on your mind forever (Larsson, 05.01.24)?

According to Larsson there was a difference between the Nordic countries and the English-speaking world as to anonymizing pictures of dead people and children in particular. The links to these pictures, however, were an important part of fact-checking: "We [in the Nordics] are very much in favor of anonymizing people [in pictures]. [It is about] personal integrity and then it applies even to dead people, [especially dead] children. […] On the other hand, I think fact-checking [is] important, [so then I will] include the link" (Larsson, 05.01.24).

In an article summarizing the first month of the information war between Israel and Hamas, TjekDet also included the following warning: "This article contains graphic descriptions as well as links to videos and images of acts of violence that some readers may find disturbing" (Damgaard Frisch et al., 2023). Søndergaard argued that they had the responsibility AT TJEKDET to document their fact-checks to the European Fact-Checking Standards Network (EFCSN), but also to warn their readers about graphic pictures and videos:

> The reason we did it in this article […] is because we have [an] obligation to EFCSN […] to document our claims. […] In order to document these things, we must link to pictures and videos of some really macabre things [such as] child corpses, necrophilia and violence. We want to make the readers aware of [it and] we also think it is a good idea to [warn] at the top of the article and say from the start: Be aware [that] this article [contains links to] violent content (Søndergaard Petersen, 07.12.23).

We also have the response from the audience after the fact-checks have been published. All of the three fact-checkers addressed the importance of feedback from their readers and a willingness to go into dialogue even with critical voices. Here it is crucial to strike a balance between openness and dialogue with the public, on the one hand, and not let the public relations and education steal too much time and resources on the other hand.

I Hear What You Say – Public Dialogue

A fourth strategy of building trust through fact-checking is to keep an open dialogue with the public. A good relationship between the fact-checker and their readers is important, both for providing tips for stories to fact-check, but also to correct or get feedback on the published articles. Østrem in Faktisk noticed how they most often were contacted by their most critical readers, usually with special interest in a topic going in one or another direction: "[They are] critical of us - that we do too little or that we make mistakes. [...] Many have very specific tips and offer to write the articles for us" (Østrem, 18.12.23).

The fact-checkers at Faktisk try to go into dialogue with the most critical voices, but also realize that the conversation is time consuming, taking away the focus from their main mission of fact-checking: "However, I am grateful to see that many participants in the comment section take on that job. [They] try to moderate and say: Yes, but read the [whole] article now" (Østrem, 18.12.23). Looking at the responses from the public, Østrem also admits that it creates fact-checkers with a tough skin:

> We receive some direct [criticism too] when you write about a story where someone perceives us as [...] very biased or in opposition to their world view. People [in the newsroom] have become a bit, regrettably I must say, tough. They tolerate [much], but it is not how it's supposed to be. We like to discuss things, but often it gets out of hand. [...] Even if we do a good job and we feel that this was a good article, we are still not sure about the reaction it will receive (Østrem, 18.12.23).

Larsson from Källkritikbyrån noticed the difference between their day-to-day followers, who appreciated their work, and those who were critical of a specific fact-check, where the latter more often made contact to voice their criticism. However, dialogue and respect go a long way:

> When we do [fact-checks], we have a safe base of people who like us [...] and appreciate such reviews. [But] it's not like they especially seek us out to say it was a good article. Those who contact us and have something to say, [they] are angry. [...] They think we have an opinion on something or that we have published an opinion piece, when we have [only] examined a factual matter as part of an opinion piece. [...] Then we need to explain our role, [and] when you take them seriously and talk to them, then [they] calm down and sometimes they even say they are sorry (Larsson, 05.01.24).

Søndergaard Petersen from TjekDet also received a lot of feedback from the public, usually in the form of criticism. Sometimes this led to editorial changes to their articles, like the previously mentioned "From the river to the sea – Palestine will be free" (Engelbrecht & Brandt, 2023) expression. He did, however, question the value of fact-checking the Israel-Hamas conflict, due to the deep emotional divide and differences of opinion:

> In general, I think [fact-checking] is important because it can help people understand the world they live in, and it is particularly important when they give their vote in various [...] elections. [But] I don't think it is particularly important with Israel-Hamas. I actually think that the fact-checks [in an] inflamed conflict like this

> have less effect, because people are so locked in their attitudes and opinions. Even if everything points to the claims they believe in being wrong, they cling to them anyway (Søndergaard Petersen, 07.12.23).

Søndergaard Petersen argued that instead to trying to convince people with a set opinion, the real value of TjekDet's fact-checks in deadlocked conflicts like this was in monitoring and correcting editorial media coverage: "Fact-checking can have value during war and conflict in relation to the established media, because they [can] run off with some half-truths and undocumented claims. I think we can achieve something [here] because we are more thorough than they are. We have made them change headlines or make corrections" (Søndergaard Petersen, 07.12.23).

A Network of Fact-checkers

The fifth and final strategy of building trust is to get tips and seek help from the network of other fact-checking organizations. When faced with the challenge of fact-checking a war and conflict far away from home with little or no access to the area and a multitude of pictures and videos going viral, Faktisk, Källkritikbyrån, and TjekDet rely on the network of other fact-checkers – in the Nordic countries, Europe, and internationally. Østrem in Faktisk stressed the value of the network for tips and internal discussion:

> We have benefited from being part of both [the] European fact-checking network, EFCSN, and the international [fact-checking network]. There are often very good [discussion] threads [to follow] there, because many people work on the same issues. [It's] really helpful to see what other [fact-checkers] are working on and get tips. [Last], we have a Nordic collaboration with academics and fact-checkers in the Nordic countries, [where] we inform each other [about] what we are working on (Østrem, 18.12.23).

According to Søndergaard Petersen at TjekDet, they cooperated with all EFCSN fact-checkers, as well as "Faktisk, Källkritikbyrån, and Faktabaari in Finland [on certain stories]" (Søndergaard Petersen, 07.12.23), but no cooperation with media and editorial offices in Denmark. He saw the opportunities for a database with verified visual content, like the one offered by Faktisk Verifiserbar in Norway, but argued that TjekDet would rather prioritize their main mission of publishing fact-checks.

Larsson from Källkritikbyrån mentioned the Nordic fact-checking network NORDIS and her appreciation of having "colleagues" to discuss the fact-checking process:

> We collaborate by taking part in the same [network]: NORDIS. It has been very rewarding for me [since I am] working for myself. [It] feels like you have colleagues, you have the same problem, [and] you can brainstorm with them. [However] we do

> not cooperate on fact-checking specific stories together, other than the large project we did on the child welfare rumours in various countries (Larsson, 05.01.24).

The child welfare rumors, mentioned in the above quote, was a joint project with Källkritikbyrån, Faktisk, TjekDet, and Faktabaari (Faktabaari, 2024) looking into viral rumors about the child protective services in the four Nordic countries (Larsson, 2023b). An interesting part of the project was to show how the rumors had "migrated from Denmark and Sweden [to Finland and Norway]" (Larsson, 05.01.24). Many of the stories claimed that children were kidnapped by the social services to sell them for profit.

Sometimes the rumors about the child protective services in Sweden were linked to the Israel-Hamas conflict. In an article (Larsson, 2023d), fact-checking an incident during a pro-Palestine demonstration in Stockholm, Larsson found that one of the speakers from the demonstration had strong ties to online communities that are against the child protective services:

> It started with [this story] going viral, that [a random person] grabs the microphone [during a pro-Palestinian demonstration] and makes his own little speech. […] He says [that] in Ulf Kristersson's Sweden, 1,000 children are being kidnapped [by the child protective services]. I have [however] been involved in these anti-child welfare services groups on Facebook for quite some time [and there] I got a completely different picture. So, this is a kind of fact-check that I would call: "There is more to the story" (Larsson, 05.01.24).

According to Larsson, the joint work of fact-checkers in Sweden, Denmark, Norway, and Finland on the child welfare rumors allowed the fact-checkers to get a clearer picture of the extent of the phenomenon, usually found in closed online forums: "These groups that we are talking about now, people with an immigrant background, they talk and live in various online social forums. […] We're not good at listening to [such communities]. What are they talking about and what misinformation is being shared [between them]" (Larsson, 05.01.24). She clearly saw the value of network and cooperation across borders to get a "helicopter perspective" (Larsson, 05.01.24) on the stories they fact-check.

Conclusion

This chapter has addressed the challenge of handling fake news and disinformation related to an escalating conflict and war far away from home. By combining research interviews with representatives for the three Scandinavian fact-checkers Faktisk, Källkritikbyrån, and TjekDet with relevant examples of fact-checks from the first two months of the Israel-Hamas conflict, this study has aimed to answer the following research question:

How did Faktisk, Källkritikbyrån, and TjekDet seek to build trust through fact-checking in the two first months of the Israel-Hamas conflict?

In the analysis I have chosen to focus on five different strategies that the fact-checkers use to build trust and credibility in the fact-checking process. They are honest about uncertainty (1), present longer explanations and conclusions rather than color categories (2), provide visual proof, but also protect the audience (3), encourage and take part in public dialogue (4) and rely on Nordic, European, and international networks of fact-checkers (5).

When fact-checking war and conflict in a faraway area with limited time and access, fact-checking organizations like Faktisk, Källkritikbyrån, and TjekDet are often faced with uncertainty and difficulty in giving a clear answer to the validity of a claim, a picture, or a video. They all stress the importance of being honest about this uncertainty, from posting "This is what we know" articles to editing previous fact-checks or cancelling publication all together.

While the fact-checkers from the launch often used color categories to mark their conclusions, now they write more detailed conclusions, explaining the fact-checking process and trying to educate their audience. While some of the fact-checkers admit that there is a tabloid and visual appeal of colors and clear-cut conclusions, they still argue that this type of fact-checking takes away the focus from what they want to present. However, in cooperation with Meta's *Third-party fact-checking initiative* all the fact-checkers are faced with a demand for categorizing their results.

Another aspect of fact-checking is providing proof, often visual proof, from the war scene. Documenting their claims is an important part for the fact-checkers to build trust. On the other hand, the visuals are often very graphic and can be considered disturbing for the audience. The fact-checkers do sometimes include links to violent videos and images, but make sure to warn their audience before they follow these links. Faktisk, Källkritikbyrån, and TjekDet all need to find a balance between documentation and being a responsible publisher.

All of the three fact-checking organizations have experienced varied responses from the audience after a fact-check had been published. These responses are both positive and negative, but the readers seeking them out to voice their opinion are often critical of their coverage. Their criticism might be based on a misunderstanding of the fact-checkers' role and the process of fact-checking. However, all the fact-checkers emphasize the need for respect and a two-way conversation, even if a public dialogue was time consuming and they sometimes had to face both unreasonable critique and threats.

Last, all the fact-checkers valued the network of other fact-checking organizations. They use the networks to get tips on stories or viral claims to fact-check, and to discuss the cases they are working on. Through the network, the fact-checkers also become more aware of the viral spread of misinformation or unverified pictures and videos across national borders. This might be especially valuable in situations with a faraway conflict and war, with little or no access to the area, like we see in the Israel-Hamas conflict.

References

AFP. (2023, October 18). *No evidence that a video features a doll rather than a dead child in Gaza*. https://factcheck.afp.com/doc.afp.com.33YE6K9

Allern, S. (2019). Journalistikk som "faktasjekking". In P. Bjerke, B. K. Fonn, & B. R. Mathisen (Eds.), *Journalistikk, profesjon og endring* (pp. 167–187). Orkana akademisk.

Allern, S., & Pollack, E. (2019). *Källkritik!: journalistik i lögnens tid*. Studentlitteratur.

APNews. (2023). *Gaza receives largest aid shipment so far as deaths top 8,000 and Israel widens military offensive*. AP News. Retrieved October 30 from https://apnews.com/article/israel-hamas-war-news-10-29-2023-de1a7d660ba2f6d80b3d7aeaae5bb0f3

BBC. (2023, November 2). *What is happening in Israel and Gaza, and what is Hamas?* Retrieved November 2 from https://www.bbc.com/news/world-middle-east-67039975

Bengtsson, M., Hansen, A. S., Hartley, J. M., Kristensen, J. B., Mayerhöffer, E., & Ramsland, T. G. B. (2021). *Conspiracy theories during Covid 19: The case of Denmark*. Roskilde Universitet. *https://doi.org/https://doi.org/10.5281/ZENODO.5556012*

Boyd-Barrett, O. (2004). Judith Miller, The New York Times, and the propaganda model. *Journalism studies (London, England), 5(*4), 435-449. https://doi.org/10.1080/14616700412331296383

Damgaard Frisch, N., Fischer Brandt, T., Askjær, R. A. O., Gregersen, K. T., & Nguyen, V. (2023, November 15). *Pallywood, video-jihad og reklamer til børn: Her er første måned i informationskrigen mellem Israel og Hamas*. TjekDet. https://www.tjekdet.dk/indsigt/her-er-foerste-maaned-i-informationskrigen-mellem-israel-og-hamas

Egeberg, K. (2021, June 17). *Velkommen til nye Faktisk.no!* https://www.faktisk.no/artikler/0orn3/velkommen-til-nye-faktiskno

Ellul, J. (1973). *Propaganda*: *the formation of men's attitudes* (Vintage Books edition. ed.). Vintage Books.

Engelbrecht, S., & Brandt, T. F. (2023, November 7). *"From the river to the sea, Palestine will be free" råber demonstranter. Men er det antisemitisk?* TjekDet. https://www.tjekdet.dk/indsigt/river-sea-palestine-will-be-free-raaber-demonstranter-men-er-det-antisemitisk

Faktabaari. (2024). *Welcome to Faktabaari/Factbar*. https://faktabaari.fi/in-english/

Faktisk. (2024). *Faktisk*. https://www.faktisk.no/

Faktisk. (n.d.-a). *Endringer*. https://www.faktisk.no/endringer

Faktisk. (n.d.-b). *Om oss*. https://www.faktisk.no/om-oss

Graves, L. (2016). *Deciding what's true: the rise of political fact-checking in American journalism*. Columbia University Press.

Greger, M. (2023, November 14). *Slik jobber de med Gaza-videoer: – Vi teller kroppsdeler*. https://www.journalisten.no/slik-jobber-de-med-gaza-videoer-vi-teller-kroppsdeler/593619

Gregersen, K. T., & Engelbrecht, S. (2023, October 13). *Har Hamas halshugget 40 babyer i israelsk kibbutz?* TjekDet. https://www.tjekdet.dk/faktatjek/har-hamas-halshugget-40-babyer-i-israelsk-kibbutz

Hiebert, R. E. (2003). Public relations and propaganda in framing the Iraq war: a preliminary review. *Public Relations Review, 29*(3), 243-255. https://doi.org/10.1016/S0363-8111(03)00047-X

Jowett, G., & O'Donnell, V. (2012). *Propaganda & persuasion* (5th edition. ed.). SAGE.

Kalsnes, B. (2019). *Falske nyheter: løgn, desinformasjon og propaganda i den digitale offentligheten*. Cappelen Damm akademisk.

Khaldarova, I., & Pantti, M. (2016). Fake News: The narrative battle over the Ukrainian conflict. *Journalism practice, 10*(7), 891-901. https://doi.org/10.1080/17512786.2016.1163237

Källkritikbyrån. (2024). *Källkritikbyrån.* https://kallkritikbyran.se/

Källkritikbyrån. (n.d.-a). *Om oss.* https://kallkritikbyran.se/om-oss/

Källkritikbyrån. (n.d.-b). *Vår journalistik.* https://kallkritikbyran.se/var-journalistik/

Larsen, E. S. (2019). The Limits of Truth. A Case Study of Faktisk and CrossCheck. In F. Orban & E. S. Larsen (Eds.), *Discussing Borders, Escaping Traps. Transdisciplinary and Transspatial Approaches* (pp. 47–66). Waxmann. https://www.waxmann.com/buch4045

Larsson, Å. (05.01.24). *Research interview with Åsa Larsson* [Interview]. Källkritikbyrån;

Larsson, Å. (2023a, October 31). *"Liket lever" – tre fall som påstås vara fejkande offer.* Källkritikbyrån. https://kallkritikbyran.se/liket-lever-tre-fall-som-pastas-vara-fejkande-offer/

Larsson, Å. (2023b, March 6). *"Staten kidnappar barn" – så skiljer sig ryktena åt i de nordiska länderna.* Källkritikbyrån. https://kallkritikbyran.se/staten-kidnappar-barn-sa-skiljer-sig-ryktena-at-i-de-nordiska-landerna/?fbclid=IwAR3utvlu5oIiLre7HrEi-h9yOKiT-zjWudUgEMd4bgkp00z9oNmokM8PWXM

Larsson, Å. (2023c, October 11). *Så sprids obekräftade uppgifter från kriget i Israel.* Källkritikbyrån. https://kallkritikbyran.se/sa-sprids-obekraftade-uppgifter-fran-kriget-i-israel/

Larsson, Å. (2023d, October 25). *Talaren från Family united presenterades av LVU-kritisk arrangör.* https://kallkritikbyran.se/talaren-fran-family-united-presenterades- av-lvu-kritisk-arrangor/

Lønrusten, K. G., Akerbæk, E., & Lindin, I.-K. (2023, October 11). *Nyheten om 40 døde babyer og barn er ikke bekreftet.* https://www.faktisk.no/artikler/jx872/nyheten-om-40-dode-babyer-og-barn-er-ikke-bekreftet

Magallón-Rosa, R., Fernández-Castrillo, C., & Garriga, M. (2023). Fact-checking in war: Types of hoaxes and trends from a year of disinformation in the Russo-Ukrainian war. *El profesional de la informacion, 32*(5). https://doi.org/10.3145/epi.2023.sep.20

McGonagle, T. (2017). "Fake News": False fears or real concerns? *Netherlands Quarterly of Human Rights, 35*(4), 203-209. https://doi.org/10.1177/0924051917738685

Nimmo, B. (2015). *Anatomy of an Info-War: How Russia's Propaganda Machine Works, and How to Counter It*. StopFake. Retrieved 19.05.15 from https://www.stopfake.org/en/anatomy-of-an-info-war-how-russia-s-propaganda-machine-works-and-how-to-counter-it/

NorwegianGovernment. (2022, July 6). *Israel – Palestine: A negotiated two-state solution.* https://www.regjeringen.no/en/topics/foreign-affairs/peace-and-reconciliation-efforts/norways_engagement/israel_palestine/id2522237/

NorwegianGovernment. (2024, May 22). *Norway recognises Palestine as a state.* https://www.regjeringen.no/en/aktuelt/norway-recognises-palestine-as-a-state/id3040194/

Picheta, R., & Salman, A. (2024, February 29). *More than 30,000 killed in Gaza since Israel-Hamas war began, health ministry says.* Retrieved Feburary 29 from https://

edition.cnn.com/2024/02/29/middleeast/gaza-death-toll-30000-israel-war-hnk-intl/index.html

Ponsonby, A. P. (1928). *Falsehood in war-time: containing an assortment of lies circulated throughout the nations during the great war.* Allen & Unwin.

Sikt.no. (2023). *Ref.nr. 379715 Truth – the first casuality. Fake News and Fact-Checking in the Israel – Palestine conflict.* https://minforskning.sikt.no/prosjekt/65c3149d-7db5-466e-afbf-0799ebe2877e

Søndergaard Petersen, A. (07.12.23). *Research interview with Andreas Sønderhaard Petersen* [Interview]. TjekDet.

Tambini, D. (2017). *Fake News: Public Policy Responses. https://core.ac.uk/download/pdf/80787497.pdf?fbclid=IwAR2plm*I40dVowxOs

TjekDet. (2024). *Tjek det.* https://www.tjekdet.dk/

TjekDet. (n.d.-a). *Fejl og fakta.* https://www.tjekdet.dk/fejl-og-fakta

TjekDet. (n.d.-b). *Om os.* https://www.tjekdet.dk/om-os

Verifiserbar. (2023). *Om oss.* https://www.verifiserbar.no/

Østrem, O. (18.12.23). *Research interview with Olav Østrem* [Interview]. Faktisk, Pressens Hus.

Media Trust in the Digital Era: Exploring Audience Trust Perception Among Fact-checkers in Mali

Samba Dialimpa Badji and Nelson Bahati

Introduction

"Whether it is true or not, we wish for the fatal fall of this currency. Besides, what circulates in TikTok is true in our opinion".

The comment above is from one Facebook user on a fact-checking article debunking a video shared on TikTok and stating the end of the Euro currency in Europe. Beyond the debate on the effectiveness of fact-checking, this reaction raises the question of audience trust in fact-checkers, and even more, that of knowing what it does to a fact-checker to feel that his/her readers, or some of them, do not trust him while he is supposed to help them access correct information. Thus, this chapter aims to understand what audience trust means for Malian fact-checkers by addressing the following question: how do Malian fact-checkers perceive audience trust and what strategies do they put in place to build audience trust?

Trust is a fundamental element in social interactions (Hancock et al., 2023). It plays a key role in many economic and social outcomes (Algan & Chahuc, 2014). Even if there is no commonly accepted definition of trust, Kramer (1999) noted that most theorists have conceptualized trust as a "psychological state" or a "choice behavior". As a psychological state, he said, "trust entails a state of perceived vulnerability or risk that is derived from individuals' uncertainty regarding the motives, intentions, and prospective actions of others on whom they depend" (p. 571). As a choice behavior, he added, trust involves risky choices where individuals make rational choices to "maximize expected gains or minimize expected losses from their transactions" (p. 572). In this sense, trust is a voluntary expectation on the part of one party that the other will perform a particular action that will be beneficial (Coleman, 1974). Thus, trust involves taking risks with the expectation that all parties involved will act competently and dutifully (Lewis & Weigert, 1985). Based on the previous definitions, and following Hanitzsch, van Dalen, and Steindl (2018) regarding the media, we refer to audience trust as the willingness of the audience to be vulnerable to fact-checkers based on the expectation that they will perform adequately. In journalism, the importance of trust is reflected in the existence of projects and initiatives intended to strengthen trust in the media such as The Trust Project, the Journalism Trust Initiative from Reporters Without Borders, or the Trust, Media and Democracy Initiative of the Knight Foundation. Regarding fact-checking, trust is embedded in the International Fact-Checking Network's Codes of Principles which advocates for non-partisanship, transparency of sources, transparency of funding and organization, transparency of methodology, and an open and honest corrections policy (IFCN, n.d.).

Without trust from the audience, the media will not have the legitimacy to fulfill their watchdog function (van Dalen, 2020). The concept of media trust is linked by some scholars to perceptions of journalistic quality (Prochazka & Schweiger, 2018), others link it to the concept of media credibility. Media credibility refers to the quality of a piece of information or a media being believable or worthy of trust (Bucy, D'Angelo, & Bauer, 2014). In this sense, credibility is related to a specific evaluation of media content and refers to the perception of the accuracy of the information at a given point in time (Fawzi et al., 2021). Thus, the perception of credibility is in some aspect a factor of the audience's trust in fact-checking work (Liu, Qi, Wang, & Metzger, 2023). For this study, we will use interchangeably media trust and audience trust to refer to the trust of the audience toward fact-checkers and fact-checking organizations.

Trust in the media has constantly fallen in most countries in recent years (Fletcher, 2020). A trend confirmed by the *Digital News Report 2023* of *Reuters Institute* (Newman et al., 2023) and the 2024 Edelman Trust Barometer (Edelman, 2024). However, trust in media overall remains relatively high in Africa (Newman et al., 2023; Edelman, 2024). Several reasons explain the decline of audience trust in the media. Some research highlights the role of the Internet which has allowed the public to have alternative sources of information (Tsfati, 2010); thus other research found that audience distrust is linked to the perception of bias – political and commercial – and poor journalism (Newman & Fletcher, 2017).

Journalism and Fact-Checking in the Context of Conflict in Mali

The ongoing crisis in Mali has a detrimental impact on the safety and security of media professionals (MFWA, 2021). Several media have been suspended and journalists are subject to attacks and arbitrary arrests (Bocande, Lagarde, & Marong, 2023). In 2022, Malick Konate, a journalist, blogger, and fact-checker, fled the country and went into exile in Senegal, then in France, following threats received after contributing to a French television documentary on the presence in Mali of the Russian private military company Wagner (L'Oeil de la Maison des Journalistes, 2023). The deterioration of the media situation in Mali in recent years is reflected in the Reporters Without Borders world ranking on the state of press freedom, in which Mali fell from the 25th place in 2011/2012 to the 113th place in 2023.[1]

Mali, the second largest country in West Africa with 1,241,238 million square kilometers, has been the scene of a multidimensional crisis since 2012 (Coulibaly & Lima, 2013). The crisis started with a Tuareg rebellion demanding independence for northern Mali before evolving into a jihadist insurgency, to which inter-community violence has been added (Pellerin, 2019).

1 Based on an analysis of Reporters Without Borders reports from 2011/2012 to 2023.

Before the outbreak of the crisis, Mali was among the African countries with a free and dynamic press. The Malian press notably supported the democratic demands that led to the fall of the dictatorship of General Moussa Traore – who ruled the country from 1968 to 1991 – and the establishment of democracy in the country in the early 1990s (Perret, 2005). For three decades, the Malian media landscape had been characterized by its diversity and independence, before facing increasingly severe restrictions in recent years, especially after the coup d'etat that occurred in 2021 (MFWA, 2022). A restrictive stance that seems to have the support of an important segment of the population as 53% of Malians are in favor of the government controlling what the media publishes (Afrobarometer, 2024). Furthermore, the Malian information ecosystem has been also impacted by the changes that have disrupted the media industry globally in the digital era with the emergence of a new form of journalism characterized by the speed in production and dissemination, great interactivity, and customized sharing (Mutsvairo, 2016). Therefore, the media are forced to adapt (Mabweazara, 2014). The main challenge is to preserve journalistic standards while adopting and mastering new tools and codes that come with technology (Aboutayeb, 2022). While introducing major changes in journalism, technology has democratized the production and dissemination of information. Through blogs and social media platforms, citizens are able to play a prominent role in the public debate (Castells, 2007). These spaces of free expression favored the emergence of a dynamic blogosphere in Mali, with bloggers producing content that borrows from journalistic discourse, contributing to democratizing information, expanding the public space as well as enriching the public debate (Sissoko & Dembele, 2023). These new actors have gained influence to the point of being given the same importance or even replacing traditional media (Togola & de Bruijn, 2023).

However, the other side of the coin is the emergence of other influential online actors who use social media platforms to disseminate false information and hate speech (Ouedraogo, 2022). In Mali, they are called 'videomen'. They are influencers using social media platforms such as Facebook, YouTube, Instagram, and TikTok to broadcast live videos with content varying between authenticity, propaganda, manipulation, and disinformation (Sissoko et al., 2024). While some of these influencers, particularly those who support the government, increasingly play a prominent place in the media landscape, critical voices, including media professionals, are muzzled (Lorgerie, 2022). In this difficult context for journalists, seven organizations are involved in fact-checking as a strategy to combat disinformation and its effects. These organizations are Association des Blogueurs du Mali, Benbere Verif, Mali Check, Studio Tamani, Wuya, Mopti Check, and Sahel Check. Fact-checking involves verifying all information constructed inappropriately and falsely shared publicly (Vizoso & Vazquez Herrero, 2019). It is also seen as a reform movement in journalism aiming to restore trust in the media (Kyriakidou et al., 2023). However, organizations doing fact-checking in Mali are often accused of bias and of being pawns of Western powers, France in particular, from which they receive funding (Laplace, 2022). In this context, it is therefore worth understanding what perception these fact-checkers have of audience trust.

Literature Review

Since the early 2010s there has been increased research on the relationship between the media and the audience (Uth, Stehle, Wilhelm, Detel, & Podschuweit, 2023). A significant part of the literature is dedicated to audience trust in a context marked by concerns regarding the decreasing level of trust in the media, the emergence of alternative sources of information competing with legacy media, and the proliferation of disinformation (Jakobsson & Stiernstedt, 2023). The academic interest in audience trust is mostly oriented toward trust in the media as an institution (Kohring & Matthes, 2007; Otto & Köhler, 2018), or in trust in specific media types (Kiousis, 2001). Tsfati et al. (2022) went beyond general media trust to suggest that media trust may differ depending on the topic of coverage and that topical media trust can be distinguished from general media trust. Moehler & Singh (2011) were interested in understanding why in post-authoritarian African democracies citizens trust government-owned broadcast media more than private broadcasters. They explain that it is related to "low political sophistication, undemocratic and uncritical attitudes toward political and pro-government partisan orientations" (p. 277).

Comparing mainstream media and new media in Ghana and Nigeria, Mustapha et al. (2022) argue that mainstream media is still highly trusted regarding political information, even if social media is used as an alternative to getting political information. Wasserman & Madrid-Morales (2019) found that lower levels of media trust in South Africa are linked to higher levels of perceived exposure to disinformation. This finding is corroborated by that of Valenzuela, Halpern, and Araneda (2022) regarding misinformation and media trust in Chile. Regarding fact-checking, research that addresses audience trust is more focused on the effectiveness of this journalistic practice (Primig, 2022; Liu, Qi, Wang, & Metzger, 2023; van Erkel et al., 2024).

As this review shows, most of the research on media trust has an audience perspective focus. A few of them investigated journalists' trust in the media (Steindl et al., 2023), and fewer have studied journalists' perception of audience trust (Tsfati, 2004). Investigating Israelis journalists' perception of audience trust, Tsfati (2004) argues that there is a positive correlation between perceived audience trust and identification with professional standards such as neutrality, verification, and factualness. To our knowledge, this is the first study on the perception of audience trust among fact-checkers, which makes this exploratory study relevant. This study contributes to advancing the research on media trust, by shedding light on Malian fact-checkers' perception of audience trust and their strategies to build and maintain trust.

The exploration of Malian fact-checkers' perception of audience trust and their trust-building strategies is done through the lens of imagined audience and engaged journalism theories.

Imagined audience

To communicate efficiently, each person needs to know his audience. Audience refers here to the people for whom the message is intended (Sullivan, 2020). The way one perceives the composition of his audience determines somehow his way of communicating and interacting with them (Coddington, Lewis, & Belair-Gagnon, 2021). Faced with the impossibility of knowing exactly the composition of the audience, people tend to conceptualize it mentally (Litt, 2012). Following Litt (2012) and Marwick & Boyd (2011), Litt and Hargittai (2016) found that people's imagined audiences on social media are not stable. The two authors categorized two types of social media users' imagined audiences: an "abstract" imagined audience that was "vague and general", and a "target" imagined audience that was "more specific and directed" (Litt & Hargittai, 2016). While Napoli (2011) argued that the digital era enables a greater journalism audience understanding, Coddington, Lewis, & Belair-Gagnon, (2021) found that journalist's perception of their audience comes from a complex variety of sources which include interactions via email, social media, and comment sections.

Understanding what the fact-checking audience looks like from the fact-checkers' perspective and where fact-checkers' perception of their audience comes from could highlight how fact-checkers engage with their audience in order to build and maintain trust.

Engaged Journalism

In their definition of engaged journalism, Green-Barber and McKinley (2019) listed four characteristics: inclusiveness, prioritizing the audience's expectations in terms of information, creating space for the audience to collaborate in all aspects of the journalistic process, and the effort to establish and preserve relationships of trust between journalists and the audience. This means that the media must not only pay attention to the needs of their audiences but also have permanent contact with them to earn and maintain their trust and loyalty (Nelson, 2018), with the ultimate goal of enhancing democracy (Rosenberry & St. John, 2010). In this sense, engaged journalism is a new form of public journalism with enhanced interaction capabilities enabled by technology (Schmidt, Nelson, & Lawrence, 2022). The idea behind public journalism, a concept that appeared in the early 1990s (Rosen, 1999; Harcup, 2014; Min, 2016) is to make audience engagement part of the journalistic process (St. John & Johnson, 2021). The digital era enabled what Rosenberry & St John (2010) call "public journalism 2.0" in which improved technical capacities reinforce the way journalists and citizens are partners in driving the news agenda (Rosenberry & St. John, 2010). Another form of engaged journalism is what Nip (2010) calls guided professional reporting. This mode of citizen connection in newsmaking allows the audience's involvement in several stages of the news process while the journalists keep the responsibility of producing the work (Nip, 2010, p. 138). In this sense, engaged journalism is a relevant theory of analyzing and understanding how Malian fact-checkers attempt to build and maintain audience trust through the interactions they have with their readers and the community in general.

Methods

Data have been collected through semi-structured in-depth interviews of five editors from five of the seven fact-checking organizations existing in Mali; for security reasons, we have decided not to disclose the names of these organizations as we believe that could make it possible to identify the respondents. For instance, during two of the interviews, respondents asked about the security of the Teams platform with concerns that communication could be intercepted by country authorities. As data was collected as part of his Ph.D. project research, one of the authors registered an academic research application. In addition, respondents were informed about the purpose of the interviews and gave their consent. We also guarantee confidentiality and anonymity. Even if a sample of five respondents is small, our sample selection was not aimed at the representativeness of fact-checkers but at understanding the lived experience of those interviewed (Van Manen, 1997) and how they perceive, describe, and make sense of audience trust (Patton, 2002). In addition, as those who write fact-checking articles work on a freelance basis, we found it relevant to focus on editors who are in charge of safeguarding editorial standards. Due to geographical constraints, all five interviews were conducted online between March 7 and March 25 2024 using the Teams platform and lasted between one hour and one hour and a half. Remote data collection with online interviews has become a common practice since the Covid-19 pandemic and is a practical alternative when the researcher's presence or travel is difficult (Keen et al., 2022). However online interviews can be challenging in some places due to Internet problems (Akyirem et al., 2024). The interviews were conducted in French. After transcription, they were anonymized and translated into English for the data analysis with the co-author who does not read nor speak French.

Following the thematic analysis, the interview transcripts were manually coded to extract meanings from the data (Braun & Clarke, 2006). In this chapter, respondents are identified as follows: Respondent 1 (R1), Respondent 2 (R2), Respondent 3 (R3), Respondent 4 (R4) and Respondent 5 (R5).

Findings

Three themes have emerged from the data analysis. The first theme is related to the significance of audience trust for fact-checkers. The second theme is about fact-checkers' views of their audience. The third and last theme is related to the strategies to build audience trust.

Definitions, Meanings, and Importance of Audience Trust

There is a lack of consensus among fact-checkers regarding the definition of the concept of "trust". The respondents frequently employed a diverse range of terms, including evidence, truth, confidence, belief, challenge, and feedback to define the constituent components of trust. For example, Respondent 1 (R1) posited that audience confidence in engaging with fact-checkers is tantamount to trust, while Respondent 2 (R2) argued that if audiences question fact-checkers, such as by seeking specific information and contacting them for verification, it indicates trust in their work. The findings reveal a consistent pattern of value among the respondents on the notion of the "meaning of audience trust" within the framework of fact-checking. The analysis of the five respondents reveals that audience trust is more than just a component of their profession; it encompasses all aspects of their fact-checking pursuits.

For instance, the respondents tended to believe that, without trust, their work and efforts would be ineffective, signaling a bleak outlook, as seen in the interview excerpts below:

> *"This is the very essence of our work. If the public does not trust it, it means that everything we do is doomed to failure. The trust of our readers is what pushes us to move forward and persevere… Without this trust, I believe we are working in nothingness"* (R5, 25th March 2024).

> *"It means a lot to me because we are working to clean up the public debate. If we do not have the trust of the people who read us, it can be discouraging, and it can give the impression that the work we do is not of much use"* (R3, March 19th 2024).

The findings reveal that the nature of their fact-checking activity influenced the respondents' perception of audience trust. They held the belief that their work revolved around establishing the trust of the public, and the absence of such trust had an impact on the implementation of fact-checking.

> *"Audience trust is important because we work, as journalists/fact-checkers we work on subjects of public interest, it is for the public that we work, so if this public does not trust us our work has no meaning"* (R4, March 20th 2024).

Fact-checkers' Views of the Audience

Respondents identify two distinct categories of audiences. One segment of the audience exhibits a lack of appreciation for alternative ideologies and instead maintains their own interpretation of the information, displaying indifference towards the veracity of the information they encounter. Conversely, another segment of the audience displays curiosity and engages in independent reading to ascertain the accuracy of the information they receive.

The respondents believe that doing fact-checking for a segment of the audience that prioritizes their existing thoughts or ideological affiliations over factual accuracy undermines the significance of fact-checking efforts (R5, March 25th 2024).

Furthermore, the respondents consider that the majority of the audience believes in all information or content they receive without asking questions or putting it into perspective. They explain this attitude by the fact that the audience is not aware of the existence of disinformation or its dangers:

> *“Very few people know, for example, that you can edit a photo, that you can take a video out of the context it was taken, that you can transform a voice and give the voice of another person”* (R1, March 7th 2024).

In addition, respondents see a part of the audience as fickle and only believing any information that supports their beliefs, regardless of whether it is correct or not. One respondent likens this attitude to a strategy of self-protection in a context marked by restriction of freedoms and repression of dissent voices. Thus, systematically believing any information that is in line with the narratives of the military government is a kind of survival instinct:

> *"Like a fox, a cunning animal, that knows when to hunt, when not to hunt, and which prey to hunt, this audience believes or does not believe in information depending on the orientations of those in power"* (R3, March 19th 2024).

Trust-building Strategies

Strategies for building audience trust are strongly linked to respect for journalistic ethics standards, strict compliance with fact-checking procedures, independence, and non-partisanship. As one respondent points out:

> *“Everything we do, we strive to do it professionally. That is to say, if we have a topic, we take the necessary time, we try to find the necessary sources, the testimonies that can adequately support our fact-checking article so that at the end of our work we can present to the audience a work which is free from reproaches, a work that is much more professional"* (R1, March 7th 2024).

Transparency is also considered a major factor in building and maintaining audience trust. This includes transparency regarding funding sources. To this end, fact-checking organizations publish the list of their donors on their website:

> *“We explain to the public how we are financed. We say who finances us, and how we finance our activities. Because we think it’s essential so that people don’t think that we work on behalf of this or that organization or that we are at the service of this or that entity"* (R5, March 25th 2024).

Interactions with the audience are another way for fact-checkers to build and maintain audience trust. These interactions allow fact-checkers to collect the audience's concerns and see to what extent they can address them. In addition, the public is encouraged to contact fact-checking organizations to submit content for verification. For instance, as a means of facilitating their interactions with their audiences, fact-checkers established WhatsApp groups, both at the regional and national levels, where audiences, including journalists, interact with fact-checkers on various concerns.

As a strategy to build and maintain trust with their audience, fact-checkers revealed that they explain their work to their audiences, demonstrating how they establish falsehood in journalistic news articles. For instance, they engage their audiences during training sessions about how they arrive at facts in their articles. In these training sessions, fact-checkers present audiences with fact-checking articles, allowing them to read and understand the fact-checker's approaches to the content.

> *"We just explain what we do. And every time we have the opportunity, we even teach people to read our fact-checking articles. In my sessions, I always take two hours to show an article and get people to read and understand what we are putting in it"* (R1, March 7th 2024).

Discussion

Drawing on data from five (5) respondents, our findings revealed that fact-checkers in Mali perceived audience trust as entwined with the process of fact-checking, and the nature of the fact-checking process highly influenced the fact-checkers' perceptions of their work as relying widely on audience trust. We established that it is challenging to create trust in a politically polarized audience, mostly during periods of conflict and political events such as elections, when the audiences are highly divided in nature.

The notion of "trust" was found to be central among the fact-checkers, despite the fact that we observed a lack of consensus about the definition of the concept of "trust" among all our respondents. We found that fact-checkers believed audience trust incorporates in the activities of their work, and the notion of trust relates to audiences believing and engaging in the fact-checkers craft, such as audiences having confidence in the fact-checkers work, the audiences' ability to challenge the fact-checkers' work, the audiences' feedback, and using the fact-checkers' work as evidence of truth to verify disinformation content. This kind of perceived engagement underscores what Kim et al. (2022) considered an important aspect of the fact-checking process, as audience engagement widely signifies and substantiates the existence and usage of fact-checking in contemporary journalism. This perceived intersectionality of audience trust and fact-checking among fact-checkers connotes a complete openness of cooperation and inclusiveness in the process of fact-checking between the fact-checkers and their audiences.

The analysis suggests that, with the exception of one participant, the remaining respondents supported an approach to fact-checking that involves actively involving their audiences. This includes activities such as providing fact-checking training sessions, establishing feedback mechanisms, and encouraging interactions. The aim is to bridge the existing gap between fact-checkers and their audiences. This fact-checkers' proposed working model aligns with Green-Barber & McKinley's (2019) understanding of the burgeoning phenomenon of "engaged journalism." According to these authors, this model should foster cooperation, establish a comprehensive working environment between journalism and its audiences, and enable audience participation in the journalistic process. More specifically, by allowing members of the audience to suggest content to fact-check, fact-checkers apply guided professional reporting, a form of engagement where the audience is involved in the news process while the journalists keep the responsibility of producing the work (Nip, 2010). We found out that fact-checkers have seen the significant potential to foster closer engagement with the audience while carrying out their activities. For instance, they perceived invaluable support from their audience in terms of information consumption and interactions, viewing this collaboration as a significant achievement in their working processes. They believed that their credibility in providing fact-checking services could be strengthened alongside the growing cooperation and trust from their audiences and showed concern that their work would appear less valuable and insignificant without the trust and engagement of their audiences.

The findings showed that fact-checkers not only perceived audience trust as integral to their fact-checking processes but also identified divisions within their audiences. For instance, all the informants interviewed for this research concurred that their audience fell into two distinct categories.

The first category does not appreciate alternative ideologies in information but rather believes in their own understanding of ideas and information they encounter. This group tends to favor fact-checking and media content that aligns with their political leanings, as suggested by Pedro Baptista et al. (2023). They tend to maintain their own interpretation of information using their own beliefs, values, and political orientations. Furthermore, they may prioritize confirming information they perceive as true rather than verifying potentially false information, as noted by Ognyanova (2024).

On the other hand, there's a category that exhibits curiosity and a willingness to engage with the fact-checkers' work, driven by a desire for accuracy in the information they encounter. The analysis shows that fact-checkers highly value this audience segment, emphasizing the importance of audience engagement and participation in their work, including reading, commenting, and providing feedback. More significantly, respondents consistently cited audiences' political beliefs and ideological orientations as key factors contributing to this division. Nyhan and Reifler (2012) support this perspective. They contended that political ideologies impact audiences' acceptance or rejection of fact-checking and media content. The respondents emphasized the audience's susceptibility to political misinformation, arguing for the need for scrutiny to achieve accurate comprehension. Moreover, this category of misinformation was believed to present a significant challenge, highlighting the importance of fact-checking efforts in addressing and countering its effects (Kyriakidou et al., 2022).

In countries like Mali, the politically charged information landscape is believed to exacerbate divisions among audiences and media perceptions. This political influence shapes audience understanding of information from both fact-checking institutions and the media, with politically biased audiences likely to gravitate towards information that confirms their existing ideologies (Faragó et al., 2020).

Hence, the fact-checkers' advocacy for continued engagement with politically polarized audiences, as evidenced by the respondents, appears to stem from their belief that political information must undergo analysis and verification to provide audiences with accurate information. In their study on fact-checking and disinformation in UK media, Kyriakidou et al. (2022) concluded that integrating fact-checking into journalistic practices is essential. This practice not only helps audiences understand politics, but it also protects them from political misinformation and holds politicians accountable for spreading it.

This underscores the importance of fact-checking initiatives in mitigating the effects of political misinformation and fostering informed public discourse. Moreover, the prevalence of misinformation poses a significant threat to the credibility of news and media, leading to a growing lack of trust in both news institutions and the news itself among different audience groups (Hameleers et al., 2022). The respondents highlighted the challenges faced in fact-checking processes, particularly emerging from audience polarization and the evolving media landscape, such as the proliferation of social media and the emergence of online influencers called "videomen" in Mali's media environment (Sissoko et al., 2024). These challenges have contributed to a crisis of trust in both fact-checking and traditional news media information in Mali, according to the respondents.

This supports the rationale behind fact-checkers' advocacy for engaging and collaborating with their audiences to build trust in fact-checking information and raise awareness about the growing threats of misinformation. Some practitioners view this engagement as a means of building trust and enhancing the relationship between journalism and its audiences (Wenzel & Nelson, 2020). Others perceive it as a deeper cultivation of interactions and relationships between the media and its audiences. Scholars argue that such engagement not only enriches news programming but also serves as a strategic approach to bolstering trust in journalism (Schmidt et al., 2022).

Our findings indicate that the concept of "engaged audiences" in fact-checking, as advocated by the respondents in this study, is gaining traction among Malian fact-checkers. For instance, respondents reported using interactive social platforms like WhatsApp to foster dialogue and address audience concerns regarding news and other types of information encountered. This initiative aims to enhance the relationship between audiences and fact-checking information (Schmidt et al., 2022). Additionally, they have implemented training sessions to educate audiences on the significance of fact-checking. Batsell (2015) underscored the importance of journalism adapting to the current global information landscape by engaging audiences, as their relationship with, accessibility to, and attention to information are vital for the survival of journalism.

Additionally, our findings revealed that fact-checkers have embraced various strategic measures aimed at fostering and maintaining trust among their audiences. Specifically, they have prioritized adherence to journalistic ethics standards such as independence, objectivity, and non-partisanship throughout the fact-checking processes. This suggests a positive correlation between perceived trust and identification with professional standards (Tsfati, 2004).

In terms of independence, fact-checkers approached this principle from two key points of view. First, they operated without the fear of losing a portion of their audience, prioritizing accuracy over audience size. Second, they emphasized transparency by disclosing the sources of financial support for their operations. This dedication to upholding truth and independence reveals the fact-checkers' commitment to maintaining trust and integrity in their work.

Furthermore, our findings underscore the complex nature of fact-checking practices and audience trust. While the respondents acknowledged implementing various strategies to cultivate trust among their audiences, such as transparency, independence, disclosure of financial sources, and objectivity, they recognized the impossibility of satisfying every individual's information needs. This sentiment aligns with the arguments put forth by the majority of the respondents who emphasized the importance of prioritizing ethical approaches in fact-checking as a means of working effectively, even if it results in a smaller audience. Essentially, the focus is not on garnering a large audience to trust fact-checking information but rather on appreciating and serving those who do trust in the craft. This perspective highlights the significance of integrity and commitment to truth in the fact-checking craft, regardless of audience size.

Conclusion

In this chapter, we sought to understand what audience trust means for Malian fact-checkers and what strategies they put in place to build this trust. Our results showed that even if there was no consensus on the definition of the notion of trust among them, fact-checkers attach great importance to trust from the audience. We found that building audience trust is embedded in their work, through their commitment to independence and transparency as well as their permanent engagement with their audiences. With only 5 respondents, the scope of this study is limited, and its results cannot therefore be generalized. However, we believe that this study contributes greatly to the literature on audience trust in the media by approaching the subject from the perspective of media professionals, in this case, fact-checkers. We believe that as important as it is to understand the determinants of audience trust or distrust in the media, it is also instructive to understand media professionals' perceptions of audience trust or distrust and how this affects them, hence the relevance of this chapter. Furthermore, we believe that our results are enlightening and that they lay the foundations for more in-depth studies on fact-checking and public trust. One avenue of research would be to understand how public trust or distrust can influence the way fact-checkers work.

References

Aboutayeb, M. (2022). La technologie vs la presse: quel impact ? *Akofena,* 005 (1), 287-296.

Afrobarometer (2024). World Press Freedom Day: Africans strongly endorse media's role in holding governments accountable. https://www.afrobarometer.org/wp-content/uploads/2024/05/News-release_World-Press-Freedom-Day-Africans-strongly-endorse-medias-role-in-holding-governments-accountable_30-April.pdf

Akyirem, S., Ekpor, E., Aidoo-Frimpong, G. A., Salifu, Y., & Nelson, L. E. (2024). Online interviews for qualitative health research in Africa: A scoping review. International Health, 16(1), 4–13. https://doi.org/10.1093/inthealth/ihad010

Albarran, A.B. (2010). *The Transformation of the Media and Communication Industries.* Media Markets Monographs. N11.

Algan, Y., & Cahuc, P. (2014). Chapter 2 - Trust, Growth, and Well-Being: New Evidence and Policy Implications. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 2, pp. 49–120). Elsevier. https://doi.org/10.1016/B978-0-444-53538-2.00002-2

Ariely, G. (2015). Trusting the Press and Political Trust: A Conditional Relationship. *Journal of Elections, Public Opinion and Parties, 25*(3), 351–367. https://doi.org/10.1080/17457289.2014.997739

Batsell, J. (2015). *Engaged Journalism: Connecting with Digitally Empowered News Audiences.* Columbia University Press.

Bocande, A., Lagarde, J., Marong, S. (2023). *Dans la peau d'un journaliste au Sahel.* Reporters Sans Frontières. Report. https://rsf.org/sites/default/files/medias/file/2023/04/Dans%20la%20peau%20d%27un%20journaliste%20au%20Sahel_FR_0.pdf

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Bucy, E. P., D'Angelo, P., & Bauer, N. M. (2014). Crisis, Credibility, and the Press: A Priming Model of News Evaluation. *The International Journal of Press/Politics,* 19(4), 453-475. https://doi.org/10.1177/1940161214541682

Butcher, P. (2019). *Disinformation and democracy: The home front in the information war.* European Policy Center. Brussels. https://www.epc.eu/content/PDF/2019/190130_Disinformationdemocracy_PB.pdf

Castells, M. (2007). Communication, power and counter-power in the network society. *International Journal of Communication, 1*(1), 238–266.

Coddington, M., Lewis, S. C., & Belair-Gagnon, V. (2021). The Imagined Audience for News: Where Does a Journalist's Perception of the Audience Come From? *Journalism Studies, 22*(8), 1028–1046. https://doi.org/10.1080/1461670X.2021.1914709

Coleman, J.S. (1974). *Power and the Structure of Society.* Norton, New York.

Colón, A. (2017). You are the new gatekeeper of the news. *The Conversation.* https://theconversation.com/you-are-the-new-gatekeeper-of-the-news-71862

Coulibaly, H. & Lima, S. (2013). Crise de l'État et territoires de la crise au Mali. *Echo-Géo.*

Edelman (2024). 2024 Edelman Trust Barometer. Global Report. https://www.edelman.com/sites/g/files/aatuss191/files/2024-02/2024%20Edelman%20Trust%20Barometer%20Global%20Report_FINAL.pdf

Fawzi, N., Steindl, N., Obermaier, M., Prochazka, F., Arlt, D., Blöbaum, B., Dohle, M., Engelke, K. M., Hanitzsch, T., Jackob, N., Jakobs, I., Klawier, T., Post, S., Reinemann, C., Schweiger, W., & Ziegele, M. (2021). Concepts, causes and consequences of trust in news media – A literature review and framework. *Annals of the International Communication Association, 45*(2), 154–174. https://doi.org/10.1080/23808985.2021.1960181

Fletcher, R. (2020). *Trust will get worse before it gets better. Forward looking essay for Journalism, Media and Technology Trends and Predictions Reports 2020.* https://www.digitalnewsreport.org/publications/2020/trust-will-get-worse-gets-better/

Faragó, L., Kende, A., & Krekó, P. (2020). We Only Believe in News That We Doctored Ourselves. *Social Psychology, 51*(2), 77–90. https://doi.org/10.1027/1864-9335/a000391

Green-Barber, L., & McKinley, E. G. (2019). *Engaged Journalism: Practices for Building Trust, Generating Revenue, and Fostering Civic Engagement.* Available at: https://s3-us-west-2.amazonaws.com/lindsaygreenbarber.com/assets/IA+Engaged+Journalism+Report+1.31.19.pdf

Guess, A., Nyhan, B., & Reifler, J. (2018). *Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US. presidential campaign* (United States of America) [Report]. Dartmouth College. fake-news-2016.pdf (bpb-us-e1.wpmucdn.com)

Hagberg, S., & Körling, G. (2012). Socio-political Turmoil in Mali: The Public Debate following the Coup d'État on 22 March 2012. *Africa Spectrum, 47*(2–3), 111–125. https://doi.org/10.1177/000203971204702-306

Hameleers, M., Brosius, A., & de Vreese, C. H. (2022). Whom to trust? Media exposure patterns of citizens with perceptions of misinformation and disinformation related to the news media. *European Journal of Communication, 37(3)*, 237–268. https://doi.org/10.1177/02673231211072667

Hancock, P. A., Kessler, T. T., Kaplan, A. D., Stowers, K., Brill, J. C., Billings, D. R., Schaefer, K. E., & Szalma, J. L. (2023). How and why humans trust: A meta-analysis and elaborated model. *Frontiers in psychology, 14*, 1081086. https://doi.org/10.3389/fpsyg.2023.1081086

Hanitzsch, T., Van Dalen, A., & Steindl, N. (2018). Caught in the Nexus: A Comparative and Longitudinal Analysis of Public Trust in the Press. *The International Journal of Press/Politics, 23*(1), 3–23. https://doi.org/10.1177/1940161217740695

Harcup, T. (2014). *A Dictionary of Journalism.* Oxford: Oxford University Press.

IFCN (n.d). *The commitments of the Code of Principles.* https://ifcncodeofprinciples.poynter.org/the-commitments

Internews (2020). *First Grants Awarded to Journalists and Newsrooms from Internews Information Saves Lives Rapid Response Fund.* https://internews.org/first-grants-awarded-journalists-and-newsrooms-internews-information-saves-lives-rapid/

Jakobsson, P., & Stiernstedt, F. (2023). Trust and the Media: Arguments for the (Irr)elevance of a Concept. *Journalism Studies, 24*(4), 479–495. https://doi.org/10.1080/1461670X.2023.2169191

Keen, S., Lomeli-Rodriguez, M., & Joffe, H. (2022). From Challenge to Opportunity: Virtual Qualitative Research During COVID-19 and Beyond. *International journal of qualitative methods, 21.* https://doi.org/10.1177/16094069221105075

Kim, H. S., Suh, Y. J., Kim, E., Chong, E., Hong, H., Song, B., Ko, Y., & Choi, J. S. (2022). Fact-Checking and Audience Engagement: A Study of Content Analysis and Audience Behavioral Data of Fact-Checking Coverage from News Media. *Digital Journalism, 10*(5), 781–800. https://doi.org/10.1080/21670811.2021.2006073

Kiousis, S. (2001). Public Trust or Mistrust? Perceptions of Media Credibility in the Information Age. *Mass Communication and Society, 4*(4), 381–403. https://doi.org/10.1207/S15327825MCS0404_4

Kohring, M., & Matthes, J. (2007). Trust in News Media: Development and Validation of a Multidimensional Scale. *Communication Research, 34*(2), 231-252. https://doi.org/10.1177/0093650206298071

Kramer, R.M. (1999). Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual Review of Psychology,* 50(1), 569-598.

Kyriakidou, M., Cushion, S., Hughes, C., & Morani, M. (2022). Questioning Fact-Checking in the Fight Against Disinformation: An Audience Perspective, *Journalism Practice,* 17(10), 2123-2139, https://doi.org/10.1080/17512786.2022.2097118

Laplace, M. (2022) Mali : l'Union européenne accusée de vouloir corrompre les blogueurs maliens. *Jeune Afrique.* https://www.jeuneafrique.com/1345156/politique/mali-lunion-europeenne-accusee-de-vouloir-corrompre-les-blogueurs-maliens/

Lewis, J. D., & Weigert, A. (1985). Trust as a social reality. *Social Forces, 63*(4), 967–985. https://doi.org/10.2307/2578601

Litt, E. (2012). Knock, Knock. Who's There? The Imagined Audience. *Journal of Broadcasting & Electronic Media, 56*(3), 330–345. https://doi.org/10.1080/08838151.2012.705195

Litt, E., & Hargittai, E. (2016). The Imagined Audience on Social Network Sites. Social Media + Society, 2(1). https://doi.org/10.1177/2056305116633482

Liu, X., Qi, L., Wang, L., & Metzger, M. J. (2023). Checking the Fact-Checkers: The Role of Source Type, Perceived Credibility, and Individual Differences in Fact-Checking Effectiveness. *Communication Research,* 0(0). https://doi.org/10.1177/00936502231206419

L'œil de la Maison des Journalistes (2023). Portrait. „Nous n'avons rien fait de mal, juste tendu notre micro et allumé la caméra". https://www.oeil-maisondesjournalistes.fr/2023/06/05/portrait-nous-navons-rien-fait-de-mal-juste-tendu-notre-micro-et-allume-la-camera/

Lorgerie, P. (2022). Au Mali, les journalistes sont muselés au profit des influenceurs. *Ouest France.* https://www.ouest-france.fr/monde/mali/au-mali-journalisme-patriotique-contre-liberte-de-la-presse-4956420e-ca2a-11ec-8113-3d2243031bc2

Maati, A., Edel, M., Saglam, K., Schlumberger, O., & Sirikupt, C. (2023). Information, doubt, and democracy: how digitization spurs democratic decay. *Democratization,* 1–21. https://doi.org/10.1080/13510347.2023.2234831

Mabweazara, H. M. (2014). Introduction, Digital *Journalism, 2*(1), 2–11, https://doi.org/10.1080/21670811.2013.850195

Marwick, A. E., & Boyd, D. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society, 13*(1), 114–133. https://doi.org/10.1177/1461444810365313

MFWA (2021). Enlèvement du journaliste Olivier Dubois au Nord du Mali: dangers d'informer en zones sous menaces djihadistes. Media Foundation for West Africa, Accra. https://www.mfwa.org/fr/issues-in-focus/enlevement-du-journaliste-olivier-dubois-au-nord-mali-dangers-dinformer-en-zones-sous-menaces-djihadistes/

MFWA (2022). Etat de la liberté de la presse au Mali. Rapport 2021. https://www.mfwa.org/wp-content/uploads/2023/01/Mali-Rapport-State-of-Press-Freedom.pdf

Min, S.-J. (2016). Conversation through journalism: Searching for organizing principles of public and citizen journalism. *Journalism, 17*(5), 567-582. https://doi.org/10.1177/1464884915571298

Moehler, D. C., & Singh, N. (2011). Whose News Do You Trust? Explaining Trust in Private versus Public Media in Africa. *Political Research Quarterly, 64*(2), 276–292. http://www.jstor.org/stable/23056391

Mustapha, J. M., Shilina, M. G., Agyei, O. S., & Ocansey, C. R. (2022). News Media Trust and Sources of Political Information in West Africa: Mainstream vs. New Media in Ghana and Nigeria. *RUDN Journal Of Studies In Literature And Journalism, 27*(1), 200-208. https://doi.org/10.22363/2312-9220-2022-27-1-200-208

Mutsvairo, B. (2016). Recapturing citizen journalism: process and patterns. In Mutsvairo, B. (Ed.), *Participatory politics and citizen journalism in a networked Africa: a connected continent* (pp. 1–15). London: Palgrave Macmillan

Mwangi, M. (2023). Technology and Fake News: Shaping Social, Political, and Economic Perspectives. *Biomed Sci Clin Res, 2*(2), 221–236.

Napoli, P.M. (2011) *Audience Evolution: Technology and the Transformation of Media Audience.* Columbia University Press, New York.

Nelson, J. L. (2018). The Elusive Engagement Metric. *Digital Journalism, 6*(4), 528-544. https://doi.org/10.1080/21670811.2018.1445000

Newman, N. and Fletcher, R. (2017) B*ias, Bullshit and Lies: Audience Perspectives on Low Trust in the Media*. https://doi.org/10.2139/ssrn.3173579

Newman, N., Fletcher, R., Eddy, K., Robertson, C. T., & Nielsen, R. K. (2023). *Reuters Institute Digital News Report 2023.* Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2023-06/Digital_News_Report_2023.pdf

Nip, J. Y. M. (2010). 'Routinization of Charisma. The Institutionalization of Public Journalism Online'. In Rosenberry, J. & St John, B. (Eds.). *Public Journalism 2.0: The Promise and Reality of a Citizen Engaged Press* (pp. 135-148). Routledge.

Nyhan, B., & Reifler, J. A. (2012). *Misinformation and fact-checking: Research findings from social science.* New America Foundation Media Policy Initiative Research Paper. Washington, DC

Ognyanova, K. (2024). Fact-Checking: Journalistic Strategies and Audience Outcomes in Diverse National Contexts. *Journalism & Mass Communication Quarterly,* 10776990241238247. https://doi.org/10.1177/10776990241238247

Otto, K., & Köhler, A. (Eds.). (2018). *Trust in media and journalism. Empirical perspectives on ethics, norms, impacts and populism in Europe.* Springer.

Ouedraogo, L. (2022). *Mali's fake news ecosystem: An overview.* Center for Democracy and Development. Abuja. Research Report.

Patton, M. Q. (2002). *Qualitative evaluation and research methods.* Thousand Oaks, CA: Sage.

Pedro Baptista, J., Gradim, A., Loureiro, M., & Ribeiro, F. (2023). Fact-Checking: A Journalistic Movement yet to Be Discovered? Audience Attitudes and Familiarity

Levels in Portugal. In F. J. García-Peñalvo & A. García-Holgado (Eds.), *Proceedings TEEM 2022: Tenth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 719–727). Springer Nature. https://doi.org/10.1007/978-981-99-0942-1_74

Pellerin, M. (2019). *Les violences armées au Sahara. Du djihadisme aux insurrections?* Policy Center for the New South/IFRI https://www.ifri.org/sites/default/files/atoms/files/pellerin_violence_armees_sahara_2019_002.pdf

Perret, T. (2005). Médias et démocratie au Mali: Le journalisme dans son milieu. *Politique africaine, 97*, 18-32. https://doi.org/10.3917/polaf.097.0018

Primig, F. (2022). The Influence of Media Trust and Normative Role Expectations on the Credibility of Fact Checkers. *Journalism Practice,* 1–21. https://doi.org/10.1080/17512786.2022.2080102

Prochazka, F., & Schweiger, W. (2018). How to Measure Generalized Trust in News Media? An Adaptation and Test of Scales. *Communication Methods and Measures, 13*(1), 26–42. https://doi.org/10.1080/19312458.2018.1506021

Rosen, J. (1999). *What Are Journalists For?* Yale University Press. http://www.jstor.org/stable/j.ctt32bw3z

Rosenberry, J. & St John, B. (2010). *Public Journalism 2.0: The Promise and Reality of a Citizen Engaged Press.* Routledge.

Schmidt, T. R., Nelson, J. L., & Lawrence, R. G. (2022). Conceptualizing the Active Audience: Rhetoric and Practice in "Engaged Journalism". *Journalism, 23*(1), 3-21. https://doi.org/10.1177/1464884920934246

Shin, J. & Thorson, K. (2017). Partisan selective sharing: The biased diffusion of fact-checking messages on social media. *Journal of Communication,* 67, 233–255.

Sissoko, E.F., & Dembélé A., (2023). La digitalisation et le journalisme : L'impact des réseaux sociaux et du journalisme citoyen sur la scène médiatique malienne. *Revue Francophone d'Économie et de Gestion,* 2(1), 152-176.

Sissoko, E.F., Tangara, T., Kone, B., Kone, M. (2024). L'économie de la vidéomania au Mali : Exploration d'un nouveau métier médiatique et son impact sociétal. *Revue Française d'Economie et de Gestion, 5*(2). https://www.revuefreg.fr/index.php/home/article/view/1499

St. John, B., & Johnson, K. A. (2021). Public Journalism Without the Public: Problematizing the Public Sphere and Press Credibility in Academic Journals, 1991–2018. *Journal of Communication Inquiry, 45*(2), 178-198. https://doi.org/10.1177/0196859920961029

Steindl, N., Obermaier, M., Fawzi, N., & Lauerer, C. (2023). Explaining media trust among journalists and recipients: Different experiences, different predictors. *Journalism*. Online first. https://doi.org/10.1177/14648849231190698

Sullivan, J. (2020). History and concept of the audience. In *Media Audiences: Effects, Users, Institutions, and Power* (Second Edition ed., pp. 1-26). SAGE Publications, https://doi.org/10.4135/9781071872406

Togola, M., & de Bruijn, M. (2023) Les réseaux sociaux dans la dynamique des conflits au centre du Mali: un exemple de journalisme citoyen à travers la plateforme numérique KI, C*anadian Journal of African Studies / Revue canadienne des études africaines, 57*(2), 305-325, https://doi.org/10.1080/00083968.2023.2177689

Tsfati, Y. (2004). Exploring Possible Correlates of Journalists' Perceptions of Audience Trust. J*ournalism & Mass Communication Quarterly, 81*(2), 274-291. https://doi.org/10.1177/107769900408100204

Tsfati, Y. (2010). Online News Exposure and Trust in the Mainstream Media: Exploring Possible Associations. *American Behavioral Scientist, 54*(1), 22-42. https://doi.org/10.1177/0002764210376309

Tsfati, Y., Strömbäck, J., Lindgren, E., Damstra, A., Boomgaarden, H., & Vliegenthart, R. (2022). Going Beyond General Media Trust: An Analysis of Topical Media Trust, its Antecedents and Effects on Issue (Mis)perceptions. *International Journal of Public Opinion Research, 34*(2). https://doi.org/10.1093/ijpor/edac010

Uth, B., Stehle, H., Wilhelm, C., Detel, H., & Podschuweit, N. (2023). The journalism-audience relationship in the digital age: A theoretical literature review. *Journalism,* 0(0). https://doi.org/10.1177/14648849231221611

Valenzuela, S., Halpern, D., & Araneda, F. (2022). A Downward Spiral? A Panel Study of Misinformation and Media Trust in Chile. *The International Journal of Press/Politics, 27*(2), 353-373. https://doi.org/10.1177/19401612211025238

van Dalen, A. (2020). Journalism, trust, and credibility. In K. Wahl-Jorgensen & T. Hanitzsch (Eds.), *The Handbook of Journalism studies* (2. ed., pp. 356-371). Routledge. https://doi.org/10.4324/9781315167497-23

van Erkel, P. F. A., van Aelst, P., de Vreese, C. H., Hopmann, D. N., Matthes, J., Stanyer, J., & Corbu, N. (2024). When are Fact-Checks Effective? An Experimental Study on the Inclusion of the Misinformation Source and the Source of Fact-Checks in 16 European Countries. *Mass Communication and Society,* 1–26. https://doi.org/10.1080/15205436.2024.2321542

Van Manen, M. (1997). *Researching lived experience: human science for an action sensitive pedagogy.* London and New York: Routledge Taylor and Francis Group

Vizoso, Á., & Vázquez-Herrero, J. (2019). Fact-checking platforms in Spanish. Features, organisation, and method. *Communication & Society, 32*(1), 127-142. https://doi.org/10.15581/003.32.37819

Wasserman, H., & Madrid-Morales, D. (2019). An Exploratory Study of "Fake News" and Media Trust in Kenya, Nigeria and South Africa. *African Journalism Studies, 40*(1), 107–123. https://doi.org/10.1080/23743670.2019.1627230

Wenzel, A., & Nelson, J. L. (2020). Introduction "Engaged" Journalism: Studying the News Industry's Changing Relationship with the Public. *Journalism Practice, 14*(5), 515–517. https://doi.org/10.1080/17512786.2020.1759126

Zollo, F. (2019). Dealing with digital misinformation: a polarised context of narratives and tribes. *EFSA Journal* 17(S1):e170720, 15 pp. https://doi.org/10.2903/j.efsa.2019.e170720

Crisis Pregnancy Centers: Digital Rhetoric, Misinformation, and Trust

Kirsti K. Cole and Johanna M. Wagner

Introduction

With the explosive rise of social media, algorithms, and digital engagement, political and social ideologies are tools through which to sort the flow of data. A simple Google search for "abortion care," for example, will result in vastly different outcomes depending on a user's location and search history, as well as a plethora of keywords. In the United States, abortion has been a fraught issue since Roe v. Wade was decided in 1973. Since then, anti-choice activists have taken advantage of the digital revolution by crafting various tools that engage in misinformation regarding the topic. From deceptive imagery to false medical information, people who believe women do not have rights regarding their own bodies attempt to gain the public's trust while misleading them and misrepresenting themselves at every step of the pregnancy process.

Some of the most organized, coordinated, and insidious deception is online and comes from one source collectively known as Crisis Pregnancy Centers (CPCs). The rhetoric of these anti-abortion organizations has moved quietly and effectively from the physical to the digital; the word-of-mouth communication, leaflets, posters, and billboards that marked (most of) their information distribution in the twentieth century have given way to online websites in the twenty-first. Some websites are basic, some elaborate, but what they all have in common is that they radiate a constant stream of misinformation; after all, their sole purpose is to keep women from accessing abortion.

This chapter argues that Crisis Pregnancy Centers have burgeoned across the US since the latter 20th century, and misinformation, deceptive facts and figures, as well as outright lies guide their digital footprint. The histories of these centers intersect with invisible phenomena ever-present in twenty-first-century American politics: *anti-intellectualism* – a strain of animosity toward education and expertise in the American public, coined by Hofstadter (1963) – and, *The War on Women* (Melich, 1996; Faludi, 1991) – a branch of the American culture wars that devalues women generally and is specifically bent on restricting their freedom to control their own bodies (Wagner, 2019). These phenomena exacerbate the pervasive and devious ways in which CPCs act on the web. Of particular importance since the Supreme Court overturned Roe v. Wade in 2022, we argue, are the intersections of these phenomena with disingenuous digital political rhetoric, and their impacts on structural and interpersonal trust throughout the country.

Drawing on theories of digital rhetoric (Eyman, 2015; VanKooten, 2016), digital writing research (McKee & DeVoss, 2007), and hyperlink network analysis

(Park & Thelwall, 2003), the chapter wades into how CPCs use their website material (text, graphics, audio, images, video) to masquerade as legitimate medical facilities. The chapter examines the websites of eighteen CPCs in five states. The choice of these states – Arizona, Nevada, Minnesota, Wisconsin, and Rhode Island– was entirely practical as they have not previously been examined in other studies about CPCs to date. Websites were chosen randomly, with the only criterion being that they represent cross-sections of urban and rural populations within the state. The reason we chose both urban and rural areas is because of the differences we found in the website schemes. The simpler – usually rural – websites were minimalist in design, and monolithic in images and messaging, while more complex variation in design, messaging, and diverse[1] imagery were found on the urban websites. These findings will be discussed more fully later in the chapter.

CPCs are nonprofit organizations that "purport to provide free services to women who are considering terminating their pregnancies. However, their 'paramount, and typically undisclosed, mission is to convince women not to have abortions'" (Gilbert qtd in Brown, 2018, p. 223). As of 2020, the number of CPCs operating across every state in the US totalled 2,527 (Swartzendruber & Lambert, 2020). During the Trump Administration, Title X was revised to allow CPCs access to the federal Department of Health and Human Services Family Planning Grants (United States Congress, 2021) even though they provide no access to contraception. While over 60% of CPCs provide some kind of "medical" care such as ultrasound or testing for sexually transmitted infections, they deceive potential clients by providing inaccurate and in some cases medically dangerous information (Bryant & Swartz, 2018).

An interesting aspect of abortion rhetoric in the US is the Constitutional right to free speech. Free speech has become a *de facto* tactic for spreading medical misinformation. For example, after the state of California attempted to regulate unlicensed CPCS through the 2015 California Reproduction Freedom, Accountability, Comprehensive Care, and Transparency Act, the Supreme Court ruled in favor of CPC's First Amendment rights by striking down the state law that required CPCs to disclose that they are not health facilities (Rubin, 2018). Legal protections such as these mean that CPCs are not regulated in the same ways as healthcare facilities and in most cases are not regulated at all (Ahmed, 2015).

A combination of religious ideology, anti-intellectualism, and legal undermining of human rights in the US has capitalized on a context created by Republican politicians begun in the early 1980s, entrenched by George W. Bush in the early 2000s, and engrained by the Trump administration and his Supreme Court in the 2020s. Though these actions are directly contrary to the will of the American people (61% say abortion should be legal and accessible, Hartig, 2022), the per-

1 It should be noted that the use of "diverse" in this chapter refers to visual imagery used on the websites, which denotes more the websites' ideas of *diversity* than the authors of this chapter. Diversity seems indicated via visual racial or cultural markings, such as skin tone and hair texture and color.

vasive political rhetoric and real legal protections of CPCs compound the lack of public trust and digital activism surrounding abortion and abortion care.

Methods: Rhetoric, Knowledge Production, and Meaning

As this chapter explores how CPCs use digital communication in their marketing, it is necessary to focus on digital rhetoric methods. Although it remains undertheorized, digital rhetoric builds on current rhetorical theory. Still, because it concentrates on the digital, it inevitably moves away from the importance of the spoken or written word. As Eyman suggests, the power of rhetoric is double, (2015). Depending on the working perspective, one can "use" it, and "study" it. It is a "guide for the production of persuasive discourse" and an "analytic method" (2015, p. 16). Likewise, Buchanan notes "rhetoric is both the practice of persuasive communication and a formal art of studying such communication" (1985, p. 6). As academics, we exercise this doubleness by tapping into our knowledge of rhetoric to investigate the various underlying rhetorical processes employed in human communication and interaction. This is not simply an academic exercise; it is the discovery and identification of the ways in which knowledge and meaning are produced at a given time and place. For Eyman, focusing on the digital, and deciphering its ability for "knowledge production and meaning-making" (2015, p. 17) is at the core of digital rhetoric methods. Eyman emphasizes this strain of thought, "rhetoric is synonymous with meaning, for meaning is in use and context, not words themselves. Knowledge and belief are products of persuasion," (Eyman 2015, p. 16). Following Eyman's, as well as Bizzell's and Hertzberg's ideas, this chapter focuses especially on the ways meanings are made by CPCs digital communication, and how that meaning subtly reproduces foundational ideologies that construct CPCs. McKee and DeVoss (2007) define "digital writing research" as research that focuses in part "on the interactions of people who use digital technologies to communicate" (3).

In this chapter, we are interested in how meaning is made by employing digital artifacts, and as such, we situate our research in DePew (2007) who argues for the importance of triangulation, which means looking not just at texts but at contexts and users. After all, questions of meaning are what drive the paranoid style and anti-intellectual thought in the US: emotional belief overcomes reason in the paranoid style, while anti-intellectual thought resists knowledge outside one's belief. CPC digital rhetoric intersects with US anti-intellectual thought and the paranoid political style to shore up CPC supporters' beliefs. Among other things, the CPC's engagement with both practices signifies a microcosm of societal mistrust of authority, believing solely in its own ability to extract "true" meanings behind events and issues in the culture. It also presents a mistrust of individuals, as the CPC positions itself to convey such meaning. The rise of CPCs and their associated digital presentation demonstrates a crisis of trust in the individual as well since CPCs do not trust the individual to understand the world around them, especially regarding pregnancy and abortion.

Using the Crisis Pregnancy Center Map assembled by Swartzendruber and Lambert (2021), we chose 5 states not yet explored in literature published on CPCs. We collected 18 sites to provide a representative sample of the digital footprint. These samples are locationally relevant and include clinics in Phoenix, AZ, Tucson, AZ, Sierra Vista, AZ, Las Vegas, NV, Reno, NV, Pahrump, NV, Madison, WI, Milwaukee, WI, Stevens Point, WI, Minneapolis, MN, Mankato, MN, Detroit Lakes, MN, Providence, RI, and Westerly, RI. For each state selected we chose a CPC website from a large urban center, a mid-sized city (usually with populations between 45,000-80,000), and a small, rural community. This selection process gave us a broad data spread in terms of ways in which CPCs tailor their messaging and differentiate their design.

Context: Anti-intellectualism as US Ideology

Anti-intellectualism is not a new phenomenon in the US. It is not even modern. In his germinal text *Anti-Intellectualism in American Life* (1962), American historian Richard Hofstadter states that American "anti-intellectualism is, in fact, older than our national identity," and appears prone to "cyclical fluctuations" (1962, p. 6). He labels anti-intellectual thought as a "resentment and suspicion of the life of the mind and of those who are considered to represent it; and a disposition constantly to minimize the value of that life" (1962, p. 7). Among other things, he asserts, anti-intellectual thought suggests that the "plain sense of the common man […] is an altogether adequate substitute for, if not actually much superior to, formal knowledge and expertise acquired in the schools" (1962, p. 19).

Building upon the common sense vs. expertise paradigm, a style identified by Hofstadter as "paranoid," is attached to US political rhetoric. In his essay "The Paranoid Style in American Politics" Hofstadter portrays this style as grounded in anger, which "evokes the sense of heated exaggeration, suspiciousness, and conspiratorial fantasy" (1964, p. 77). For Hofstadter, it is because usage of the mode is not limited to fanatics or those we might describe as "disturbed," but by "more or less normal people that makes the phenomenon significant" (1964, p. 77). The paranoid style should act as a warning, according to Hofstadter, as it "has a greater affinity for bad causes than good" (1964, p. 77). He notes further, "The paranoid style is an old and recurrent phenomenon in our public life which has been frequently linked with movements of suspicious discontent" (1964, p. 77). In this style, we begin to see the ways in which an angry, frightened, and distrustful electorate would be fertile ground to sow anti-intellectual thought, which often relies on emotion and disdain rather than factual evidence or truth. Between the anti-intellectual strain and the paranoid style, American politics has become a festering nest of hysterical unreason and dangerous hyperbole.[2] These ways of

2 This thought and style have encroached on important aspects of American life in recent years. Currently, the peaceful transition of power in US democracy has become a lightning rod for extremists: significant numbers of Republican voters display a propensity to believe in various

looking into American political thought and discourse seem most vital, according to Hofstadter, when they "become effective in our affairs, gravely inhibit[ing] or impoverish[ing] intellectual and cultural life" (1962, p. 7). Since American political thought has become increasingly anti-intellectual and paranoid in the last fifty years, especially as espoused by extremist Republican (GOP) politicians and pundits, we scrutinize these phenomena, especially as we investigate the digital rhetoric of Crisis Pregnancy Centers and those who advocate for them.

Anti-intellectualism in the US is entrenched in the rise of CPCs. Just as CPCs are entrenched in American Christian evangelical religious traditions, so too is the anti-intellectualism that seeks to block information that might run counter to religious ideology (sexual health, LGBTQIA+, BIPOC histories and lived experiences, etc.). Anti-intellectualism is anti-science and works in two ways to undermine accurate information: first, it undermines the structural trust of the population towards institutions and government (since the Reagan era), and second, it undermines the interpersonal trust that happens between groups (Democrats and Republicans, Christians and secular humanists, etc.), or between individuals.

Discussion: CPCs, Digital Rhetoric, and Trust

"Crisis pregnancy centers have the stated goal of preventing abortions, and, based on many of their websites, appear to use tactics that scare women in order to dissuade them" (Bryant et al., 2014, p. 605). CPCs, also known as pregnancy resource centers (PRCs), pregnancy care clinics (PCCs), anti-abortion pregnancy centers (PLPCs), or fake women's health centers are non-profit organizations established by anti-abortion groups in the United States. Hundreds of CPCs are also operational in Europe, Africa, Canada, and Latin America. The major organizations that operate CPCs are Care Net, Heartbeat International, Birthright International, and the National Institute of Family and Life Advocates (NIFLA). Most CPCs are religiously affiliated (largely non-denominational evangelical Christians, Gibbs, 2007) with these national or international organizations, following the path of the Pearson Foundation, the first such organization established by a Catholic man in Hawaii after abortion was legalized there in 1967 (Morrison, 2019). Robert Pearson formed the Foundation and established the playbook for CPCs, including instructions for CPC management and propaganda brochures and images, some of which are still in circulation today (Griswold, 2019). Birthright International established the first network of CPCs in Canada in 1968, Heartbeat International was founded in 1971, and Care Net was established in Maryland in 1980. In 2007, Care Net and Heartbeat International, both explicitly evangelical, accounted for the majority of CPCs in the US. In 2020, operational CPCs totaled 2,527 (Swartz-

conspiracy theories – concerning illegal voting and stolen elections – as well as becoming involved in anti-government activities, such as the January 6, 2021 insurrection at the Washington D.C. Capitol building.

endruber & Lambert, 2020), while abortion clinics providing medical services were 790 (Johnson 2022).

In the only comprehensive study of its kind, Swartzendruber and Lambert (2020) established the CPC Map to provide accurate data on CPCs. They found that "state funding for CPCs was positively associated with the number of CPCs, and a greater number of CPCs predicted the introduction of extreme state legislation restricting abortion" (9).

Traditionally CPCs provide pregnancy testing and counseling services but have become increasingly medicalized in recent years. Medical services might include limited obstetric ultrasounds to confirm pregnancy, and testing for some sexually transmitted infections, though CPCs do not conform to "national quality family planning service recommendations that define a core set of services to prevent missed opportunities for comprehensive prevention and treatment" (Swarzendruber et al., 2018). In fact, most CPCs seem to intentionally mislead their clients:

> Lay volunteers who are not licensed clinicians at CPCs often wear white coats and see women in exam rooms. They also purport to provide medical advice on a variety of issues, including sexually transmitted infections, early pregnancy, and abortion. Because centers are sometimes located close to abortion clinics and have names and logos similar to nearby abortion clinics, women could mistakenly seek care there rather than at the intended clinic. They also seek to target women who are most likely to seek abortion, particularly low-income women and women of color.[3] (Bryant & Swartz, 2018, p. 270).

CPCs remain largely unregulated and their websites misrepresent who they are and what they do. The misinformation perpetuated by these organizations is particularly concerning because of the ways they target multi-marginalized women in the US. By targeting low-income women of color, CPC rhetoric has the manifest potential to negatively impact existing socio-political inequalities around an intersectional understanding of race and ethnicity, class, gender, mobility, and place. For example, the study of Bryant et al. (2014), found that "most crisis pregnancy centers listed in state resource directories for pregnant women provide misleading or false information regarding the risks of abortion" (p. 601). The reality that is being produced, and the misinformation around the health of the mother

3 It is imperative to note that the number of low-income women and women of color seeking abortions is higher than other groups because of systemic factors. According to Dehlendorf, Harris, and Weitz, "[d]isparities in abortion rates are related to disparities in unintended pregnancy, and associated disparities in contraceptive use. Structural factors, including economic disadvantage, neighborhood characteristics, lack of access to family planning, and mistrust in the medical system underlie these findings" (2013, p. 1775). CPCs and other anti-abortion organizations use anti-intellectual and paranoid impulses to flip the factual evidence of inequality on its head: rather than acknowledging abortion as an effect of societal health and welfare inequality, CPCs argue that abortion itself is a cause of exploitation of these risk groups. In all actuality, this anti-intellectual rationale is tied to a gut-feeling that weighs religious belief more heavily than science, crucially exhibiting CPCs exploitation of the very people they purportedly aim to protect.

and the concept of the fetus are anti-science, and the anti-science stance of CPCs is reinforced by the rhetoric of paranoid, anti-intellectualism that tells women they cannot trust their doctors; therefore, typically underqualified care providers are redefining what science tells us. In Bryant et al.'s analysis of 254 CPC websites (2014), 204 had misleading or false information. The health of the mother was the predominant focus of the websites. For example, 51 websites made the claim that having an abortion can cause breast cancer and 186 websites claim that having an abortion causes severe mental health issues (604). At least 10% of the websites claimed that abortion causes "fetal pain" (604). There has, of course, been extensive research on abortion outcomes and multiple studies disprove any link at all between mental health issues and abortion (Major, 2009; Steinberg & Finer, 2011, APA, 2008).

Results: Shaping Trust through Website Design

Medical Misinformation: Half-Truths and Empowerment

Most CPCs are not licensed or equipped to perform any diagnostic or medical procedures and rarely have medical doctors on staff (Holtzman, 2017; Bryant & Swartz, 2018; McKenna & Swartz, 2018; Borrero et al., 2019). They may have access to pregnancy tests and limited ultrasounds; they may have an associated nurse/nurse practitioner, but more commonly, they have no trained medical personnel at all (Borrero et al., 2019; Bryant & Swartz, 2018; McKenna & Murtha, 2021). Websites for CPCs, however, paint a very different picture. CPC websites embed medically inaccurate information with otherwise truthful data and statistics. This combination results in deceptive digital rhetoric: kernels of truth swaddled in misinformation or half-truths. Half-truths, according to Thomas L. Carson's book Lying and Deception (2010), are forms of "spin," the process of "putting an interpretation on events or facts" that make them "biased and unreliable but not necessarily incorrect" (p. 57). Half-truths "selectively emphasize facts that tend to support a particular interpretation or assessment of an issue and selectively ignore or minimize other relevant facts that tend to support contrary assessments" (p. 57-58). Relating to CPCs, there are a number of examples of this kind of 'spin.'

For example, websites that list fetal development exaggerate the growth of the fetus in early development (FCPS, 2020). In order to "prove" to women seeking abortion support that the information they provide about fetuses is accurate, many offer what they call "limited ultrasounds." These ultrasounds are described as follows, "to help you confirm a viable pregnancy, detect a fetal heartbeat, and estimate how far along you are based on fetal measurements. If desired, we will also supply you with a Proof of Pregnancy Form to submit for insurance or benefits" (FCaPC). These ultrasounds are offered to confirm pregnancy and establish a due date, but most clinics have no medical staff who could accurately provide such information (FcaPC, n.d., FCPRC, n.d.; HR, n.d.; OW, n.d.; WCC, n.d.;

WSCMIL, n.d.; WRMCSM, 2023; PWC, 2018). In fact, many CPCs offer to train the very individuals who use their services, as volunteers.

There is a half-truth model prevalent in these organizations because mixing unreliable data with medically accurate terminology can prolong pregnancy. Prolongation is a goal typically because the states in our study have time limitations on legal abortions. For example, Wisconsin's abortion ban—written just one year after Wisconsin became a state (1849)— classifies abortion as a felony, punishable by up to 6 years in prison and a maximum fine of $10,000. The only exception is "to save the life of the mother"; there is no exception for rape, incest, or patient health. Of the 60 CPCs in the state of Wisconsin, 53 provide limited medical services. The majority of these clinics are labeled "care centers" or "pregnancy help centers." Although Rhode Island allows abortion until viability, many providers do not offer them up to that point; what is more, if under the age of 18, one must secure a parent's/legal guardian's permission. In Nevada, women can only choose to have an abortion within 24 weeks of pregnancy. In Arizona, abortions are limited to 15 weeks. There are 73 CPCs in Minnesota, 46 of which provide limited medical services. Choosing to have an abortion is a right protected by the Minnesota Constitution (Doe v. Gomez), and there is no restriction based on gestational age. Though the state laws differ across our sample set, the websites for the CPCs are largely identical in their presentation of information about fetal development and ultrasound use. Most of our sampled websites use the same inaccurate statistics claiming that: "as many as 25% or more pregnancies" (WCC)—"26%" at Care Net (2023)— end in miscarriage. According to the Mayo Clinic, however, the number is between "10% to 20%" ("Miscarriage" 2023). In addition to misinformation about early pregnancy and detection, CPCs also offer inaccurate information about abortions.

In our sample, CPC websites that do offer information regarding abortion ensure an emphasis on its "risks," spinning abortion with negative connotations (AWC, 2023; CareNet, 2023; FcaPC, n.d.; FCPRC, n.d.; HR n.d.; FCPS, 2020). Of our sample of 18 sites, over half prominently describe risks of abortion, offer an "abortion pill reversal" which claims to "reverse chemical abortion" (FCaPC), and describe side effects that emphasize ongoing personal, familial, and, frequently, spiritual regret. On the Mother of Life Pregnancy Center website, in the "Be Informed" drop-down menu, there is a link called "Abortion Pill Reversal." This links to a website of the same name (APR, 2023). This website is rife with medical misinformation including nonexistent studies. In their FAQ dropdown menu, they claim that:

> Initial studies of APR have shown it has a 64-68% success rate. Without the APR treatment, the first abortion pill may fail to abort the pregnancy on its own. In other words, your pregnancy may continue even without APR if you decide not to take the second abortion drug likely prescribed or provided to you. APR has been shown to increase the chances of allowing the pregnancy to continue. However, the outcome of your particular reversal attempt cannot be guaranteed.

This display of medical misinformation is couched in the half-truths outlined by Carson. The website focuses on progesterone in its descriptions of abortion and pregnancy: "Progesterone is the natural hormone in a woman's body that is necessary to nurture and sustain a pregnancy. The first pill in the abortion pill regimen blocks progesterone's actions. By giving extra progesterone, we hope to outnumber and outcompete the first abortion pill in order to reverse the effects and provide an opportunity to save the pregnancy" (APR, 2023). While it is accurate that progesterone helps sustain (*nurture* is not involved with hormonal function) pregnancy, the information that follows about the abortion and reversal pills play on a selective emphasis in order to sway people who may not know enough about pregnancy to understand the misinformation presented.

Together, these rhetorical choices build up cognitive chains of work to reassure women that CPCs are truly "medical" facilities and that women can trust them with their healthcare. In fact, on the website of one CPC in Phoenix, if one scrolls to the bottom of the page under the tab "About us / Our Staff & Mission," one finds the following statement: "*First Way is not a diagnostic medical clinic and can not give medical advice. Please contact your doctor for any symptoms of concern*" (FWPC, 2023). This kind of buried admission notwithstanding, another way CPCs reify their credibility as "medical" centers is through specific kinds of testing and jargon affiliated with the medical community. Some CPCs advertise STD testing and wellness exams, exploiting users' beliefs that "tests" and "exams" are performed by licensed medical practitioners. Terminology regarding the "medical" aspects of their offerings is scattered throughout their web pages: medical jargon such as the pregnancy hormone HCG can be found easily, offering an aura of "medical" knowledge; the term "ectopic pregnancy" generally recurs on various pages of the websites, enabling a kind of frequency illusion; and some sites offer "fetal development" timelines. Referencing medical information and exams, and utilizing medical jargon functions much like the visual rhetoric discussed earlier, raising the sites' medical ethos.

The misinformation used by CPCs is an attempt to gain credibility by framing their "services" as empowering. All 18 of our sample sites assure visitors that they will have clear and reliable knowledge ("find out for sure," "make an informed decision," FCaCP). The sites also use medical/legal language to indicate that the staff at CPCs will inform visitors more thoroughly than a medical clinic or abortion provider ("it's your right," "You owe it to yourself to get complete and accurate medical information. If you're considering an abortion, you have a legal right to be fully informed before making any decision," OW). The rhetorical moves at play, visually and textually, on these websites, are attempts to gain the trust of women as a group and individually. Crafting the facade of medical clinics, CPCs tap into the ethos of scientific reasoning and judgment of the medical community. The irony, of course, is that CPCs themselves do not trust the science offered by the medical community, nor do they trust the judgment of pregnant women searching for services. Driven by religious belief that rejects abortion (and often even contraception), and anti-intellectualism, which harbors frank antagonism toward "elitist" (medical) knowledge, CPCs abuse the trust of their local communities that they are, in fact, what they say they are. Moreover, they actively defy accepted medical practices and protocols vital to the well-being of pregnant women (McKenna & Murtha, 2021). When compounded with misleading medical information, the empowerment rhetoric used by these websites attempts to establish user's trust while simultaneously swaying their opinions away from abortion. Of our sample, particularly urban-based organizations (Phoenix, Las Vegas, Madison, and Minneapolis) have co-opted the "my body my choice" argument. The websites use the language of abortion rights with no irony. One website, in particular, whose rhetorical dance seems more sophisticated than most is one from Las Vegas, NV. Not only do they use the terminology of abortion rights, emphasizing the term "choice" on their home page, but also brazenly imply a countercultural positionality that insists abortion is some kind of consumerism mainstream medicine and culture ("You Deserve The Truth," "Get The Facts," FCPS).

By co-opting the language of choice, misleading women with medical half-truths, and promising empowerment through deceptive knowledge, the content of CPC websites works to shape a user's understanding of and reaction to pregnancy, sex, and abortion. It is not only the language used on the websites, though, that influences audience perceptions, it is also the design of the websites themselves.

Credibility by Design: Website Analysis and User Accessibility

CPC websites use visual media as an heuristic through which pregnant women make cognitive connections that may turn them away from abortion. How does this happen? Through a sophisticated layering of visual cues in design. There are interesting rhetorical moves CPC websites make to establish trust with their users, but which emphasize a general problem that Laura J. Gurak (2018) argues arises in digital rhetoric, especially when handling such specialized information as medical knowledge. Namely, there is a "flattening of traditional information and knowledge hierarchies," which means "anyone with a mobile device and a Twitter feed [can] create what appear to be equivalent truths" (p. 124). Medical information that was only "accessible [...] to trained specialists, is now available to everyone," and regular users of technology can "misinterpret ideas and information that take years of training and experience to comprehend" (p. 124). CPCs take advantage of this democratization of information and rely on regular people misunderstanding their organizations' expertise. They establish trust by *performing* visually in ways that evoke "medical" communities. For example, the Aid to Women Center website shows a visual of an ultrasound, an image that is repeated throughout our sample set and is used to ensure website visitors that medical treatment is available.

Although the anti-intellectual strain in the States means many Americans have a healthy skepticism surrounding doctors in general – as emphasized during the recent pandemic and the various arguments regarding vaccines and masks – CPCs intuit that women continue to assume medical professionals will appropriately attend to their healthcare, which means the website must bend to its users' assumptions. Thus, for example, when opening the homepage of the First Choice Pregnancy Services website in Las Vegas, a woman in a white lab coat, along with a stethoscope befittingly draped about the neck, is superimposed in a pop-up video on the screen. Likewise, the Free Ultrasound tab and icon of First Choice Pregnancy Center in Pahrump (Nevada), and the Hands of Hope, Tucson (Arizona), as well as the Home pages of Life Choices Community Pregnancy Clinic in Carson City (Nevada), and Phoenix Women's Clinics (Arizona) showcase images or videos of figures with laboratory coats and/or scrubs.

"Scientific" props are used as visual cues by CPCs in performatively appropriating the idea of a "medical" website. The use of these cues is meant to build user credibility and is designed to make CPC spaces appear safe, professional, medically accurate, and official. Images of pregnancy tests and ultrasound equipment appear prominently and repeatedly while scientific-sounding terms describe them. Pregnancy tests, for example, are not only free, but "high-sensitivity lab-quality," "clinical grade" (FCPS, 2020; FCPC, n.d.), "laboratory rated" (CareNet, 2023), "laboratory-grade" (PWC, 2018), and "medical-grade" (RCWC, 2022, AWC, 2023). 12 of the 18 sample sites have images of exam rooms with "medical" paraphernalia, such as ultrasound equipment, examination tables, and sterile-appearing rooms with associated medical supply cabinets and counters. And finally, ultrasounds – like pregnancy tests – are almost always offered free, or at "low cost" (Aid, 2023), which can foster unsubstantiated connections for the user between the CPC and government-assisted healthcare programs such as Medicaid.

It is not just medical imagery that CPC sites use to promote accessibility and credibility. All 18 of the sites, regardless of the level of design sophistication (typically the more rural the location of the center, the less design-intensive the website), show images of smiling, happy, excited, young, and racially diverse women, some pregnant, some not. These images are drawn from standard, open-access stock images but always display women in one of two ways, either happy or peacefully contemplative.

The structure of these sites, pairing medical imagery and imagery of thoughtful, diverse, and typically smiling women work together to visually craft credibility and trust as enticing features of their virtual and physical spaces. They direct and cue a user, on almost every page of each website, to focus on the image and on buttons that say "Contact Us!" or "Free Pregnancy Test Today!" or "Walk-Ins Welcome." Through Hyperlink Network Analysis (HNA), we understand an individual website to function not just as a space, but as an actor who influences trust and credibility in other websites (Park, 2003). The visual design of CPC websites is similar enough, and they all use similar design principles to impact trust across the hyperlinked networks in which CPCs operate. In the era predating the widespread use of the internet, anti-abortion activists drew their inspiration from the Pearson Foundation, however, that playbook has drastically changed. Gone from the websites are images used widely from the original Pearson Foundation. These included "many pictures of bloody fetuses in waste cans and one of a gurney carrying a woman who is apparently dead and is covered by a sheet. It ends by comparing abortion to the final solution" (Stacey, 2019).

As noted earlier, one interesting component in our representative sample set is that all of our websites are linked to larger umbrella organizations (noted near the beginning of the chapter), but it is nearly impossible to find hyperlinks to those pages on the websites. They are buried deep in the front-ends of the websites, and require multiple levels of clicks in order to find links between the organizations. We find this another example of the deceptive practice of CPC websites, because, though the site designs are very similar, they attempt to portray themselves as unique, grassroots parts of the local communities in their "service" area. Just as the sophistication of the design varies across urban and rural websites, so does the diversity of the images included. Some lean more, though not exclusively, white (typically the rural sites) and some attempt visual inclusivity (typically the urban sites).

When combined with the ways in which CPCs co-opt the language of choice from the abortion debate in the US, this image rebranding based in web design works to do three things: first, to convince women that they are advertising legitimate medical offices. Second, to engineer trust through a powerful combination of visual, alphabetic, and aural cues that claim to empower women in ways that standard medical care cannot or will not. Finally, to visually appeal to their local, embedded community by displaying a safe space for diverse, but always happy or contemplative, women.

Conclusion

Our work is meant to "contend with the overlapping rhetorical, technological, and ethical frameworks" (DeHertough, 2018) that affect how meaning is made in Crisis Pregnancy Center websites through the text, context, and user experience prompted by the design of these spaces. The overturning of Roe v. Wade in June 2022 was a major victory for a religious and conservative vocal minority. This minority has been quietly working for decades to dismantle the 1973 decision by appointing justices to the Supreme Court who would carry out such a task. Of course, this decision has rocked everyday life in the US, especially for women. Now CPCs thrive more than ever, as the uniform federal response regarding abortion – and by extension, contraception – has been so splintered. Some states have reverted back to, or newly enacted, draconian laws restricting women's reproductive rights, while others struggle to craft new laws to protect those rights. Many clinics that perform abortion procedures have and will close, leaving even fewer resources for women who need help with pregnancies, family planning, and abortion care.[4] McCann and Walker (2023) have found that "at least 61 clinics, Planned Parenthood facilities, and doctors' offices stopped offering abortions in the last year (2022)" (Section 4). This brings the total to at least 120+ closed abor-

4 As of 6 October 2022, The Guttmacher Institute reports that at least sixty-six clinics across 15 states have stopped offering abortions. Prior to the overturn of Roe, these states had a total of seventy-nine clinics that offered abortions; now, that number stands at thirteen (Kirstein et al., 2022).

tion clinics (McCann & Walker, 2023; Kirstein et al., 2022). Fewer legitimate medical clinics available to pregnant women or women with unplanned or unwanted pregnancies means that CPCs with dubious intentions and deceptive rhetoric become overly influential throughout the country.

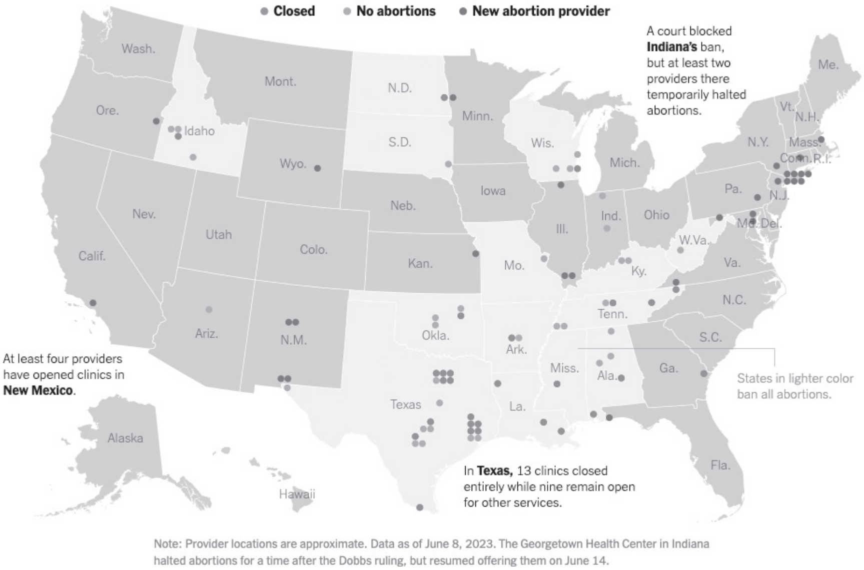

Figure 1: New York Times Map of Abortion Provider Locations (McCann & Walker, 2023)

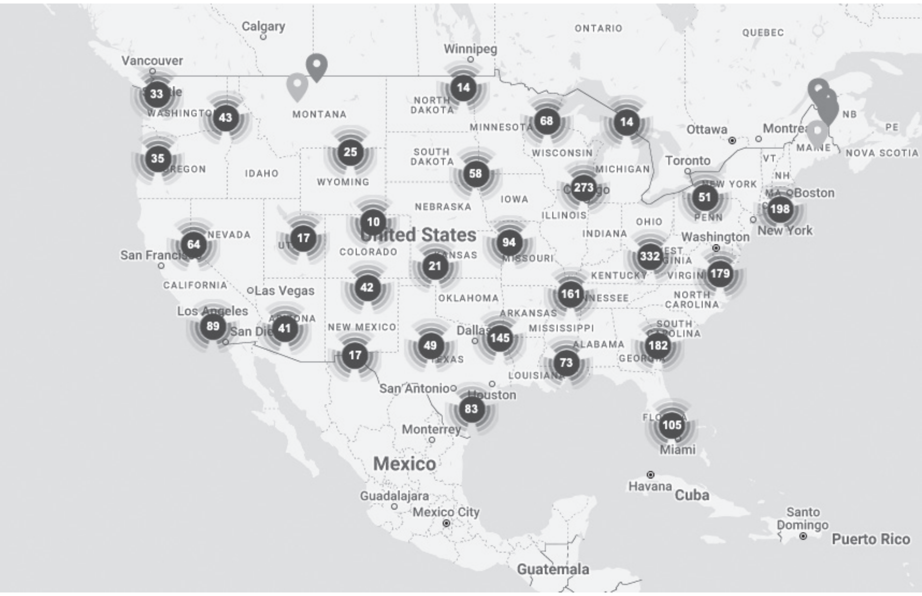

Figure 2: Crisis Pregnancy Center Map (Swartzendruber & Lambert, 2021), reproduced by consent of the authors

The relationship between the expansion of CPCs and the narrowing of medical facilities provides a stark representation of the reality that American women face in regard to abortion and abortion care.

As the CPC goal is to keep women from having abortions, their deceptive online misinformation and stall tactics can keep women from accessing it, thereby meeting that goal. Disturbingly, their tactics will likely lead to additional unplanned and unwanted children born every year, and the science is very clear on how a family fares when a child is unplanned or unwanted. Unintended pregnancies, according to the Committee on Unintended Pregnancy at the Institute of Medicine in Washington D.C., "carry appreciable risks for children, women, men, and families. That is, unintendedness itself poses an added, independent burden beyond whatever might be present because of other factors" (1995). A child resulting from unwanted pregnancy, in particular, is more likely to lack early prenatal care; the mother is "more likely to expose the fetus to harmful substances"; the child "is at greater risk" of being underweight or "dying in its first year of life, of being abused, and of not receiving sufficient resources for healthy development"; a mother of these children can be "at greater risk of physical abuse herself, and her relationship with her partner is at great risk of dissolution"; finally, the "mother and father may suffer economic hardship and fail to achieve their educational and career goals" (Committee, 1995).

These are real consequences that affect real people, consequences that vastly change the trajectory of lives for the worse. Regarding such complex issues with a myopic lens literally leads to brutal outcomes, and losing sight of the real humans affected by the simplistic view of anti-abortion groups can be cruel at best, and life-threatening at worst. For CPCs, the fetus is the most important factor in the abortion equation. These organizations will do all they can to compel the woman to carry the fetus to term. This is especially alarming because of our understanding of the arduous impediments unplanned and unwanted children and their families face.

What CPC websites engage in is deceptive advertising, which, argues Carson, "harms people" (2010, p. 182). He explains that these practices "harm consumers by causing them to have false beliefs about the nature of the products being advertised and thereby causing them to make different purchasing decisions than they would have made otherwise," and likely "purchas[ing] things unsuitable for their needs" (p. 182). Women's bodies and health, both mental and physical, as well as the health of their relationships and their families, are all affected by this deception.

The expanding mass of digital, medical disinformation from CPCs should be a concern for all who believe we should be able to differentiate between trustworthy, scientific medical information and its opposite. We must train young people and members of our communities in digital literacies, especially in discerning how digital, anti-abortion rhetoric works to misinform. We must also work toward local, state, and national regulations restricting the ability of groups (religious or otherwise) to spread medical misinformation online regarding abortion and women's healthcare. We are all patients at some point, and we should be able to access our physicians, healthcare workers, and trust that they have our best interests at heart, rather than their own personal beliefs. As noted by Brown (2018), and reiterated in this study: "It does violence to the physician-patient relationship, and the trust that it requires, when this relationship is leveraged for ideological gains" (2). The undermining of the physician-patient relationship is another arm of the war on women: women are taught that they cannot trust doctors while CPC volunteers showcase themselves as trustworthy through visual, alphabetic, aural rhetoric, and website design. American anti-intellectualism rhetorically positions itself to confront the very definition of truth. Deception has become an explicit component of political discourse in the American landscape, especially since the Trump administration's integration of "fake news" into the official discourse of the executive branch of the US government. The combination of disinformation and the unpacking of trust around authority continues to ground the contemporary war on women.

References

Crisis Pregnancy Centers

Aid to Women Center, Metro Phoenix (AWC). 2023. Website. Last modified 2023. http://aidtowomencenter.org/ultrasounds.

Care Net Pregnancy Center of Cochise County, Inc, Sierra Vista (Care Net). 2023. Website. Last modified 2023. https://www.carenetsv.com/.

Choices Women's Pregnancy Center, Waverly, RI (CWPC). N.d. Website. https://choicesrhodeisland.com/.

First Care Pregnancy Center, Minneapolis, MN (FCaPC). n.d. Website. https://firstcaremn.org/.

First Choice Pregnancy Center, Pahrump (FCPC). n.d. Website. http://www.pregnancypahrump.com/.

First Choice Pregnancy Resource Center, Stevens Point, WI (FCPRC). N.d. website. https://www.firstchoiceprc.org/.

First Choice Pregnancy Services, Las Vegas (FCPS). 2020. Website. Last modified 2020. https://www.firstchoicelv.com/services.

First Way Pregnancy Center, Phoenix (FWPC). 2023. Website. Last modified 2023. http://1stway.net/ultrasound-imaging.php.

Hands of Hope Tucson (HHT). 2021.Website. Last modified 2021. https://handsofhopetucson.com/.

Health Resources, Detroit Lakes, MN (HR). n.d. Website. https://hrcdl.org/.

Life Choices Community Pregnancy Clinic, Carson City (LCCPC). 2020. Last modified 2020. https://lifechoicescarson.com/.

Mother of Life Pregnancy Center, Providence, RI (MLPC). n.d. website. https://motheroflife.org/.

Options for Women, Mankato, MN (OW). n.d. Website. https://www.optionsmankato.org/

Phoenix Women's Clinics (PWC). 2018. Last modified 2018. https://lcwcaz.org/womens-health/.
Real Choices Women's Center, Fernley; Reno (RCWC). n.d. https://www.realchoicesnv.org/.
Women's Care Center, Madison, WI (WCC). n.d. Website. https://www.womenscare-center.org/.
Women's Resource Medical Centers of Southern Nevada, Inc., Las Vegas (WRMCSN). 2023. https://freepregtest.com/.
Women's Support Center of Milwaukee, Inc., Milwaukee, WI (WSCMIL). N.d. website. https://www.wscmil.org/.

Secondary Sources

Abortion Pill Reversal. (2023). https://abortionpillreversal.com/abortion-pill-reversal/overview.
Ahmed, A. (2015). Informed Decision Making and Abortion: Crisis Pregnancy Centers, Informed Consent, and the First Amendment. *The Journal of Law, Medicine & Ethics,* Spring 43(1), 51-58. https://doi.org/10.1111/jlme.12195.
APA Task Force on Mental Health, Abortion. Report of the APA Task Force on Mental Health and Abortion; 2008 [Washington, D.C.]. http://www.psychiatry.org/advocacy-newsroom/position-statements.
Borrero, S., S. Frietsche, & C. Dehlendorf (2019). Crisis Pregnancy Centers: Faith Centers Operating in Bad Faith. *Journal of General Internal Medicine (JGIM)*, 34(1), 144-145. https://doi.org/10.1007/s11606-018-4703-4.
Brown, T.R. (2018). Crisis at the Pregnancy Center: Regulating Pseudo-Clinics and Reclaiming Informed Consent. *Yale Journal of Law and Feminism*, 30(2), 221-274. https://openyls.law.yale.edu/handle/20.500.13051/7117.
Bryant, A. et al. (2014). Crisis Pregnancy Center Websites: Information, Misinformation and Disinformation. *Contraception,* 90, 601-605. https://doi.org/10.1016/j.contraception.2014.07.003.
Bryant, A.G., & J.J. Swartz (2018). Why Crisis Pregnancy Centers Are Legal but Unethical. *AMA Journal of Ethics*, 20(3), 269-277. https://doi.org/10.1001/journalofethics.2018.20.3.pfor1-1803.
Buchanan, R. (1985). Declaration by Design: Rhetoric, Argument, and Demonstration in Design Practice. *Design Issues*, 2(1), 4-22. http://www.jstor.org/stable/1511524
Carson, T.L. (2010). Lying and Deception: *Theory and Practice.* Oxford: Oxford UP. https://doi.org/10.1093/acprof:oso/9780199577415.001.0001.
Dehlendorff, C., L.H. Harris, & T.A. Weitz (2013). Disparities in Abortion Rates: A Public Health Approach. *American Journal of Public Health,* October, 103:10, 1772-1779. https://doi.org/10.2105/AJPH.2013.301339.
DePew, K. (2007). Through the Eyes of Researchers, Rhetors, and Audiences: Triangulating Data from the Digital Writing Situation. *Digital Writing Research: Technologies, Methodologies, and Ethical Issues,* 49-69, Cresskill (NJ): Hampton P.
Eyman, D. (2015). *Digital Rhetoric: Theory, Method, Practice.* University of Michigan Press.
Faludi, S. (1991). *Backlash: The Undeclared War Against Women.* New York: Three Rivers Press.

Gibbs, N. (2007). The Grass-Roots Abortion War. *Time*. February 15, 2007. Archived from the original on February 18, 2007.

Gilbert, K.E. (2013). Commercial Speech in Crisis: Crisis Pregnancy Center Regulations and Definitions of Commercial Speech. *Michigan Law Review,* 111(4), 591-616. https://repository.law.umich.edu/mlr/vol111/iss4/3.

Griswold, E. (2019). The New Front Line of the Anti-Abortion Movement. *The New Yorker.* November 11, 2019. https://www.newyorker.com/magazine/2019/11/18/the-new-front-line-of-the-anti-abortion-movement.

Gurak, L.J. (2018). Ethos, Trust, and the Rhetoric of Digital Writing in Scientific and Technical Discourse. *The Routledge Handbook to Digital Writing and Rhetoric,* edited by Jonathan Alexander and Jacqueline Rhodes, 124-131. New York: Routledge. https://doi.org/10.4324/9781315518497-13.

Hartig, H. (2022). About six-in-ten Americans say abortion should be legal in all or most cases. *Pew Research Center,* June 13, 2022. https://pewrsr.ch/3xtjvXq.

Hofstadter, R. (1963). *Anti-intellectualism in American Life*. New York: Alfred A. Knopf.

Hofstadter, R. (1964). *"The Paranoid Style in American Politics" and Other Essays*. Cambridge: Harvard UP.

Holtzman, B. (2017). Have Crisis Pregnancy Center Finally Met Their Match: California's Reproductive FACT Act. *Northwestern Journal of Law & Social Policy,* 12(3), 78-110. http://scholarlycommons.law.northwestern.edu/njlsp/vol12/iss3/3.

Johnson, T. (2022). A Quarter of All Abortion Clinics Would Close in US If Roe Is Overturned. Bloomberg. https://www.bloomberg.com/news/articles/2022-06-17/over-200-abortion-clinics-in-the-us-would-disappear-if-roe-v-wade-is-overturned.

Lee, I.T., & K.T. Barnhart (2023). "What is an Ectopic Pregnancy?" *JAMA Patient Page*, JAMA Network, The American Medical Association. https://doi.org/10.1001/jama.2022.22941.

Major B., M. Appelbaum, L. Beckman, M.A. Dutton, N.F. Russo, & C. West (2009). Abortion and Mental Health: Evaluating the Evidence. *Am Psychol* 64, 863–90. https://doi.org/10.1037/a0017497.

Mayo Clinic (2023). "Pregnancy Week by Week." Mayo Foundation for Medical Education and Research (MFMER). https://www.mayoclinic.org/healthy-lifestyle/pregnancy-week-by-week/in-depth/healthy-pregnancy/hlv-20049471.

Mayo Clinic (2023). "Miscarriage." Mayo Foundation for Medical Education and Research (MFMER). https://www.mayoclinic.org/diseases-conditions/pregnancy-loss-miscarriage/symptoms-causes/syc-20354298.

McAnn, A., & A.S. Walker. (2023). "One Year, 61 Clinics: How Dobbs Changed the Abortion Landscape." *New York Times*, June 22, 2023. https://www.nytimes.com/interactive/2023/06/22/us/abortion-clinics-dobbs-roe-wade. html.

McKee, H. A., & DeVoss, D. N. (Eds.). (2007). *Digital Writing Research: Technologies, Methodologies, and Ethical Issues.* New York: Hampton Press.

McKenna J., & T. Murtha (2021). "Designed to Deceive: A Study of the Crisis Pregnancy Center Industry in Nine States." The Alliance, State Advocates for Women's Rights & Gender Equality, 1-70 https://alliancestateadvocates.org/wp-content/uploads/sites/107/Alliance-CPC-Study-Designed-to-Deceive.pdf..

Melich, T. (1996). *The Republican War Against Women: An Insider's Report from Behind the Lines.* New York: Bantam Books.

Mencimer, S. (2020). The Trump Administration Is Giving Family Planning Funds to a Network of Anti-Abortion Clinics. *Mother Jones.* https://www.motherjones.com/

politics/2019/11/the-trump-administration-is-giving-family-planning-funds-to-a-network-of-anti-abortion-clinics-obria/.
Mohr, J.C. (1978). *Abortion in America: the Origins and Evolution of National Policy.* Oxford University Press.
Morrison, J. (2019). Resuscitating the Black Body: Reproductive Justice as Resistance to the State's Property Interest in Black Women's Reproductive Capacity. *Yale Journal of Law and Feminism* 31, 35–56. https://openyls.law.yale.edu/bitstream/handle/20.500.13051/7124/Resuscitating_the_Black_Body_Reproductive_Justice.pdf?sequence=2&isAllowed=y.
Napili, A. (2021). "Title X Family Planning Program: 2019 Final Rule." Congressional Research Service. https://crsreports.congress.gov/product/pdf/IF/IF11142.
Park, H. W., & M. Thelwall. (2003). Hyperlink Analyses of the World Wide Web: A Review. *Journal of computer-mediated communication, 8*(4), JCMC843.
Petchesky, R.P. (1987). Fetal Images: The Power of Visual Culture in the Politics of Reproduction. *Feminist Studies,* 13(2), 263-292. https://doi.org/10.2307/3177802.
Rubin, R. (2018). At 'crisis pregnancy centers,' critics say, ideology trumps evidence. *The Journal of the American Medical Association,* 320(5), August 7, 2018, 425-427. https://doi.org/10.1001/jama.2018.8661.
Stacey, D. (2019). "The Pregnancy Center Movement: History of Crisis Pregnancy Centers." *Mother Jones*. https://www.motherjones.com/files/cpchistory2.pdf.
Steinberg, J.R., & L.B. Finer (2011). Examining the association of abortion history and current mental health: a reanalysis of the National Comorbidity Survey using a common-risk-factors model. *Soc Sci Med,* 72, 72–82. https://doi.org/10.1016/j.socscimed.2010.10.006.
Swartzendruber, A., & D. Lambert (2021). Crisis Pregnancy Center Map. https:/crisispregnancycentermap.com. Published August, 2021.
Swartzendruber, A., & D.N. Lambert (2020). "A Web-Based Geolocated Directory of Crisis Pregnancy Centers (CPCs) in the United States: Description of CPC Map Methods and Design Features and Analysis of Baseline Data." *JMIR Public Health and Surveillance,* 6(1), March 27 2020. https://doi.org/10.2196/16726.
Swartzendruber, A., A. Newton-Levinson, A.E. Feuchs, A.L. Phillips, J. Hickey, & R.J. Steiner (2018). Sexual and reproductive health services and related health information on pregnancy resource center websites: a statewide content analysis. *Women's Health Issues*, 28(1), 14-20. https://doi.org/10.1016/j.whi.2017.10.007.
Taylor, J.S. (2008). *The Fetal Sonogram: Technology, Consumption, and the Politics of Reproduction.* New Brunswick: Rutgers UP.
United States Congress. (2021). Title X Family Planning Program: 2019 Final Rule. *Congressional Research Services Report,* November 10, 2021. https://crsreports.congress.gov/product/pdf/IF/IF11142.
VanKooten, C. (2016). Methodologies and Methods for Research in Digital Rhetoric. *Enculturation: A Journal of Rhetoric, Writing, and Culture, 17.*
Wagner, J.M. (2019). Ta(l)king Women out of the Picture: How Extreme Republican Discourse Framed Women Outside Women's Issues: 2010–2012. In *Discussing Borders, Escaping Traps: Transdisciplinary and Transspatial Approaches*. Münster and New York: Waxmann.
Waxman, H. (2016). "False and Misleading Health Information Provided by Federally Funded Pregnancy Resource Centers" (PDF). *Consumer Health Sourcebook.* United

States House of Representatives Committee on Government Reform. July 16. Archived from the original (PDF) on January 18, 2018.

Zwicky, A. (2005). "Just Between Dr. Language and I." *Language Log,* University of Pennsylvania, August 7. http://itre.cis.upenn.edu/~myl/languagelog/archives/002386.html.

Trust and Climate Change in Wikipedia Discussion Pages

Anje Müller Gjesdal, Marie Chandelier, Claire Doquet, Øyvind Gjerstad, Céline Poudat

Introduction

Climate change is a subject field where scientific knowledge plays a crucial role for public debates, and where trust in such knowledge has been challenged, potentially impacting public debates and political decision-making (Jasanoff, 2012; 2017; Bogert et al., 2023). This chapter compares the expression of trust in discussion pages about climate change in French and Norwegian Wikipedia. According to the French Agency for Ecological Transition (ADEME, 2022); the French population agrees that there is scientific consensus about climate change (72%), and that climate change is caused by humans (81%). 51% state that they have experienced the consequences of climate changes "often or sometimes" (ADEME, 2022). In Norway, only 61% agree that climate change is mainly caused by human activity, something which is considerably lower than other countries surveyed (Delebekk & Flem, 2023). As such, France and Norway present differences that make it interesting to compare these two countries.

Previous research has shown that there is a correlation between trust in scientific knowledge and climate-friendly behaviour (Cologna & Siegrist, 2020), indicating the interest of investigating trust in this context. In this chapter, we analyse the expression of trust and distrust in scientific knowledge as it plays out in a digital context, illustrated by the case of discussions of climate change topics on Wikipedia, the online collaborative encyclopaedia.

We will enquire whether the differences between Norway and France is reflected in discussions related to articles about climate change on Wikipedia, a heterogeneous discourse community that presents contrasting viewpoints (Carbou & Sahut, 2019). Since its creation in 2001, Wikipedia has become one of the most widely used sources of information globally - in February 2024, it was the 7th most visited website in the world.[1] Because of its open-editing nature, allowing anyone to edit its articles, trust has always been a concern, and bias and vandalism have been challenges (see e.g. Adler, 2011). There have been several cases of misinformation, affecting trust in Wikipedia, particularly in professional, educational and academic frameworks, where reliability is paramount (e.g. Kazimi & Guliyeva, 2022). However, other studies indicate that Wikipedia has a high level of accuracy (see e.g. Giles, 2005). Several mechanisms in Wikipedia contribute to this, including two basic, complementary principles for Wikipedia editorial policy. First, articles must be written from a Neutral Point of View (NPOV), presenting different perspectives on the subject without bias, in a fair and balanced man-

1 According to Similar Web, consulted in March 2024.

ner, assigning each viewpoint proportional weight according to its recognition. Second, citing sources is a way of adding nuance to the former principle, as the article information must be verified by way of external sources, whose acceptability must meet specific criteria, including verifiability and reliability.[2] Interestingly, these principles were established quite early to regulate the expression of personal opinions and evaluate the validity of knowledge based on the credibility of the sources cited (Sahut, 2014). Nonetheless, the two principles are not always sufficient to regulate the disagreements between contributors, and to ensure knowledge reliability. The NPOV principle is an ideal which is hard to achieve and even to define, and there is a lack of consensus about its meaning. While scientific sources have been progressively prioritised, other types of sources are also used, depending on the topic of the page. Disagreements on the validity of the source are numerous, particularly on controversial topics, such as pseudo-sciences but also climate change.

In order to study the expression of trust in this context, we analyse a corpus of discussion pages, in which users discuss topics related to climate change. The corpus consists of French and Norwegian discussion pages from Wikipedia, written and published within a time frame of 20 years. We present a qualitative analysis of how participants use external references in the discussion pages in order to see how trust and distrust are expressed in this context. The analysis is based on the linguistic framework of "Representation of Other's Discourse" (*la représentation du discours autre*, Authier-Revuz, 2020). Specifically, we focus on reported speech that cites external sources in Wikipedia discussion pages related to the topic of climate change.

To enable the investigation of the expression of trust and distrust in our corpus, we will address the following research questions:

1. How is scientific knowledge presented and reformulated in the corpus?
2. Do the discussion participants express trust or distrust in the science of climate change?

The remainder of the chapter is structured as follows: Section 2 presents the background of the study, including literature on trust and climate change, as well as the theoretical framework of ROD (Representation of Other's Discourse). In 3, we present the materials and methods of the study, while section 4 presents the findings. Finally, section 5 contains the discussion, conclusion and the limitations of this study.

2 https://en.wikipedia.org/wiki/Wikipedia:Verifiability#Reliable_sources

Background and Theoretical Framework

Climate Change and Trust in Science

Climate change issues started to receive political and public attention in the early 1990s, and it seemed at the time that it would be possible to avoid an increase in the temperature of the earth, dramatic weather events and changed living conditions. However, today, global climate emissions are still increasing, and climate change cannot be avoided. In 2022, the world was 1.1 °C warmer than in pre-industrial times, and global average temperatures may rise by around 3°C or more, if emissions of greenhouse gases are not reduced (Skea et al., 2022).

The concept of trust is highly relevant for the issue of climate change, as previous research (Cologna & Siegrist, 2020) has shown that trust is essential to achieve political action to mitigate climate change and implement measures of climate change adaptation. Cologna & Siegrist (2020, p. 6) also find that trust in science is the type of trust that correlates most strongly with climate-friendly behaviours, while trust in institutions has a weaker correlation. However, although there is scientific consensus on the anthropogenic nature of climate change, this is an issue where there have been misinformation campaigns by so-called climate denialists (Gundersen et al., 2022; Bogert et al., 2023), creating a consensus gap between science knowledge and some parts of public opinion (Bogert et al., 2023). Bogert et al. (2023) analyse the relationship between trust in science and beliefs consistent with the scientific consensus on climate change, as well as correlations between media use and such beliefs. They find that a higher level of trust in science is associated with what they term pro-social beliefs (i.e. belief in consensus over anthropogenic climate change), which could also be associated with media use. Interestingly, they find that what they term user-generated media (including Wikipedia) has a more positive effect on pro-social beliefs than traditional, centralised media, indicating that digital media do not have a clear-cut status when it comes to trust. This points to the interest of investigating trust in the context of Wikipedia.

The question arises whether trust and distrust in science is associated with climate scepticism. What Capstick and Pidgeon call "epistemic scepticism" (2014) can be sub-divided in doubts or refusal regarding the existence of climate change, the role of human influence on the climate system, or the severity of the consequences, corresponding to Rahmstorf's (2004) *trend, attribution* and *impact* scepticism, respectively. These categories appear to be empirically relevant. In a study of climate change scepticism in the UK, Capstick and Pidgeon identify trend, attribution and impact scepticism among participants. Additionally, surveys in Norway and Belgium have shown that these categories are all present in citizens' views on climate change (Catellani et al., 2024). The question is why the same categories of epistemic scepticism appear across different countries. Following van Rensburg (2015), the answer might be that such views form in direct opposition to the main claims by the Intergovernmental Panel on Climate Change (IPCC), namely that there is a clear warming *trend,* that this warming can be *attributed* to anthropogenic GHG (Greenhouse gases) concentrations, and that the warming will adversely *impact* human beings and natural habitats.

Representation of Other's Discourse

To investigate how trust in science is expressed in the context of discussions about global warming, we examine how speakers cite other discourses, particularly scientific ones. To analyse this, we apply the analytical framework *représentation du discours autre* (Representation of Other's Discourse, ROD) developed by the French linguist Jacqueline Authier-Revuz (2020).

Language is equipped with tools for self-representation, which are employed in the operation of representing other's discourses. According to Authier-Revuz, language is inherently heterogeneous and dialogical. Every discourse is intrinsically influenced by others, whose presence is revealed through the process of representing other's discourse. The Other's Discourse (OD) can be represented in two ways: directly or indirectly.

Direct representation shows the Other's Discourse in its materiality; the other's words are displayed verbatim. This is exemplified in Direct Speech (DS), where inverted commas are used to assign words to a secondary speaker:

"'The concentration of CO2 is about 370 ppm,' we read. 'This is very high compared to natural conditions.'" (Changement climatique)[3].

This way of showing the other's words is also found in Autonymous Modalisation (AM), where a statement in quotation marks is inserted into the main discourse:

"The 'earth's' atmosphere contains 380 ppm of CO2[4]." (Effet de serre)

In both cases, attention is drawn to the words of the other, which are displayed directly.

The second way of representing OD is *Indirect representation*, where the content of Other's Discourses is reported independently of their verbal expression, as it is reformulated by the reporting discourse. This is accomplished through two modes: Indirect Speech (IS) and Modalisation in Second Order Assertion (MSA). Indirect Speech is illustrated as follows:

3 On lit : « la concentration en CO2 est d'environ 370 ppm. Cette valeur est fortement élevée comparée aux conditions naturelles […] ». The discussion page where the example is taken from is given in parenthesis. We use English translation in running text and the original quotes in footnotes. All translations from French and Norwegian have been done by the authors.

4 L'atmosphère de la « Terre » contient 380 ppm de CO2.

"All scientists know that the orthodox opinion is the true and only scientific opinion. There are only a handful of crooks funded by Shell and Exxon who claim the opposite[5]." (Changement climatique).

Modalisation in Second Order Assertion (MSA) is illustrated in the following quote:

"According to the Global Humanitarian Forum, global warming is already responsible for the deaths of 300,000 people a year[6]." (Réchauffement climatique).

In both cases, it involves reporting the content of another discourse without repeating the exact words.

To sum up, the Representation of Other's Discourse is achieved through devices that either directly quote (DS and AM) or rephrase the words (IS and MSA). From the reader's point of view, this either shows a discourse, or leads directly to its meaning. On the part of the speaker, it can be hypothesised that the rephrasing involves an appropriation of the other's discourse, whereas the quotation corresponds to a certain externality of the speaker with respect to this discourse.

Corpus and Methods

Corpus

The corpus consists of Wikipedia discussion pages for topics that are fundamental to the issue of climate change. In order to compare French and Norwegian Wikipedia, we created a comparable corpus of discussion pages, i.e. a corpus with a fairly similar size and identical or very similar topics. We started from a set of articles on key topics in the climate change debate, and retained the discussion pages that generated a comparable volume of discussions in both languages. Further, we kept articles that expressed different viewpoints and preferably conflicting ones, in the form of disagreements, or even conflicts between the contributors.

Initially, the following articles were selected for analysis: 'Denial of global warming', 'Greenhouse effect', 'Issues of global warming', 'Greenhouse gases', 'Hockey stick graph', 'Intergovernmental Panel on Climate Change' (IPCC), 'Tipping Points in the Climate System', 'Climate Policy', 'Kyoto Protocol', 'Global Warming', and 'Climate Sensitivity'. Two Wikipedia discussion pages were excluded from the analysis as they turned out to have no equivalent in the other language ('Klimapådriv' ('climate forcing') in Norwegian and 'Controverse sur le réchauffement climatique'), and 'Climate Sensitivity' was excluded as it generat-

5 Tous les scientifiques savent que l'opinion orthodoxe est la vraie et seule scientifique. Il n'y a qu'une poignée d'escrots financés par shell et exxon qui prétendent le contraire.

6 Selon le Forum humanitaire mondial, le réchauffement climatique est déjà responsable de la mort de 300 000 personnes par an.

ed little discussion in both languages. In most cases, there was more discussion in one language. Thus, the pages 'Denial of global warming', 'Issues of global warming', 'Tipping points in the climate system', and 'Climate policy' generated much more discussions in Norwegian than in French. The opposite is true of the 'Hockey Stick Graph' and 'Kyoto Protocol' pages, which are much more developed in French than in Norwegian. As a general observation, and for all the articles, climate change issues seem to generate more conflictual discussions in Norwegian than in French, although in terms of volume, the French discussion pages are more extensive - which is not surprising, given the different sizes of the French-speaking and Norwegian communities. The descriptive statistics of the corpus are summarized in in Table 1.

Table 1: The descriptive statistics of the Corpus description

French	**No. of pages[7], discussion threads, messages**	**No. of words**	**Norwegian**	**No. of pages, discussion threads, messages**	**No. of words**
'Réchauffement climatique'	3, 123, 734	69400	'Global oppvarming'	4, 70, 358	29700
'Groupe d'experts intergouvernemental sur l'évolution du climat'	1, 15, 40	3977	'FNs klimapanel'	2, 18, 78	6497
'Effet de serre'	2, 70, 105	12421	'Drivhuseffekt'	5, 9, 38	3117
'Gaz à effet de serre'	1, 11, 26	4885	'Klimagass'	1, 2, 10	628
			'Klimaskepsis'	1, 36, 93	12102
Total	**7, 219, 905**	**90683**	**Total**	**14, 136, 577**	**52044**

Methods

In order to investigate the expression of trust in scientific sources, we analyse reported speech referring to sources that are external to Wikipedia and the participants in the discussion pages. We use an annotation model for the Representation of Other's Discourse (ROD) based on Authier-Revuz (2020), which has been developed as part of a ROD research project, carried out in dialogue with French specialists (Poudat, Chandelier & de Lucca, 2023). The annotation model is based on the analytical framework of ROD, presented above (section 2). The model distinguishes between five modes of representing other's discourse (ROD) (Authier-Revuz, 2020), four of which are present in our corpus: Direct Speech (DS), Indirect Speech (IS), Autonymous Modalisation (AM) and Modalisation in Second Order Assertion (MSA). In addition, we have annotated the sources cited, using

7 Current and archived pages.

the following five categories: ‘scientific’, ‘journalistic’, ‘political’, ‘other explicit’, ‘non explicit’. The first category, scientific sources, includes scientific institutions such as the IPCC as well as individual researchers. The second category, ‘journalistic’, refers to sources from national and international media. The category ‘political’ refers to politicians or political parties. The category ‘non explicit’ refers to cases where no explicit source is given, as in the following example: “certain groups twist and turn reality”[8]. Finally, the category ‘other explicit’ refers to sources that are identified, but that do not belong to the scientific, journalistic or political categories, as in the following example: “This statement has been taken from a WWF report (which has not been peer-reviewed) written by, among others, Andy Rowell, an anti-smoker and food security campaign leader who has worked for the WWF and Greenpeace[9]”. In a last step, the instances of reported speech were analysed with regards to the speakers attitude to the contents of the reported utterance, on a scale from -2 (very negative evaluation) to +2 (very positive evaluation), 0 being neutral. We take a positive evaluation to be indicative of trust, and a negative evaluation to be indicative of distrust.

Findings

Type of Other’s Discourse

Distribution

The analysis showed differences between the French and Norwegian corpora when it comes to the distribution of the modes of representation of other’s discourse. Overall, the most frequent modes for both corpora are Indirect speech (IS) and Direct speech (DS). IS is more frequent in Norwegian, while IS and DS are almost equally frequent in French. Autonymous Modalisation (AM) and Modalisation in Second Order Assertion (MSA) are significantly less used in both corpora. These results point to a tendency in the French corpora towards direct quoting, which is present in the DS and AM modes.

---

8 “visse grupper vrir og vender på virkeligheten” (Drivhuseffekt).

9 “Denne påstanden er hentet fra en WWF-rapport (som ikke er fagfellevurdert) skrevet av blant annet Andy Rowell en anti-røyker og matsikkerhetskampanjeleder som har arbeidet for WWF og Greenpeace” (FNs klimapanel).

Figure 1 presents the distribution of the different modes of representation of other's discourse (OD) in the two corpora.

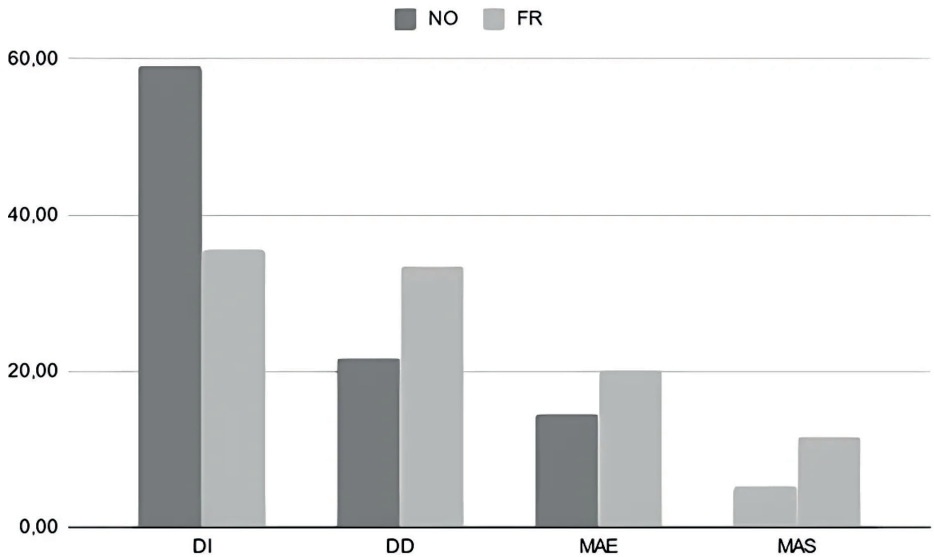

Figure 1: Distribution of modes of OD in the French and Norwegian corpora

In some cases, different modes of discourse can interact. For instance, an autonymous modalisation (AM) can occur within a stretch of indirect speech (IS), as in the following example:

"However, a mandarin from the US Academy of Sciences (only the president) claims that 'global warming' is a load of rubbish10." (Changement climatique)

This example is interesting due to the complexity of the ROD (IS + AM), but also because of the use of the categorising verb "claim" ("prétendre"), which has an axiological impact on the statement in ROD. It is an illustration of the fact that one is reporting not only a discourse, but an enunciation. The introductory verb of discourse categorises the speech act performed by the other discourse, but also gives clues about the positioning of the second speaker on the OD.

Articles and Modes

As mentioned previously, the French and Norwegian corpora are not identical in their size and composition. With this caveat, we present the most important differences in the distribution of OD forms, between topics and between languages. Tables 2 and 3 present the distributions of OD forms across topics in the two corpora. Given the differences in size and composition, we present the distributions in percentages. The articles are presented in descending order according to their size in number of words (see Table 1 for details).

10 Cependant, un mandarin de l'US Science Academy (le président seulement) prétendrait que le «global warming», c'est du pipeau.

Table 2: Distribution of forms of ROD in the French corpus, percentages.

Article	DS	IS	AM	MSA	Total
Réchauffement climatique (global warming)	41,72%	41,25%	32,22%	30,77%	38,41%
Effet de serre (greenhouse effect)	47,68%	50,00%	51,11%	59,62%	50,55%
Gaz à effet de serre (greenhouse gas)	6,62%	3,75%	5,56%	3,85%	5,08%
GIEC (IPCC)	3,97%	5,00%	11,11%	5,77%	5,96%
Total	**100,00%**	**100,00%**	**100,00%**	**100,00%**	**100,00%**

Table 3: Distribution of forms of ROD in the Norwegian corpus, percentages.

Article	DS	IS	AM	MSA	Total
Global oppvarming (global warming)	48,81%	75,53%	60,17%	21,43%	64,84%
Klimaskepsis (climate scepticism)	20,24%	11,18%	27,12%	26,19%	16,21%
FNs klimapanel (IPCC)	16,67%	7,17%	9,32%	33,33%	10,85%
Drivhuseffekt (greenhouse effect)	14,29%	5,91%	3,39%	11,90%	7,61%
Klimagass (climate gas)	0,00%	0,21%	0,00%	7,14%	0,50%
Total	**100%**	**100%**	**100%**	**100%**	**100%**

The four discussion pages in the French corpus reveal a striking particularity: half of all the OD forms are found in the 'Greenhouse Effect' page, which is the second largest page in the corpus. In comparison, the 'Global Warming' page, which is the largest page in the corpus, contains only 38% of the annotated OD forms. We also observe disparities in the Norwegian corpus, but these can be attributed to differences in the size of the pages. In the Norwegian corpus, it is the largest page, 'Global oppvarming' ('Global Warming'), which contains the highest number of OD (over 64%). We also note that this topic discussion has a high rate of IS. In Norwegian, IS is used in particular to refer to external sources to support one's own statements, this observation could be indicative of controversy and the fact that the scientific consensus is contested in this part of the corpus (Poudat, Gjesdal & Gjerstad, 2023).

Type of Source

What kind of sources do the Wikipedians cite when they discuss climate change? To what extent do they engage with scientific sources? In this section we look into the distribution of the different types of sources in the French and Norwegian corpora. Figure 2 shows the distribution of the different cited sources in the Norwegian corpus. The category of scientific sources is clearly the most frequent in the Norwegian corpus, indicating the interest and importance of this type of source for the participants in the discussions. It is followed by 'non explicit' sources, i.e. sources that are not explicitly identified but rather refer generically to persons or groups, as in the following example:

"The problem, many would say, is structural and something the governments of the world first and foremost need to address[11]" (Global oppvarming).

Further, the category 'other explicit' refers to persons or groups that are identified but who do not belong to the categories 'scientific', 'journalistic' and 'political', as in the following example:

"Instead of trusting the statements of a well-known climate sceptic like Steve McIntyre (who is not a climate scientist, who works in the coal industry and who has made it his personal mission to blindly criticise Michael Mann[12]" (Global oppvarming).

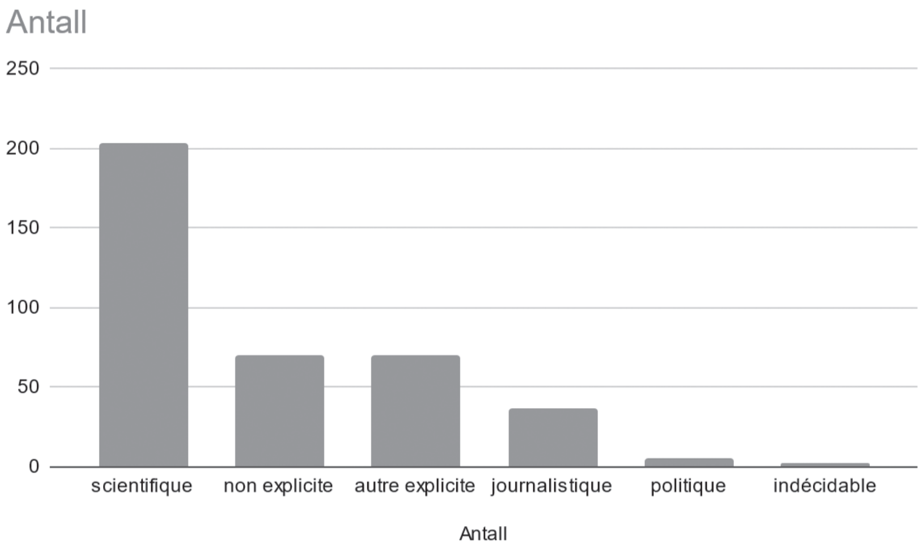

Figure 2: Distribution of sources in the Norwegian corpus

The distribution of sources in the French corpus can be found in Figure 2.

11 "Problemet, vil mange si, er strukturelt og noe verdens myndigheter først og fremst må ta tak i."
12 "I stedet for å stole på uttalelsene til en velkjent klimaskeptiker som Steve McIntyre (som ikke er en klimaforsker, som jobber i kullindustrien og som har gjort det til sitt personlige oppdrag å blindt kritisere Michael Mann)."

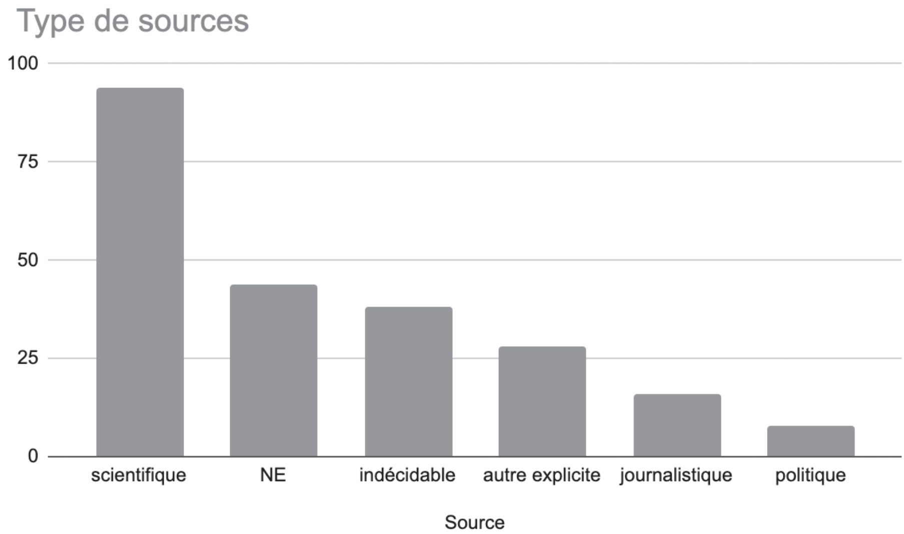

Figure 3: Distribution of sources in the French corpus

When we compare the corpora, we see clear similarities when it comes to the types of sources that are cited. As in the Norwegian corpus, mostly scientific sources are cited. To a lesser extent, the French Wikipedians also refer to external content without explicitly mentioning the source, as in the following example:

> "I've even read people claim, without batting an eyelid, that the greenhouse effect doesn't exist."[13]

Contributors use a variety of explicitly mentioned sources, such as NGOs, movies, and documentaries. Journalistic and political sources are the least quoted.

Attitudes to Sources

Regarding the attitudes towards the quoted sources, we observe that in most cases scientific sources are either positively (55%) or neutrally (35%) evaluated in the French corpus. Scientific authority is at the heart of the debate, both climate sceptics and non-sceptics quote scientists to defend their views:

> "Yes the anthropogenic cause of global warming is only a truth revealed in France but strongly contested in the United States, India, China and in the rest of the

13 "J'ai même lu des gens qui affirmaient, sans sourciller, que l'effet de serre n'existait pas" (Changement climatique).

> world. Many renowned scientists such as the director of MIT's Institute of Climatology, do not accept this thesis."[14]
>
> "Changement climatique"/archives2 page, "Climategate" thread

Climate sceptics tend to refer to scientists with indefinite formulations, as we can observe in the extract ('many renowned scientists') while contributors who defend that climate change is mostly caused by human activities tend to refer to institutions like the IPPC:

> "IPCC's 2500 scientists makeup the overwhelming majority of climatologists, glaciologists and others, and claim that the probability that global warming is manmade is over 90%."[15]
>
> "Changement climatique/Bon article" page, "votes" thread

In contrast to the French corpus, the Norwegian discussions tend to be more critical of scientific sources. As many as 20 percent of these sources are negatively evaluated, which is the same proportion as positively evaluated scientific sources. There thus appears to be a lower degree of respect for scientific authority on the issue of climate change among Norwegian Wikipedia contributors than among their French speaking counterparts. The IPCC stands out as a target of criticism among both sceptics and non-sceptics:

> "In the IPCC Working Group 1 report," it is stated,
> "There are also suggestions of increased intense tropical cyclone activity in some other regions where concerns over data quality are greater. [...] It is more likely than not (> 50%) that there has been some human contribution to the increases in hurricane intensity."
> According to Professor Les Hatton, who holds a PhD in meteorology, these statements would never have passed a peer review, as no observed data supports the statements, only model runs.[16]
>
> "FNs klimapanel"/Archive, "Antall intense orkaner øker" thread

> "The sentence says that the IPCC, by not correcting their own mistakes and thereby dismissing criticism, appears omnipotent. By not correcting their own errors (e.g.

14 "Oui la thèse de la cause humaine du réchauffement climatique n'est qu'une vérité révélée en France mais fortement contestée aux Etats-Unis, en Inde, en Chine et dans le reste du monde. De nombreux scientifiques renommés n'admettent pas cette thèse comme le directeur de l'institut de climatologie du MIT."

15 "Les 2500 scientifiques du GIEC constituent l'écrasante majorité des climatologues, glaciologues et autres, et affirment que la probabilité que le réchauffement climatique soit d'origine humaine est de plus de 90%"

16 "I IPCC arbeidsgruppe 1s rapport sies det «There are also suggestions of increased intense tropical cyclone activity in some other regions where concerns over data quality are greater. [...] It is more likely than not (> 50%) that there has been some human contribution to the increases in hurricane intensity.» Ifølge professor Les Hatton, med en PhD i meteorologi så vil disse utsagnene aldri kunne passert en fagfellevurdering, da ingen observerte data støtter opp om utsagnene, bare modellkjøringer."

> Himalaya), they appear as custodians of truth. Anyone who is not self-critical when it comes to facts believes they are managing the truth".[17]
>
> "Global oppvarming"/Archive 1, "Faglig unøyaktighet" thread

Furthermore, as in the French corpus, Norwegian climate sceptics tend to use vague formulations when citing scientific sources to support their arguments, such as 'scientists', 'tens of scientists', and 'a number of other research environments':

> "If people are weaned on IPCC material, they don't know any better. The population is divided in the middle, and there are more scientists than you would want to know who are critical of the content".[18]
>
> "Klimaskepsis"/"Kraftig omskrivning av klimaskepsis-artikkelen er nødvendig. Bærer stor preg av demonisering av klimaskeptikere" thread

Discussion

Trust in science has been shown to be correlative with climate-friendly behaviour. At the same time, climate science is sometimes contested in public debates (see e.g. Bogert et al., 2023). As a consequence, it is interesting to study how trust in science is expressed in online communication about climate change. In this chapter, we have presented a qualitative analysis of how participants in the discussion pages of the collaborative encyclopaedia Wikipedia present and reformulate scientific knowledge, and how they express trust and distrust in science as they engage in discussions about key concepts related to climate change.

We will now return to the research questions stated at the beginning of the chapter:

1. How is scientific knowledge presented and reformulated in the corpus?
2. Do the discussion participants express trust or distrust in the science of climate change?

Regarding question 1, we see that in the Norwegian corpus, external sources are mainly quoted indirectly, while in the French corpus, indirect and direct reporting are equal in size. Previous research indicates that the representation of other's discourse in the form of reported speech can also be indicative of authors' attitudes to the reported content, when scientific knowledge is reported, and that

17 "Setningen sier at IPCC gjennom å ikke rette opp i egne feil og på den måten avfeie kritikk framstår som allmektige. Ved å ikke rette egne feil (f.eks Himalaya), framstår de som forvaltere av sannhet. Enhver som ikke er selvkritiske når det kommer til kjensgjerninger tror de forvalter sannhet."

18 "Er man flasket opp med IPCC-stoff, vet man ikke bedre. Befolkningen er delt på midten, og det er flere forskere enn du vil vite om som stiller seg kritisk til innholdet."

the type of OD may be indicative of the authors' level of expertise. In a study of reported discourse in student dissertations, Boch and Grossman (2002) focused on how expert discourse, whose presence is essential in the genre of 'research dissertation', is represented. They show that the distribution between IS and DS, between reformulating and showing, is dependent on the competence of the writers: the more expert the writers are, the more they reformulate. Boch and Grossman attribute this to the risky nature of reformulating expert statements (as it can reveal misunderstandings) and find that less confident authors prefer to directly quote the experts. How do these observations apply to the corpus of Wikipedia discussion pages, and how do they relate to trust in the scientific theories cited? While quoting an OD directly serves to guard against the risk of inadequate reformulation, it simultaneously creates a detachment from one's own speech. When reformulating, speakers involve themselves with their own words in the statement they construct. When quoting, they show the words of others while keeping themselves at a distance. This difference between using their own words and using the other's words can signify a willingness or reluctance to assume ownership of the discourse, reflecting varied levels of trust. Considering the contrast outlined earlier between reformulation and quotation in French and Norwegian, we noticed conflicting patterns in the discourse of Norwegian and French Wikipedians regarding the OD. The Norwegians' inclination toward reformulation may suggest a strong confidence in their statements, while the French might be less certain of what they are asserting. It would be interesting to further investigate how these levels of certainty intersect with attitudes toward climate scepticism versus belief in human responsibility for climate change. In essence, the juxtaposition of indirect discourses and direct discourses remains relevant.

When we look further into the scientific sources that are cited, we see that our findings largely echo previous research: both French and Norwegian climate sceptics tend to refer to vague scientific sources (e.g. 'many scientists') to support their views, as there are few concrete authoritative voices to choose from. This echoes the findings of Campion, Tessier & Bourgatte (2015, p. 185), who analyse discussion about climate change in Wikipedia and online blog comment boards in a French context. In their analysis of blogs, they find that both sides of the argument refer to scientific sources in their argumentation, but that they do so differently. In their findings, those who believe in anthropogenic climate change do not present themselves as scientific experts, but they present the discourse of "real", institutional experts. The climate sceptics, on the other hand, present themselves as capable of evaluating published research and being able to distinguish between "good" and "bad" science. However, when it comes to discussions about climate change on French Wikipedia, they find that participants have a high level of knowledge, indicating that they may, in fact, be experts. It would be interesting to investigate this further and see if this is still the case.

Regarding the second research question about trust and distrust in science, there is an interesting contrast between the two corpora. While the French discussion threads tend to signal deference to scientific sources, the Norwegian corpus contains many instances of challenges to scientific authority, among sceptics and non-sceptics alike. In both the corpora, the IPCC stands out as a frequent target of such criticism, also among some who accept the hypothesis of anthropogenic climate change, but who find that the methods and practices of the panel are not beyond reproach. This could be interpreted as a sign of distrust that can be connected to scepticism, as mentioned in section 2. Previous research (Gundersen et al., 2022; Knight, 2019; Norgaard, 2006) indicates that dependence on fossil fuels and in particular production of fossil fuels is likely to influence public opinion on climate change, particularly with regard to climate scepticism and the perception of risks associated with global warming, and this could potentially be a factor in the discussions we observe.

Conclusion

Our findings are not surprising and echo previous research that has found that debates on climate change often centre on the status of scientific facts. In the Norwegian context, Ryghaug & Skjølsvold (2016) argue that climate sceptics do not express explicit distrust of scientific facts. Instead they "mimic a scientific style of arguing by referring to so-called facts, leaning on references and links when arguing about matters of concern". Thus, there is a shared, fundamental trust in science, but a disagreement over the status of credible sources. However, Ryghaug & Skjølsvold (2016) observe a shift around 2010 where there was a move from traditional media outlets to digital formats, including blogs, but also that the scientific consensus became increasingly dominant, and conflict less visible in media coverage from that time on. This also means that the status of expert knowledge is shifting, as the new, digital text genres are not necessarily subject to moderation by peer review or editorial control (Gjesdal & Gjerstad, 2015).

In this chapter, we have presented a qualitative analysis of discussion pages on Wikipedia in order to investigate expressions of trust and distrust in science in the context of climate change topics. We will end by pointing out some limitations as well as directions for further research. Due to the limited size of the materials, the results cannot be generalised. We apply an analysis where trust and distrust is measured by way of positive and negative evaluation as a proxy measure of trust. In addition, the choice of topics may also influence the result. Although we have selected key concepts related to climate change, it is possible that a different choice of topics could have produced different results. Finally, the OD has been coded manually, which is a potential source of errors.

References

Adler, B. T., De Alfaro, L., Mola-Velasco, S. M., Rosso, P., & West, A. G. (2011). Wikipedia Vandalism Detection : Combining Natural Language, Metadata, and Reputation Features. *Proceedings of the 12th International Conference on Computatio-*

nal Linguistics and Intelligent Text Processing - Volume Part II, 277-288. https://doi.org/10.1007/978-3-642-19437-5_23

Agence de la transition écologique (ADEME). (2022). *Les représentations sociales du changement climatique. 23ème v*ague du baromètre. https://librairie.ademe.fr/changement-climatique-et-energie/5917-representations-sociales-du-changement-climatique-23-eme-vague-du-barometre.html

Authier-Revuz, J. (1996). Remarques sur la catégorie de 'l'îlot textuel'. *Cahiers du français contemporain,* 3, 91-115.

Authier-Revuz, J. (2020). *La représentation du discours autre: principes pour une description.* De Gruyter. https://doi.org/10.1515/9783110641226

Boch, F., & Grossman, F. (2002). Se référer au discours d'autrui. Quelques éléments de comparaison entre experts et néophytes. *Enjeux,* 54, 41-51.

Bogert, J. M., Buczny, J., Harvey, J. A., & Ellers, J. (2023). The Effect of Trust in Science and Media Use on Public Belief in Anthropogenic Climate Change: A Metaanalysis. *Environmental Communication,* https://doi.org/10.1080/17524032.2023.2280749

Campion, B., Tessier, L., & Bourgatte, M. (2015). Interactions en ligne sur les changements climatiques. Dynamiques d'échanges et affordances des dispositifs. *Hermès,* 3, 181-190. https://doi.org/10.3917/herm.073.0181

Catellani, A., Gjerstad, Ø., & Fløttum, K. (2024). Opinions sur l'action climatique en Belgique et en Norvège. Une étude comparative. Manuscript in preparation.

Capstick, S. B., & Pidgeon, N. F. (2014). What is climate change scepticism? Examination of the concept using a mixed methods study of the UK public. *Global Environmental Change,* 24, 389-401. *https://doi.org/10.1016/j.gloenvcha.2013.08.012*

Carbou, G., & Sahut, G. (2019). Les désaccords éditoriaux dans Wikipédia comme tensions entre régimes épistémiques, *Communication,* 36(2). https://doi.org/10.4000/communication.10788

Cologna, V., & Siegrist, M. (2020). The role of trust for climate change mitigation and adaptation behaviour: A meta-analysis. *Journal of Environmental Psychology, 69,* 101428. https://doi.org/10.1016/j.jenvp.2020.101428

Delebekk, N. F. Agence de la transition écologique (ADEME). (2022). *Les représentations sociales du changement climatique.* 23ème vague du baromètre. https://librairie.ademe.fr/changement-climatique-et-energie/5917-representations-sociales-du-changement-climatique-23-eme-vague-du-barometre.html

Delebekk, N. F., & Flem, S. S. (2023, May 9). *Faktisk.no: Hver fjerde nordmann er klimaskeptiker.* Nrk.no. https://www.nrk.no/norge/faktisk.no_-hver-fjerde-nordmann-er-klimaskeptiker-1.16403293

Giles, J. (2005). Internet encyclopaedias go head to head. *Nature, 438*, 900–901. https://doi.org/10.1038/438900a

Gjesdal, A. M., & Gjerstad, Ø. (2015). Web 2.0 et genres discursifs: l'exemple de blogs sur le changement du climat. *Synergies Pays Scandinaves,* 9, 49–61.

Gundersen, T., Alinejad, D., Branch, T. Y., Duffy, B., Hewlett, K., Holst, C., ... & Baghramian, M. (2022). A new dark age? Truth, trust, and environmental science. *Annual Review of Environment and Resources, 47,* 5-29. https://doi.org/10.1146/annurev-environ-120920-015909

Jasanoff, S. (2017). Back from the brink: Truth and trust in the public sphere. *Issues in Science and Technology, 33*(4), 25-28.

Jasanoff, S. (2012). *Science and public reason.* Routledge. https://doi.org/10.4324/9780203113820

Kazimi, P. F. O., & Guliyeva, N. A. G. (2022). The concept of reliable information on the global network in times of crisis. *Technium Soc. Sci. J., 30,* 742.

Knight, K. W. 2018. Does fossil fuel dependence influence public awareness and perception of climate change? A cross-national investigation. *International Journal of Sociology, 48*(4), 295-313.

Norgaard, K. M. 2006. "We don'treally want to know" environmental justice and socially organized denial of global warming in Norway. *Organization & Environment, 19*(3), 347-370.

Poudat, C., Chandelier, M., & de Lucca, G. (2023). Un modèle pour décrire et annoter les discours autres. *Actes des 11èmes Journées de la Linguistique de Corpus,* 3-7 juillet 2023, Grenoble, 275-280.

Poudat, C., Gjesdal, A. M., & Gjerstad, Ø. (2023). Controverses du changement climatique: la représentation des paroles d'autrui dans les pages de discussion sur Wikipédia francophone et norvégien. *Bergen Language and Linguistics Studies, 13*(1). https://doi.org/10.15845/bells.v13i1.4013

Rahmstorf, S. (2004). *The climate sceptics Weather catastrophes and climate change: Is there still hope for us?* Munich, Germany: Munich Re Group.

Ryghaug, M., & Moe Skjølsvold, T. (2016). Climate change communication in Norway. In von Storch, Hans (Ed.) *Oxford Research Encyclopedia of Climate Science.* New York: Oxford University Press. https://doi.org/10.1093/acrefore/9780190228620.013.453

Skea, J., Shukla, P. R., Reisinger, A., Slade, R., Pathak, M., Al Khourdajie, A., van Diemen, R., Abdulla, A., Akimoto , K., Babiker, M., Bai, Q., Bashmakov, I. A., Bataille, C., Berndes, G., Blanco, G., Blok, K., Bustamante, M., Byers, E., Cabeza, L. F., ... Winkler, H. (2022). Summary for Policymakers. In Climate Change *2022: Mitigation of Climate Change: Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press. https://doi.org/10.1017/9781009157926.001

Van Rensburg, W. (2015). Climate Change Scepticism: A Conceptual Re-Evaluation. *SAGE Open, 5*(2). https://doi.org/10.1177/2158244015579723

Framing Distrust: TikTok as a Discursive Arena for Rivalry Between the US and the People's Republic of China

Marie Støren Jareid

Introduction

The Sino-American relationship stands as arguably the most crucial bilateral dynamic globally, given that the relationship between the superpowers determines the lives and well-being of a substantial portion of the world's population. In the wider context of recent years' trade-war, technology-war, mutual sanctions and blacklisting of companies, scholars have argued that a fundamental *trust deficit* has emerged between the two, fueled by anxiety, uncertainty, and fear (Foster, 2021; Liu, 2022; Wong, 2023). While some argue for the potential of harmonious co-existence between the two, others contend that the continuous mutual distrust may trap the two in in a series of escalating provocations that has the potential to spiral out of control (Foster, 2021; Jaworsky & Qiaoan, 2021). This deteriorating relationship has increasingly manifested itself in a fight for discursive power, both as bids for soft power, and as provoking statements towards each other (Jaworsky & Qiaoan, 2021; P. S. Lee, 2016; Wong, 2023). Thus, the trajectory of this relationship is contingent upon the perspectives that leaders in both nations hold toward one another, and how this is framed.

Scholars have raised questions regarding this persistent threat perception and how narratives of distrust disables the public from fairly evaluating each other (Lehman-Ludwig et al., 2023; Ooi & D'Arcangelis, 2017; Zaidi & Saud, 2020). These concerns have only become more relevant as the narrative contestation embedded in the discursive power competition plays out both in mainstream and on social media (Jaworsky & Qiaoan, 2021). Influential social media sites have increasingly turned into discursive battlefields where narratives are either promoted, amplified, countered, or banned. For instance, American sites X (previously Twitter) and Facebook continuously remove massive networks of accounts disseminating and amplifying false narratives (Harold et al., 2021; Twitter Safety, 2019, 2020). Furthermore, discursive contestation is progressively permeating domains traditionally considered apolitical. A notable instance is TikTok, a Chinese-origin app that has swiftly risen to become one of the world's largest social media platforms. This case is particularly interesting not only due to its exponential growth but because TikTok has emerged as a significant place for political engagement all the while platform spokespersons have asserted that it is not a „go-to place for politics" (Baker-White, 2022; Bandy & Diakopoulos, 2020; Serrano et al., 2020). Thus, this chapter asks the question of how the narrative contestation between the US and China plays out in the social sphere of TikTok. By lending inspiration from both strategic narrative and framing theory, it looks at frames used by TikTok creators. Hashtags are valuable frameworks for understanding social media culture, and this chapter explores the most popular hashtags pertaining to US-China relations through qualitative content analysis. It aims to shed light on how the collective frames of strategic government narratives of mutual distrust are, through frame alignment, transferred to the playful interface of TikTok.

The Rise of Political Engagement on TikTok

The rise of user-generated content driven platforms has diversified content creation and has transformed the role of the audience from passive receivers of information to active participants or creators of media content (Fung & Hu, 2022). In this multifaceted social landscape, TikTok emerged on the international scene in 2017, and has grown exponentially in popularity to become one of the world's most popular platforms (Zhao, 2021). The platform introduced a unique audiovisual meme-culture, which provides ways to express one's creative, professional, social, and political identities in arguably more playful ways than on other platforms. This has allowed the platform to quickly grow from a site of lip-syncing and dancing to becoming an important source of information and platform to reach out to young adults (Bandy & Diakopoulos, 2020; Serrano et al., 2020). Despite a number of issues, including the repeated spread of misinformation, studies have shown that one in three young adults consults TikTok for health advice before consulting a doctor (Basch et al., 2021; Gordon, 2022).

In the realm of declining trust in mainstream media and politicians, social media has become a place for people to seek genuine opinions and advice, particularly among younger generations (Enli & Rosenberg, 2018). Political engagement on TikTok is increasingly explored by scholars as numerous social movements and political debates have started or gained traction through the platform (Bandy & Diakopoulos, 2020; Basch et al., 2021; Literat & Kligler-Vilenchik, 2023; Schaffar & Praphakorn, 2021). Political engagement here is found to significantly differ from dominant ideals of political expression as rational, serious, and detached (Literat & Kligler-Vilenchik, 2023). It is instead expressed in humorous, over-the top or cynical ways, infused with references from popular culture. It is found to be deeply emotional, spanning the spectrum from hearty laughter to genuine tears, expressing young people‘s experiences and worldviews through content creation, commenting, liking, and hashtags. Political engagement typically revolves around intergenerational tensions and opposition to current dominant political and economic paradigms, expressed through climate activism, support for LGBTQ rights and anti-racism, and discrediting of biased media reports (Bandy & Diakopoulos, 2020; J. Lee & Abidin, 2023; Literat & Kligler-Vilenchik, 2023; Serrano et al., 2020; Vijay & Gekker, 2021; Weimann & Masri, 2020). TikTok is thereby understood as a new public arena for civic discourse, redefining contemporary political communication on a global scale (Literat & Kligler-Vilenchik, 2023; Serrano et al., 2020).

The Limits and Possibilities of TikTok

TikTok's guidelines state that content that attacks a person or group because of attributes such as ethnicity, national origin, or race is strictly forbidden on the app; however, content displaying racism, xenophobia, and promotions of violence is regularly found (Weimann & Masri, 2020). This may be aided by TikTok's technical features which facilitate a large range of content creation and dissemination, including mechanisms that effectively conceal the identity of the disseminator. This is attributed to the fact that anyone can create an account and portray themselves as someone or something else. Users may select any image as their profile picture and use any name. They can generate original content or share clips from the news, the internet, or other users. The plethora of available audiovisual effects enables the creation of content that can range from casual to professional, with or without disclosing the creator‘s presence. Through autogenerated subtitles, text-to-voice capabilities featuring perfect English accents, and green screens that create any desired background, users can disseminate any story.

Furthermore, while TikTok permits personal political opinions to be expressed and shared on the platform, the dissemination of sponsored political content is strictly prohibited. Opinion leaders, or influencers, are allowed to share their individual viewpoints, but they cannot receive financial incentives or gifts to endorse or oppose political parties, leaders, or organizations (TikTok.com, 2022). However, it remains the responsibility of the individual user to appropriately label their content as sponsored. Moreover, TikTok's labelling of state-affiliated accounts and validation system is not consistent or complete (TikTok.com, 2023). Consequently, messages can be conveyed with a consistent level of authority, credibility, or perceived authenticity, irrespective of whether the disseminator is an individual, a government entity, political organization, or business, regardless of their underlying motivations.

Theoretical Reflections: Framing Distrust on Social Media

Understanding international events through the lens of narratives entails the view that we as humans navigate through life by making sense of the world around us. These narratives present socially constructed realities that make people make sense of and situate themselves in the world (Deverell et al., 2021). The constructions become powerful when they are linked to larger narratives, making them appealing because they position states, people, and the individual in relation to the Other (Deverell et al., 2021; Schmitt, 2018). Thus, words have the power to shape representations and interpretations of reality, often referred to as the 'framing' of events (Chong & Druckman, 2007; Entman, 1993; Goffman, 1974; Moy et al., 2016; Powell et al., 2015; D. Snow et al., 2013; D. A. Snow et al., 1986). Framing consists of a selection of ideas, symbols, values, and interpretations that are organized in a coherent manner and can draw from real or fictional events, experiences, or social conditions (Entman, 1993; Schmitt, 2018; D. A. Snow et al., 1986). Frames can be understood as interpretive packages or schema for what and how to communicate the most salient aspects of the events, in order for the audience to think about the event in a certain way. According to Entman (1993), there are four dimensions of narratives in which framing can occur, namely by defining a problem, causally interpreting it, passing moral judgement on it, and/or recommending a solution for the problem (Entman, 1993). The interface of TikTok is well-suited for all these dimensions, as each video necessarily presents a complete narrative independent of the previous or next video. Each piece of content defines a topic or problem, interprets this issue, and presents a solution or punchline with or without a moral judgement passed.

Frames always are defined in relation to an event, issue, or actor, and can be distinguished between issue-specific and generic, where the latter is used in this context as they are broadly applicable on several topics across issues (de Vreese et al., 2001). Some common generic frames are "human impact," "moral values," "conflict", "responsibility", and "identity" (Chong & Druckman, 2007; de Vreese et al., 2001). Although describing frames in more generic terms leads to less detailed examinations of specific events, it allows for comparisons between frames, topics, and framing practices. This approach can be useful for understanding and analyzing expressions of distrust on a platform like TikTok, where people from many different countries and cultural contexts disseminate content.

Strategic Narratives and Frames

Governments may engage in strategic framing to shape public perceptions, garner support for their policies, or legitimize their actions. In the context of international relations, this is often referred to as the creation of strategic narratives, defined as the process of communicating a story of a series of events which is crafted by political actors with the intention to influence an audience (Schmitt, 2018). In order to influence individuals' opinions, bridging of the potential gap between individual interpretations or frames of reality, and the collective (in this context the government's) interpretations, is crucial. This can be understood through the process of frame alignment, conceptualized by Snow et al (D. A. Snow et al., 1986). This framework was developed for understanding the alignment between individuals' perceptions or reality and collective action frames (D. A. Snow et al., 1986), similar to aligning political myths to strategic narratives. Political myths, which are ingrained cultural beliefs and symbols, give strength to strategic narratives when they align with myths that tap into deeply rooted collective values, aspirations, or fears. In this way, strategic narratives can be adopted in different geographical locations or among different groups in a society in order to camouflage real political differences. For political actors, it can be strategically useful to create aligned meta-narratives on general themes, as it can be in the interests of both the group crafting the strategic narrative, which could be an external country and local politicians engaged in a national competition for power (Schmitt, 2018).

Applying this theoretical framework to the rapidly developing and changing platform of TikTok presents both advantages and challenges. One notable advantage lies in the adaptability of frame analysis, allowing for the identification and interpretation of recurring thematic frames, facilitating the understanding of how meaning is constructed and conveyed within the short-form video format. Challenges emerge simultaneously, as content spans a diverse array of genres from various sources, which may make comparisons challenging. Furthermore, the rapid proliferation and changes in popularity of content may make the analysis unpredictable and quickly seem outdated. However, it is understood as a useful tool for analyzing TikTok content in the context of this chapter, as an approach to discerning semantic and implicit messages and constructs.

Trust, Ingroups and Outgroups

A defining concept in the relationship between individuals or groups is trust. Intergroup trust, the trust between groups, which in this case pertains to nation states, can be defined as the expectation that *outgroup* members possess integrity and benign motives (Mayer et al., 1995). Trust between nations is often a complex phenomenon because of the inherently competitive international environment, and is here understood as a construct reflecting the belief and confidence that actors in the international system have in each other's intentions, reliability, and ability to fulfill commitments (Song, 2015). The antithesis to trust, distrust, conversely understood as a belief that the outgroup does not have benign motives and does not possess integrity and thereby is seen as a threat to the in-group. Expressions of distrust are here understood as either implicit or explicit references to perceptions of threat, e.g., insinuating that the outgroup has malign motives or other forms negative framing of the outgroup. The next section delves into some key tropes and frames characterizing the Sino-American relationship of distrust.

Narratives of Distrust in Sino-American Relations

American administrations have over time framed China as a challenger, often placed in the role of an untrustworthy actor on the world stage (Gray, 2021; Song, 2015). Although this temporarily subsided during the Deng Xiaoping era of market reform and internationalization, the financial crisis of 2008 pushed the US-Chinese relationship away from the modus of cooperation towards a growing narrative of systemic rivalry. This wedge grew only bigger by the Xi Jinping and Trump administrations (Foster, 2021; Jaworsky & Qiaoan, 2021). Furthermore, the Trump administration framed China not only as a challenger to American power, influence, and interests, but claiming that it was actively working to erode American security and prosperity (Ooi & D'Arcangelis, 2017).

Currently, the official narrative of the US Department of State asserts its commitment to countering and confronting what it characterizes as China's *abusive*, *aggressive*, and *coercive* behaviors. This includes violations of *fundamental rights* and *freedoms*, as well as countering China's perceived *malign activities* (United States Department of State, 2021). Furthermore, it stresses the importance of sustaining key military advantages and keeping a line of *strategic competition* towards China. The *China threat theory* implies this development and assumes that rather than rising peacefully, China will actively seek to subvert the US, its allies, and the current world order (Liu, 2022; Song, 2015). Scholars have also observed a persistent othering of China, as unrelatable, as disrespecting of international law, as *sneaky*, *cheating,* and *cunning* (Ooi & D'Arcangelis, 2017). The framing reflects the image of a country that *cannot be trusted*, that may act *irrationally* and will not follow *established rules*.

The Chinese administration, on the other hand, has focused large efforts on countering US narratives. Through government statements, state-media, and external communication efforts, its administration is claiming that the US is attempting to *contain, suppress* and *hold China down*, *interfere* in its internal affairs, and *smear* its domestic and foreign policies (Ministry of Foreign Affairs of The People's Republic of China, 2022a).

The Chinese government's emphasis of countering criticism is clearly manifested through the Ministry of Foreign Affairs' websites, where lengthy *debunking* sites and *reality checks* are published (Ministry of Foreign Affairs of The People's Republic of China, 2020, 2022a, 2022b). There is a strong emphasis on the Chinese innocence, while criticism and claims, like human rights violations are framed as *cooked up* by *anti*-China scholars in the *lie of the century.* Furthermore, explicit statements frame the US as *evil* and a *liar*. This is accompanied by statements claiming that the US has *blatantly violated* the UN Charter of Human Rights, and naming the US the *most aggressive* country in the history of the world which brings *huge turmoil and disaster* to the world. Meanwhile, the Chinese self-portrayal is as a *safeguard for international peace and security* (Ministry of Foreign Affairs of The People's Republic of China, 2020, 2022a).

President Xi has emphasized the need to *tell China's story well* and create a favorable national image in order to develop international discourse power (Chen, 2022). This entails narratives about China emerging from a *century of humiliation* caused by the Western powers and Japan, and is now a strengthened nation that will no longer be *bullied* by others, and especially not the US (Dams, 2019; Jaworsky & Qiaoan, 2021; Lehman-Ludwig et al., 2023). Scholars have argued that such competing strategic government narratives increasingly have become intertwined with collective memories and national identities (Jaworsky & Qiaoan, 2021). Thus, they are likely to contribute to shaping public perceptions and influencing social discussions surrounding the Sino-American relationship. The next section outlines the methodology for scrutinizing the ways in which the strategic narratives of distrust are transferred to TikTok content by analyzing the most popular hashtags pertaining to the topic.

Methods: Hashtags Facilitating Circulation of Meaning

In order to investigate how political distrust between the superpowers plays out on TikTok, content filed under the most popular hashtags pertaining to the Sino-American relationship, namely #USChina, #USChinarelations, and #ChinaUSA were chosen as analytical objects. Sounds, images, and text constitute the content which construct and circulate meaning on social platforms, while hashtags contribute to facilitating the connection that amplifies and proliferates this meaning (Huang & Wang, 2019). These connections often link seemingly disparate topics into chains that constitute complex networks of interconnected contexts, messages, and meanings. The choice was based on previous research on TikTok and other social media platforms where topics, hashtags, and accounts have been analyzed (Bandy & Diakopoulos, 2020; Han, 2015; Huang & Wang, 2019; Miltsov, 2022; Vijay & Gekker, 2021).

However, TikTok can present a methodological challenge to research due to its opaque structure. It is not organized in groups or pages like Facebook, nor is it primarily based on hashtags and topics like X (Twitter). Furthermore, although content is ranked according to popularity, search results may be algorithmically personalized. This indicates that on platforms like TikTok, it is challenging to 're-move' oneself as a researcher from the data extracted (Southerton & Clark, 2022). Hashtags also serve different purposes and can be used unrelated to the content only for the purpose of drawing attention to the videos, which may lead to a number of irrelevant search returns (Southerton & Clark, 2022). However, bearing this in mind, hashtags are the primary organizing device and therefore deemed appropriate for this research.

An initial familiarization of the platform was performed by exploring the platform's characteristics, giving an idea of how users can use features and templates to create content, and how the platform guides and shapes users' experiences of using TikTok (Cervi & Divon, 2023), and additionally informed by previous descriptions, walk-throughs, and analyses of the app (Kaye et al., 2021; Poh & Abidin, 2021; Ryan, F. et al., 2020). Guided by principles of abductive research, non-participant observation and qualitative content analysis was used as a methodology[1] (Denzin & Lincoln, 2011; Han, 2015; Thompson, 2022). Unobtrusive research can provide valuable information and understanding of the topic without interfering with social discourses, which is of particular importance when studying sensitive topics (Miltsov, 2022). Various tags related to the current Sino-American relationship were considered before #USChina, #ChinaUSA, and #USChinarelations were chosen.[2] The tags had approximately 50 million views combined, ranking as the most popular tags pertaining to the topic.[3]

1 Although resource constraints led to only one researcher interpreting the data, this was done several times over three months in an abductive manner where the literature was revisited in between.

2 The tags #China or #USA are in comparison a lot more popular but were excluded for containing mainly irrelevant content.

3 These were the most popular hashtags on TikTok pertaining to Sino-American relations. See

For each of the three hashtags, the 25 most popular video posts were selected and analyzed, equating to a total of 75 posts. This was deemed sufficient for thick descriptions of the content and understanding of its context, while feasible within the time and labor limitations of the research. Simultaneously, the sample was assumed to be representative and comprehensive enough as more data did not add significant breadth or depth to the study. Thus, theoretical saturation was assumed.

Similar to the characteristics of media framing, the short-video format on TikTok inherently involves the verbal or visual delineation of an issue accompanied by proposed solutions or punchlines, often intertwined with some form of moral judgments. The feature of content being ranked by popularity adds an additional layer, an indication of its resonance with the audience. This setup positions TikTok as a suitable platform for the application of the chosen theoretical frameworks. Previous research has further emphasized the importance of using a research strategy on TikTok that allows for comprehensive and immersive exploration, which qualitative methods do (Miltsov, 2022, pp. 671–672). The process involved moving back and forth between the data and the literature over time, proceeded by developing codes and extracting themes manually while immersed in the data. A codebook was made[4], not to objectively measure the data but as a guide to reflect on the interpretive connotations given to the data. It contained several characteristics, including content modality, channel description, semantic themes, and latent themes. The modality was coded as: (a) image/s (+ text/ voiceover) (b) news video clip (+voiceover/ text) (c) live video, person talking to camera (+ visual effects) (d) other live video, voiceover (+ visual effects) (Cervi & Divon, 2023; Thompson, 2022).

A code is here understood as key words or short phrase that assigns essence-capturing attributes as a way to remember parts of language-based or visual data (Thompson, 2022). The coding scheme of video content was additionally inspired by similar studies (such as Cervi & Divon, 2023; Huang & Wang, 2020) and grounded in the theoretical perspectives of strategic narratives and framing. The analysis and interpretation of qualitative data often overlap, although the analysis can be seen as the process where one seeks to explain the phenomenon, whereas the interpretation can be understood as a process that gains understanding and insight from a holistic view of the data (Leavy, 2014, pp. 464–465). After the initial data gathering, the literature was expanded multiple times which led to reconsideration of codes and refinement of the categories over time (Domenico et al., 2021).

In qualitative analysis, the numeric prevalence of codes is not the important part, but rather how they contribute to the sensemaking and themeformation. Thereby, not all codes formed themes but were still part of the codebook. The

Appendix. Non-English content (Hindi, Arabic, Spanish, and Malay contents was found) and posts not pertaining to the Sino-American relationship were excluded from further analysis.

4 See example in the Appendix. The research was conducted according to SIKT's guidelines. No information or imagery was downloaded or stored in the process, no names or personal data of creators were recorded.

posts were coded one by one, but due to multiple posts discussing the same topics, they were clustered under the same theme. Thereafter, themes are extracted, which may constitute a combination of several codes in order to explain the phenomenon theoretically. Semantic themes consist of the explicit, surface level description, while latent themes go beyond the surface, using theory to conceptually explain it (Thompson, 2022), which was important in the context of this research.

Modalities and Presentation

The in total 75 most viewed videos filed under the #USChina, #ChinaUSA, and #USChinaRelations revealed some common characteristics for both the users and the modalities chosen in the presentation. Five user accounts were accountable for a total of 45 of the videos. Three predominant modalities and characteristics emerged in the examined content. Creators engaged either in direct interaction with the audience by featuring themselves in the video while explaining the subject matter; alternatively, creators opted for a voice-over approach, articulating explanations for visuals such as news clips or webpage content without being physically present in the frame. The third prevalent modality involved creators maintaining anonymity while conveying information through text or captions and hashtags, particularly in combination with visuals like news clips or webpage screenshots.

Among the non-anonymous creators in the dataset, all but one of the most popular contributors accounts belonged to middle-aged men. This represents a clear visual departure from the average TikTok creator, as over 80% are under the age of 25 (Literat & Kligler-Vilenchik, 2023). Only three accounts featured women in any of the posts, either as narrators (one account) or on camera (two accounts). The most prominent creator, contributing to 20 videos in the dataset, was a middle-aged American man. The presentation styles remained consistent with the man facing the camera, surrounded by computer screens, delivering content in a matter-of-fact manner, avoiding unnecessary elaboration or verbosity. Alternatively, some videos featured news clips without voiceover, relying on captions and hashtags like #Chinathreat, #Chinauswar, #spies, and #Chinaconflict to convey his sentiment towards the topic.

The topics revolved around issues frequently discussed in the media, including the creator's own assessment or moral judgement about the issue. These encompassed discussions on China's perceived strategic weaknesses, the alleged use of a Chinese balloon for spying on the United States, and content mocking Chinese reactions to American naval activities in the South China Sea. While most narratives followed explicit storytelling, some exceptions employed visual means, relying on implicit judgments rather than explicit verbal expressions. For instance, a video titled "The Difference Between the United States and China" utilized a comparative visual approach. The presentation featured a split screen, depicting a tank and a soldier against a dark background with an American flag on one side, and on the other side, a group of people conversing on a bench against a bright red background symbolizing China. An accompanying text highlighted that China invested 246 billion dollars in poverty alleviation from 2012 to 2020, contrasting with the claim that the United States spent 2 trillion dollars on wars in Afghanistan and Iraq by 2020.

Out of the accounts responsible for the 45 most popular videos, only one was an official news media account labeled as *state affiliated* by TikTok. This prominent Chinese state-media channel has a significant presence on YouTube as well and contributed five videos to the dataset. Its description stated, "News on China, from China, with a non-western-media perspective". The channel presented excerpts of news and interviews with prominent individuals accompanied by explanatory texts. The content followed the official Chinese political line of focusing on China being a force of good in the world while the US is malign. Only two other accounts in the data were verified, but not labelled as state affiliated. Six of the most popular videos were posted by a self-proclaimed news channel, stating that it was a "channel for alternative news", but no information about the owner was provided. It featured news clips with headlines, descriptions, and hashtags, presenting images layered with text. Topics centered around allegations against the US, its politicians, and NATO as being evil and having bad intentions. Despite lacking verification or labels, this account garnered significant attention, with up to 62,000 likes per video. Similarly, nine videos were posted by another self-proclaimed news channel. Its description stated, "Get insight on world geopolitics". However, the content predominantly concentrated on US-China and US-Russia relations. The channel shared clips from Indian and American news channels, interspersed with content from two YouTube influencers known for their critical views of „Western" media and refusal to accept any criticism of China. An additional five videos were shared by an account owned by a middle-aged Chinese man residing in the US Although it similarly covered politicized topics such as the potential for conflict between the two nations, it distinguished itself as the sole account that did not adopt an adversarial or threat-focused perspective. Notably, among the most popular accounts, this one uniquely presented a „de-escalation" frame through bringing his own perspectives and experiences while discussing similarities and challenges both countries face.

Extraction of themes

Like the characteristics of media framing, the short-video format implies some form of verbal or visual problem delineation, with a proposed solution or punchline, and often passing of moral judgments. Additionally, content is ranked by popularity, which indicates its resonance in the audience. This setup makes TikTok well-suited for the application of such theoretical frameworks.

The semantic themes extracted from the data were the following (examples of codes in parentheses): mutual threat perception, reactiveness or aggression (the potential for US-China war, high-level visits, economics/ trade war, Taiwan relations/ war), China as threat (buying land by military bases in the US, its economic and political rise, Chinese spy balloon), US as threat (imperialist, staring wars, hegemon), US warning China (US military bases, economic/ tech strength, allies across the world), China warning US (about war, economic/ tech strength, growing number of allies), Ukraine as proxy war (China-US relation with Russia/ Ukraine/ NATO), dependency/ decoupling, government handling of Covid-19, exposing/ debunking (telling what the Western/ US/ mainstream media will not tell you), competition of moral superiority (US-China comparisons, China more national parks than US), US/ China mocking/ ridiculing the other, misconceptions about China, opinions on de-escalation and China-US similarities (increase student exchange, seeing both sides, US and China as equally good/ bad, similar challenges). The coding was done in a back-and-forth process, and in the process of understanding the data and extracting themes, the literature was regularly conferred and expanded over the course of three months.

Next, the content was categorized under latent themes. The themes were clustered in framing of outgroup, framing of ingroup, and framing of similarities. The coding and theme extraction led to the final latent themes (including sub-themes in parentheses): 1. Framing of outgroup (Othering): The other is a threat (the other/ outgroup is a threat to us, warmongering); the other is responsible for our problems (the other is the root cause of problems and injustice), the other is morally inferior (the other lies about us, the other makes poor decisions, is inferior to us); the other is less ridiculing the other (the other is irrational, sensitive, overreacting). 2. Framing of self/ ingroup: the in-group is superior (stronger, has more allies, is compelled to tell the outgroup what is right and wrong), the ingroup is morally superior (the ingroup is innocent and wrongly accused of wrongdoings, the ingroup only makes humane and ethical decisions). 3. Framing of similarities: De-escalation (the other is not that bad, both face similar challenges, both do good and bad things, need for collaboration).

Reiteration and Amplification of Official Narratives

The data revealed that popular content pertaining to the Sino-American relationship largely follows the line of official government and media framing where threat perception and deep-seated distrust permeate the official discourse. All but one identified theme revolved around *us versus them* narratives, accentuating the adversarial nature of the content. Thus, the latent frames identified depict a clear pattern of Othering and distrust where both the US and China were portrayed using negative stereotypes and simplified characterizations.

Prevalent frames included the depiction of China as a threat, untrustworthy, and prone to irrational behavior. China is portrayed as a threat to the current world order and needs to be contained. In contrast, the US is framed as a bully, a self-proclaimed rule-maker, and an unwarranted global police force. These official narratives of distrust are strategic components of the larger political battle for global discourse power between the US and China. Exemplified during the COVID-19 pandemic, the discourse in the Chinese mainstream and social media was found to quickly refocus from government failures to external issues, in particular related to the US (Jaworsky & Qiaoan, 2021). The competitive discourse about ideologies and power between China and “the West” could in this way take the focus off of the government’s wrongdoings. In this way, any disagreement between China and the US or its allies can be interpreted as attempts to *contain, suppress,* and *bully* China, framing that has been found to contribute to anti-Western attitudes in Chinese online communities (Lehman-Ludwig et al., 2023; Zhang, 2022). Meanwhile, in the US the persistent framing of China as a potential enemy, *villain*, *cheat,* and a *challenger* have enabled the justification of a continuous political hardline approach and skepticism (Ooi & D’Arcangelis, 2017).

The alignment with official narratives separates this topic from political engagement described on TikTok as playful forms of social activism, often in opposition to the dominant political paradigms (J. Lee & Abidin, 2023; Literat & Kligler-Vilenchik, 2023). It may additionally contribute to strengthening and amplifying the official narratives. In the absence of conventional information intermediaries and favoring of content that induce emotional engagement, political narratives that traditionally have been conveyed in more detached and impersonal ways (Literat & Kligler-Vilenchik, 2023) become personal, emotional, and targeted on TikTok. Thereby, previously intangible and abstract issues related to geopolitics become intimate and individualized. As a consequence, the broad narratives of distrust in official framing have the potential to become personal narrative battles between the American and Chinese.

In this landscape of adversarial narratives, the *de-escalation* frame emerged and serves as a noteworthy exception. This content highlighted commonalities between the US and China, acknowledging shared challenges and explanations of past actions, offering a more nuanced and collaborative perspective. Only one account among the most popular videos consistently took on this perspective, indicating either the lack of such content discussing Sino-American relations, or that non-adversarial perspectives are less viewed on TikTok. Further, while the findings might suggest that the viewpoints presented in the popular content resonate the most with TikTok users, it is important to note that the popularity of TikTok content does not necessarily always reflect users’ choices. Users on TikTok typically do not curate their feeds actively; instead, algorithms play a pivotal role in determining the content that viewers see. Feeds are tailored to personal preferences where algorithms predict who will resonate the most with the content (Zhao, 2021). Thus, content popularity is an interplay between users, the content, and the algorithms.

Moreover, the absence of conventional information intermediaries and clear authorship on TikTok places the burden on users to assess the credibility of an array of information independently (Goodwin et al., 2023). This setup also makes questions of authenticity and credibility of both the content and its disseminators imperative. As any user can take any username and profile picture, the creator’s agenda and identity can be disguised. In the data, self-proclaimed news channels utilized compilations of news clips or visuals, accompanied by added dramatic music and user-generated textual explanations. This facilitates circulation of content where the creators are absent, making it challenging to know who they in reality are. Thus, it may be virtually impossible to know whether an account is run by a private person displaying his or her personal opinion, a political organization, business, or agency with a certain agenda.

Conclusion

The analysis of TikTok videos pertaining to the current Sino-American relationship provides valuable insights into the framing of narratives within this social sphere. Much of the research on political engagement on TikTok have the common denominator of being in some form of opposition to the contemporary political paradigms (Bandy & Diakopoulos, 2020; Cervi & Divon, 2023). However, as this chapter has demonstrated, popular political content related to the Sino-American relationship rather takes form as the amplification and strengthening of the current narratives of distrust held by both governments.

Official political rhetoric shapes people's opinions in different ways, detracting them from the ability to make fair assessments of these relations (Ooi & D'Arcangelis, 2017). From a strategic narrative point of view, this may be seen as beneficial by both governments as they face multi-faceted internal challenges such as growing economic disparities, economic slow-down, and ethnic violence, potentially challenging the legitimacy of both countries' leadership. Thus, when the narratives surrounding a chosen out-group manages to align with myths and individuals' perception of the world through evoking fear or anger, other issues can be camouflaged. In this way, uprising and collective action against the government can potentially be avoided (Ding, 2015; Schmitt, 2018), at least for a while.

As the geopolitical contestation of narratives serves as a tool in the ongoing struggle for influence and dominance in shaping international perceptions, it is the public who bears the consequences. The reiteration of frames and narratives on the highly personalized feeds may contribute to amplifying the trust deficit created by both nations by transferring it from a geopolitical level to politically engaged social media users. The echoing of simplified images of each other indicates an amplification of an environment where the complexities of the Sino-American relationship are overlooked. While these frames may be means to serve political goals domestically, the long-term consequences of this geopolitical narrative contestation in the global domain of TikTok is unknown.

Moreover, the characteristics of TikTok, where appealing content is more important than the credibility of the creator, complicates and blurs the line between personal opinions and instructed content. Uncertainties related to the agenda and/ or identity of creators' points to a broader issue of the social media landscape where politics, advertising, news, and influencer marketing blur the lines between facts and fiction, neutral and biased information. Given TikTok's pivotal role as a global conduit for information dissemination, merely dismissing being a „go-to place for politics" holds little weight. The acknowledgment of this transformative role demands a persistent commitment from TikTok, and similar platforms, to proactively address the challenges that arise within their influential spaces. This should include consistent and broader labelling of media accounts and sponsored content, in addition to enhanced fact-checking.

The current study has several limitations that, in turn, open pathways to future research.

The focus of this chapter is exclusively directed towards the analysis of TikTok content and did not allow for an in-debt examination of the individual users responsible for sharing these videos. This decision was based on pragmatic considerations of time and space constraints inherent in the scope of this chapter. However, TikTok has made an official API that simplifies quantitative research while respecting the platform's regulations. Suggestions for future work would be an in-depth exploration of user accounts engaging in these topics. Furthermore, scholars should, through ethnographic research, aim to better understand the impact of such framing in shaping the young generation's perceptions of the world they live in.

References

Baker-White, E. (2022, December 1). *On TikTok, Chinese State Media Pushes Divisive Videos About US Politicians*. Forbes. https://www.forbes.com/sites/emilybaker-white/2022/12/01/tiktok-chinese-state-media-divisive-politics/

Bandy, J., & Diakopoulos, N. (2020). #TulsaFlop: *A Case Study of Algorithmically-Influenced Collective Action on TikTok* (arXiv, abs/2012.07716). arXiv. http://arxiv.org/abs/2012.07716

Basch, C. H., Meleo-Erwin, Z., Fera, J., Jaime, C., & Basch, C. E. (2021). A global pandemic in the time of viral memes: COVID-19 vaccine misinformation and disinformation on TikTok. *Human Vaccines & Immunotherapeutics, 17*(8), 2373–2377. https://doi.org/10.1080/21645515.2021.1894896

Cervi, L., & Divon, T. (2023). Playful Activism: Memetic Performances of Palestinian Resistance in TikTok #Challenges. *Social Media + Society, 9*(1), 20563051231157607. https://doi.org/10.1177/20563051231157607

Chen, S. (2022). Discourse Power. In *The CMP Dictionary*. https://chinamediaproject.org/CMP-Dictionary/

Chong, D., & Druckman, J. N. (2007). Framing Theory. *Annual Review of Political Science, 10*(1), 103–126.

Dams, T. (2019). *China's stories to rule the world* [China Centre Report]. https://www.clingendael.org/sites/default/files/2019-10/Clingendael_Alert_China_stories_to_rule_the_world.pdf

de Vreese, C., Peter, J., & Semetko, H. (2001). Framing Politics at the Launch of the Euro: A Cross-National Comparative Study of Frames in the News. *Political Communication, 18*(2), 107–122. https://doi.org/10.1080/105846001750322934

Denzin, N. K., & Lincoln, Y. S. (2011). *The SAGE Handbook of Qualitative Research* (illustrated ed.). SAGE 2011. https://books.google.com/books/about/The_SAGE_Handbook_of_Qualitative_Researc.html?id=AIRpMHgBYqIC

Deverell, E., Wagnsson, C., & Olsson, E.-K. (2021). Destruct, direct and suppress: Sputnik narratives on the Nordic countries. *The Journal of International Communication, 27*(1), 15–37. https://doi.org/10.1080/13216597.2020.1817122

Ding, S. (2015). Engaging Diaspora via Charm Offensive and Indigenised Communication: An Analysis of China's Diaspora Engagement Policies in the Xi Era. *Politics, 35*(3–4), 230–244. https://doi.org/10.1111/1467-9256.12087

Domenico, G. D., Sit, J., Ishizaka, A., & Nunan, D. (2021). Fake news, social media and marketing: A systematic review. *Journal of Business Research, 124,* 329–341. https://doi.org/10.1016/j.jbusres.2020.11.037

Enli, G., & Rosenberg, L. T. (2018). Trust in the Age of Social Media: Populist Politicians Seem More Authentic. *Social Media + Society, 4*(1), 2056305118764430. https://doi.org/10.1177/2056305118764430

Entman, R. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication, 43*(4), 51–58. https://doi.org/10.1111/J.1460-2466.1993.TB01304.X

Foster, J. B. (2021). Monthly Review | The New Cold War on China. *Monthly Review, 73*(3). https://doi.org/10.14452/MR-073-03-2021-07_1

Fung, A., & Hu, Y. (2022). Douyin, storytelling, and national discourse. *International Communication of Chinese Culture, 9*(3–4), 139–147. https://doi.org/10.1007/s40636-022-00259-z

Goffman, E. (1974). *An essay on the organization of experience* (Reprinted 1986). North Eastern University press, 1986.

Goodwin, A., Joseff, K., Riedl, M. J., Lukito, J., & Woolley, S. (2023). Political Relational Influencers: The Mobilization of Social Media Influencers in the Political Arena. *International Journal of Communication, 17*(19), 1613–1633.

Gordon, D. (2022, December 20). *33% Of Gen Zers Trust TikTok More Than Doctors, New Survey Shows.* Forbes. https://www.forbes.com/sites/debgordon/2022/12/20/33-of-gen-zers-trust-tiktok-more-than-doctors-new-survey-shows/

Gray, J. E. (2021). The geopolitics of 'platforms': The TikTok challenge. *Internet Policy Review, 10*(2). https://doi.org/10.14763/2021.2.1557

Han, R. (2015). Manufacturing Consent in Cyberspace: China's "Fifty-Cent Army": *Journal of Current Chinese Affairs, 44*(2), 105–134. https://doi.org/10.1177/186810261504400205

Harold, S., Beauchamp-Mustafaga, N., & Hornung, J. (2021). *Chinese Disinformation Efforts On Social Media.* RAND project Airforce Santa Monica CA. https://apps.dtic.mil/sti/citations/AD1142311

Huang, Z. A., & Wang, R. (2019). Building a Network to "Tell China Stories Well": Chinese Diplomatic Communication Strategies on Twitter. *International Journal of Communication, 13*, Article 0.

Huang, Z. A., & Wang, R. (2020). Full article: 'Panda engagement' in China's digital public diplomacy. *Asian Journal of Communication, 30*(2), 118–140. https://doi.org/10.1080/01292986.2020.1725075

Jaworsky, B. N., & Qiaoan, R. (2021). The Politics of Blaming: The Narrative Battle between China and the US over COVID-19. *Journal of Chinese Political Science, 26*(2), 295–315. https://doi.org/10.1007/s11366-020-09690-8

Kaye, D. B. V., Chen, X., & Zeng, J. (2021). The co-evolution of two Chinese mobile short video apps: Parallel platformization of Douyin and TikTok. *Mobile Media & Communication, 9*(2), 229–253. https://doi.org/10.1177/2050157920952120

Leavy, P. (Ed.). (2014). *The Oxford handbook of qualitative research.* Oxford University Press.

Lee, J., & Abidin, C. (2023). Introduction to the Special Issue of "TikTok and Social Movements." *Social Media + Society, 9*(1). https://doi.org/10.1177/20563051231157452

Lee, P. S. (2016). The rise of China and its contest for discursive power. *Global Media and China, 1*(1–2), 102–120. https://doi.org/10.1177/2059436416650549

Lehman-Ludwig, A., Burke, A., Ambler, D., & Schroeder, R. (2023). Chinese Anti-Westernism on social media. *Global Media and China, 8*(2), 119–137. https://doi.org/10.1177/20594364231166541

Literat, I., & Kligler-Vilenchik, N. (2023). TikTok as a Key Platform for Youth Political Expression: Reflecting on the Opportunities and Stakes Involved. *Social Media + Society, 9*(1). https://doi.org/10.1177/20563051231157595

Liu, K. (2022). The Global Times and The China Threat Narrative: An Empirical Analysis. *Journal of Chinese Political Science, 27*(1), 1–18. https://doi.org/10.1007/s11366-021-09754-3

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review, 20*(3), 709–734. https://doi.org/10.2307/258792

Miltsov, A. (2022). Researching TikTok: Themes, Methods, and Future Directions. In *The SAGE Handbook of Social Media Research Methods* (pp. 664–675). London SAGE Publications Ltd. https://doi.org/10.4135/9781529782943.n46

Ministry of Foreign Affairs of The People's Republic of China. (2020, July 2). *What's False and What's True on China-related Human Rights Matters.* https://www.fmprc.gov.cn/mfa_eng/zxxx_662805/t1794581.shtml

Ministry of Foreign Affairs of The People's Republic of China. (2022a, June 19). *Reality Check: Falsehoods in US Perceptions of China*. https://www.mfa.gov.cn/eng/wjbxw/202206/t20220619_10706059.html

Ministry of Foreign Affairs of The People's Republic of China. (2022b, July 5). *Fact Sheet on the National Endowment for Democracy* [Fmprc.gov.cn]. https://www.fmprc.gov.cn/eng/zxxx_662805/202205/t20220507_10683090.html

Moy, P., Tewksbury, D., & Rinke, E. M. (2016). Agenda-setting, priming, and framing. *The International Encyclopedia of Communication Theory and Philosophy,* 1–13.

Ooi, S.-M., & D'Arcangelis, G. (2017). Framing China: Discourses of othering in US news and political rhetoric. *Global Media and China, 2*(3–4), 269–283. https://doi.org/10.1177/2059436418756096

Poh, Y. T., & Abidin, C. (2021). Douyin and tiktok: A cross-language systematic review of academic scholarship on sister apps. *Paper Presented at #AoIR2021: The 22nd Annual Conference of the Association of Internet Researchers.* http://spir.aoir.org

Powell, T. E., Boomgaarden, H. G., De Swert, K., & De Vreese, C. H. (2015). A Clearer Picture: The Contribution of Visuals and Text to Framing Effects: Visual Framing Effects. *Journal of Communication, 65*(6), 997–1017. https://doi.org/10.1111/jcom.12184

Ryan, F., Impiombato, D., & Fritz, A. (2020). *TikTok and WeChat: Curating and controlling global information flows* (Policy Brief 37/2020; p. 72). Australian Strategic Policy Institute.

Schaffar, W., & Praphakorn, W. (2021). The #MilkTeaAlliance: A New Transnational Pro-Democracy Movement Against Chinese-Centered Globalization? *Austrian*

Journal of South-East Asian Studies, 14(1), Article 1. https://doi.org/10.14764/10.ASEAS-0052
Schmitt, O. (2018). When are strategic narratives effective? The shaping of political discourse through the interaction between political myths and strategic narratives. *Contemporary Security Policy, 39*(4), 487–511. https://doi.org/10.1080/13523260.2018.1448925
Serrano, J. C. M., Papakyriakopoulos, O., & Hegelich, S. (2020). Dancing to the Partisan Beat: A First Analysis of Political Communication on TikTok. *12th ACM Conference on Web Science,* 257–266. https://doi.org/10.1145/3394231.3397916
Snow, D. A., Rochford, E. B., Worden, S. K., & Benford, R. D. (1986). Frame Alignment Processes, Micromobilization, and Movement Participation. *American Sociological Review, 51*(4), 464–481. https://doi.org/10.2307/2095581
Snow, D., Tan, A., & Owens, P. (2013). Social Movements, Framing Processes, and Cultural Revitalization and Fabrication. *Mobilization: An International Quarterly, 18*(3), 225–242. https://doi.org/10.17813/maiq.18.3.2886363801703n02
Song, W. (2015). Securitization of the "China Threat" Discourse: A Poststructuralist Account. *China Review, 15*(1), 145–169.
Southerton, C., & Clark, M. (2022). OBGYNs of TikTok and the role of misinformation in diffractive knowledge production. *Journal of Sociology, 59*(3), 610–627. https://doi.org/10.1177/14407833221135209
Thompson, J. (2022). A Guide to Abductive Thematic Analysis. *The Qualitative Report.* https://doi.org/10.46743/2160-3715/2022.5340
TikTok.com. (2022, August 30). *TikTok's Stance on Political Ads.* TikTok Creator Portal. https://www.tiktok.com/creators/creator-portal/en-us/community-guidelines-and-safety/tiktoks-stance-on-political-ads/
TikTok.com. (2023, January 19). *Labelling state-affiliated media entities.* TikTok. https://www.tiktok.com/transparency/en/state-affiliated-media/
Twitter Safety. (2019, August 19). *Information operations directed at Hong Kong.* https://blog.twitter.com/en_us/topics/company/2019/information_operations_directed_at_Hong_Kong.html
Twitter Safety. (2020, June 12). *Disclosing networks of state-linked information operations we've removed.* https://blog.twitter.com/en_us/topics/company/2020/information-operations-june-2020.html
United States Department of State. (2021, May 12). *US Relations With China.* United States Department of State. https://www.state.gov/u-s-relations-with-china/
Vijay, D., & Gekker, A. (2021). Playing Politics: How Sabarimala Played Out on TikTok. *American Behavioral Scientist, 65*(5), 712–734. https://doi.org/10.1177/0002764221989769
Weimann, G., & Masri, N. (2020). Research Note: Spreading Hate on TikTok. *Studies in Conflict & Terrorism, 46*(5), 752–765. https://doi.org/10.1080/1057610X.2020.1780027
Wong, B. (2023, February 27). *The Balloon Saga Reveals A Deep Trust Deficit – And That's Worrying.* China-US Focus. https://www.chinausfocus.com/foreign-policy/the-balloon-saga-reveals-a-deep-trust-deficit-and-thats-worrying
Zaidi, S. M. S., & Saud, A. (2020). Future of US-China Relations: Conflict, Competition or Cooperation? *Asian Social Science, 16*, 1–14. https://doi.org/10.5539/ass.v16n7p1
Zhang, C. (2022). Contested disaster nationalism in the digital age: Emotional registers and geopolitical imaginaries in COVID-19 narratives on Chinese social media.

Review of International Studies, 48(2), 219–242. https://doi.org/10.1017/S0260210522000018
Zhao, Z. (2021). Analysis on the "Douyin (Tiktok) Mania" Phenomenon Based on Recommendation Algorithms. *E3S Web of Conferences, 235,* 03029. https://doi.org/10.1051/e3sconf/202123503029

Appendix

Excerpt from codebook

Code 1 #: China overreacting
Modality: (c) news video clip +voiceover
Channel type: Private person. Additional info: American middle-aged man
Semantic theme (explicit): "China is again pissed off with the US for moving their naval ships in international waters in the South China Sea". This is overreacting, it is international waters, not Chinese territory.
Latent theme, frame: ridicule/ provocation (the other is overreacting, stupid, warmongering)
Other framing techniques: hashtags containing #usChinawar #conflict #Chinatreat #taiwan #trump #republican #Chinacommunistparty – pertaining to the potential for war, and the user likely being a Trump supporter
Other notes: "Talking in we-form „we were sailing there because..." "China should take us seriously..." – either being part of the military/ navy himself, or strongly identifying with his country's politics.

Hashtags pertaining to USA and China

The hashtags are ranked by popularity across the platforms Instagram, Twitter, Facebook and TikTok
usChina – (17.4 million views – relevant)
Chinausa (16.6 million views – relevant)
usChinarelations (9.3 million views – relevant)
China – (too many irrelevant hits)
tradewar – (too many irrelevant hits)
usa – (too many irrelevant hits)
usChinatradewar – (only 419 k views)
trade (too many irrelevant hits)
usChinarelations
trump – (too many irrelevant hits)
Chinanews – (not popular on TikTok, only 177k views)
usandChina – (not popular on TikTok, only 18.5k views)

Source: www.besthashtag.com

Sowing the Seeds of Distrust? Foreign Cyber Interference in the French Presidential Election in 2017 and 2022

Franck Orban

Introduction

Digitalization has gradually permeated every aspect of our modern societies. This also applies to electoral processes. Digitalization makes new ways of communicating available to candidates and contributes to the creation of new communication channels between them and voters. Thus, digitalization strengthens democracy and increases chances for dynamic participative behavior (Lorenz-Spreen et al., 2021, p. 74). But it can also be seen with more skepticism if social media, big data, digital surveillance, and control, as well as online misinformation, hate and polarization become a threat to our democracies (Porter & Tan, 2022, p. 132). Poorly controlled digitalization can alter democratic processes if information is distorted by actors who wish to deceive and alter trust between candidates and voters and to disrupt or to influence an electoral process and its outcome. Candidates, voters, or electoral processes can be victims of cyberthreats or cyberattacks. Influencing voters' electoral choices and trying to change elections' outcomes weakens in time the public consensus on democracy, and even on truth (Delerue, 2021, p. 347-348). To sum up, digitalization has a critical impact whether it is used to strengthen or undermine the integrity of the electoral environment, trust towards authorities, as well as bonds between voters and politicians (Annan, 2020, p. 9).

This chapter takes a closer look at the notion of Foreign Cyber Electoral Interference and how state actors can undermine or alter electoral processes in third countries. Our case study is the presidential elections in France in 2017 and 2022. Firstly, we analyze France's presidential election in April-May 2017 to see how foreign states and especially Russia, used cyber subversion to try to alter the electoral process and to change its outcome. For the first time in its contemporary history, France was confronted with a massive and coordinated campaign of hacking and disinformation aimed at destabilizing the core of the most important democratic process of the Fifth Republic; the election of the French president (Pajot, 2021, p. 312). Secondly, we describe how French authorities built up capabilities and a new legal framework to counter foreign cyber electoral interference after the presidential election in 2017 up to the next presidential election in 2022. Thirdly, we review the measures regarding the presidential election in April 2022 to prevent a repeat of what happened in 2017. In the conclusion, we discuss the reasons of the absence of significant foreign cyber interference in the presidential election in 2022 vs. 2017.

Foreign Cyber Electoral Interference in the French Presidential Election in 2017

Foreign electoral interference can be defined as intentional covert or covert attempts by a foreign power to determine election results, to influence public perceptions to advantage or disadvantage election candidates, or to exacerbate internal divisions in another country. A divided establishment will struggle to create the conditions for consensus on a national level. Lack of domestic support will facilitate acceptance for compromises in foreign policy to a greater extent than what might have been the case if the targeted state had been stronger. During the two last decades, the digital revolution marked a definitive transition to *Foreign Cyber Electoral Interference*. Today, cyberattacks targeting electoral processes exploit other states' systemic and institutional vulnerabilities and try to disrupt voter registration databases, electronic voting machines, the counting or tabulations of results, the transmission and publication of election outcomes, and other critical tools of election administration (Mohan & Wall, 2019, p. 113). Hacking operations are complemented by misinformation like fake-news, computational propaganda campaigns, big data, micro-targeting, and internet trolling (Tenove et al., 2018).

In that sense, French authorities carefully followed the Brexit-referendum in the UK in 2016 (House of Commons, 2019, p. 70; McGaughey, 2018, p. 334), the US presidential election the same year (Office of the Director of National Intelligence, 2017, p. 11) and the Dutch general election early in 2017 (Brattberg & Maurer, 2018, p. 7). They feared foreign cyber interference in the presidential election to be held in France in April-May 2017. The French interest in cybersecurity linked to electoral processes appeared a few years earlier. A National Cyber security Agency (ANSSI) was created in 2009 to protect government communications and strategic industries against cyber snooping and cyberattacks. The scope of its missions expanded rapidly. ANSSI was involved in the prevention and the reaction to IT incidents that could affect sensitive institutions or electoral processes. The agency organized meetings with political parties to brief them about cyber security in internal seminars run by ANSSI-experts. In October 2016, ANSSI organized a cybersecurity seminar for 40 representatives from political parties during which ANSSI-experts instructed participants to raise awareness among partymembers and leaders about the risk of cyber manipulation. All parties showed up apart from the far-right National Front party. Participants received a cyber security pamphlet for people working in politics, a 52-page guide on IT best practices, a 52-page explainer on denial-of-service attacks, a USB flash drive loaded with information about malware and computer viruses, and a travel advice booklet (Leicester, 2016).

Government agencies did not appear to make centralized efforts to guard against disinformation campaigns at once. However, Minister of Foreign Affairs Jean-Marc Ayrault stated in February 2017 at the French parliament that the risks of foreign interference against the upcoming presidential election were high on the agenda. He warned that any attempt from Russia or other countries to disrupt the French electoral process would be met with proper response (Ayrault, 2017). This statement labelled any potential leak a "manipulation" in advance. The electoral commission established a mechanism allowing candidates to request an investigation by the ANSSI if a cyber intrusion was detected. The National Commission for the Control of the Electoral Campaign for the Presidential Election (CNCCEP), which is usually set up three months preceding every French presidential election to serve as a campaign watchdog, was reactivated in that occasion. Together with the Superior Audiovisual Council, the CNCCEP would issue recommendations against late or illegal electoral propaganda and would facilitate prosecution (Couzigou, 2021, p. 9).

In March 2017, the electoral commission banned electronic voting overseas for the legislative election coming just weeks after the presidential election, again to avoid cyber manipulation. French media also played their part by adopting appropriate measures to counter online disinformation. The newspaper *Le Monde* referenced hundreds of websites and evaluated their reliability (Brattberg & Maurer, 2018, p. 9). It also participated in a temporary fact-checking project, *CrossCheck*, which brought together 37 newsrooms, including Libération, Agence France-Presse, and France Médias Monde (Strand Larsen, 2019, p. 51). There were good reasons to assume that French media had learned from experiences from the US presidential election (Toucas, 2017). At the time the French presidential election campaign grew in intensity early in 2017, the American government was still investigating Russia's possible interference in the 2016-presidential US election.

Observers expected that France would logically be the next target of Russia's information warfare strategy, likely backed up by French domestic support players. In January 2017, Defense Minister Jean-Yves Le Drian reported in an interview with French newspaper *Le Figaro* that 24,000 external cyberattacks against the Ministry of defense had been blocked in 2016 by French security devices. The minister pointed out that a few hundred of these attacks were elaborated with real intentions to harm (Le Figaro, 2017).

French concerns were not exaggerated. The presidential electoral process suffered a major digital attack in three stages. The first stage was the hacking of several institutions and political parties throughout 2016-2017. The second stage was the spreading of various fake stories via trolls on social networks or published by Russian state media outlets like Sputnik and RT. Fake stories targeted candidates like François Fillon, Jean-Luc Mélenchon, and Alain Juppé. But most of them aimed to cast doubts on Emmanuel Macron's private life or professional ethics as a candidate. On February 4, a Sputnik article featured Macron as a «US agent» backed by «a very wealthy homosexual lobby. Two hours before the decisive second round debate between Macron and his contender Marine le Pen on May 3., a forum, which already shared fake news during the US presidential election, started to disseminate documents claiming that Macron had a secret offshore bank account in the Bahamas that nobody knew about. As Marine le Pen mentioned this account during the debate on French national television, interest on social media skyrocketed (Capron, 2017). Attempts to discredit Macron and to favor le Pen also emanated from American trolls belonging to the extreme right and were determined to facilitate the election of the candidate from the National Front (Harkinson, 2017). The third stage of the cyberattack came two days before the first round of the election under a mandatory media blackout enforced by the electoral commission.

The "Macron leaks" consisted of the hacking and the online dissemination of 150,000 emails and documents originating in the Macron campaign and belonging to five individuals linked to the party *La République En Marche* (LRM) and close to its candidate Emmanuel Macron. The leaks were a combination of real emails and forgeries. Around 9GB of data was posted by a user called EMLEAKS to the document-sharing site Pastebin that allowed anonymous posting. Stolen documents were posted as #MacronLeaks on social networks in the .eml format and linked to Pastebin. They were shared on 4chan and on pro-Trump Twitter accounts in the English language, before being relayed to WikiLeaks (Le Monde, 2017). Trend Micro, a security firm, attributed in January 2017 many initial phishing attacks to Pawn Storm. The latter was considered a tool of Russia's Main Intelligence Directorate (GRU) by US intelligence agencies (Hacquebord, 2017, p. 13).

These cyber-attacks were relayed by the complicity of domestic opponents to Macron in France, especially in far-right groups, who disseminated fake news on social networks and rephrased and adapted them to fit the French cultural and political context. Among the individuals who broadcasted fake news produced by Sputnik or Russia Today during the first round of the presidential election, a majority were supporters of François Fillon and Marine Le Pen (Gaumont et al., 2018, p. 29). The pattern of manipulation started with the dissemination of the Macron Leaks on 4Chan, then to more conventional social networks like Twitter, before being disseminated by American Alt-right and French far-right communities (Jeangène Vilmer et al., 2018, p. 110). Just after the Macron Leaks were circulated, the French electoral commission published a statement urging all media outlets to respect the campaign blackout period and not to comment at all on the leaks before election day to protect the fairness of the electoral process and its outcome. At the same time, Macron's head of cyber operations, Mounir Mahjoubi, claimed that the leaks contained fake information, forcing WikiLeaks to distance itself from the Macron Leaks.

Analysis carried out by social media followers showed that several documents contained Russian characters, while others were grossly fabricated (Capron, 2017). Within a few hours, serious doubts were raised about the validity of the leaks, as it was nearly impossible to know if documents were true or fabricated. Public opinion was eventually deflected from the candidate Macron to Russia's responsibility in the cyberattack. The targeting of candidate Macron failed as Russian hackers were unsuccessful in discrediting Macron and in precipitating the election of his opponent Marine le Pen (Jeangène Vilmer et al., 2018, p. 108; Vanderbiest, 2018, p. 187).

The presidential election in France in 2017 reached the third stage of foreign cyber electoral interference of five, as described by Aaltola in the context of the US presidential election (Aaltola, 2018, p. 3). The French electoral process experienced disinformation to create suspicion and division (stage 1), theft of sensitive and digital information data (stage 2), and leaks of stolen data via suspected hacktivists (stage 3). However, it did not suffer whitewashing of the leaked data through professional media (stage 4) or a conspirative collusion between candidates and foreign states (stage 5).

There were several reasons for the failure of such cyberattacks (Jeangène Vilmer et al., 2018, p. 117). Firstly, making any attempt to interfere in the French electoral process was more demanding than was the case in the US and in the United Kingdom. A system based on the direct election of the president with two rounds and majority votes made it impossible for trolls to know in advance who would reach the second round. Secondly, the Commission for the Control of the Electoral Campaign for the Presidential Election (CNCCEP) and the National Agency for the Security of Information Systems (ANSSI) played their part perfectly, as they alerted French media, political parties, and the public to the risks of cyberattacks and information manipulation during the campaign, met and educated all campaign teams early in the electoral process, and intervened to effectively block the spread of misinformation at a time when it could have had an impact on voters. Thirdly, the main victim of the disinformation campaign, the Macron-team, anticipated the attack and developed a full strategy to counter it thanks to fake email addresses, fake passwords, and fake documents, which aimed at degrading the value of any possible cyberattacks. Fourthly, the French media environment turned out to be more resilient than anticipated, as there was a strong tradition of serious journalism in France. The latter was supported by most French readers, who still preferred conventional news sources to tabloid and alternative media. Fifthly, hackers committed mistakes. They overestimated their ability to shock and mobilize virtual communities, underestimated the resistance and intelligence of conventional media and, above all, they did not expect that the Macron-team would react fast and efficiently. Sixthly, launching an attack just a few hours before the period of electoral silence was a double-edged sword. Hackers wanted to paralyze Macron's team. By waiting until the last minute, they didn't have enough time to spread disinformation themselves. Observers reacted to the timing of the operation with suspicion. The whole operation lost credibility.

To conclude with humor, one should remember that foreign trolls wrote mainly in English. Most of the French people remain immune to Shakespeare's language. What is often portrayed as a typical French flaw proved to be an effective protection against foreign designed misinformation.

Building up Resilience against Foreign Cyber Electoral Interference after 2017

French authorities were relieved to see that the presidential election in 2017 was held in a relatively normal manner without being endangered by cyberattacks from external actors. Still, French authorities were upset by these attacks and realized that national authorities and private actors lacked proper means to counterattack attempts and misinformation in cyberspace. Fighting cyberthreats, dishonest propaganda and disinformation became a political top priority right after the election. One of the first political acts of the newly elected president Macron was to condemn Russia's disruption attempt during the French campaign firmly and publicly at a press conference with Russia's president, Vladimir Putin at the Château de Versailles on May 29, 2017 (Macron, 2017).

That did not stop further Russian interference. Between November 17, 2018, and September 9, 2019, 22% of the 6,944 videos published by RT France on its YouTube channel, exclusively focused on the Yellow Vests movement (Colon, 2023, p. 284). These reports were viewed more than 23 million times. The same was true when RT France gave extensive coverage of "Antivax" demonstrations in France during the pandemic. Russia actively contributed to exacerbating the political and social divisions in France, particularly by emphasizing the state's repression of the protesters. In January 2019, the French minister for the armed forces Florence Parly publicly admitted that a cyber war had begun and that France had to be ready to fight it (Parly, 2019).

Beside increased political awareness of the reality of cyber threats at the top level, French authorities developed capacities to prevent any repetition of the cyber-attacks in 2017. ANSSI was attached to the General Secretariat of Defense and Security (SGDSN) and was given the task of ensuring the cyber protection of the French State (protecting information networks, regulating critical infrastructures and the private sector, certifying products, and hosting the national Computing Emergency Response Team). The agency was separated from the intelligence and military branch of French cybersecurity (Liebetrau, 2022, p. 136). Specialized services were created in all intelligence services (external with DGSE, internal with DGSI, and military with DRM-DRSD). A cyber defense operational chain of command (COMCYBER) was established in Paris in January 2017 under the orders of the Armed Forces Chief of Staff, which replaced the command chain from 2013. COMCYBER had 3200 cyber fighters a year later. In time, the collaboration between ANSSI, COMCYBER, and French intelligence services was expected to increase the level of coordination between defensive and offensive cyber actions (Liebetrau, 2022, p. 117).

With respect to the political and civilian treatment of cyber threats, several laws were passed to prevent the spread of misinformation and propaganda and to counter cyberthreats. In March 2018, several members of the National Assembly who were close to the newly elected president tabled a motion entitled the "Law to combat false information." The motion sparked numerous deliberations and an enduring dispute between the two chambers. The Senate rejected the proposal on two occasions. Throughout the course of these discussions, modifications were introduced, resulting in a change in the motion's title from "Law to combat false information" to "Law against the manipulation of information." Critical voices perceived the draft on the table as a dangerous tool to strengthen state control over the information space and to limit press freedom or as an unnecessary duplication of already existing legal means (de Bellescize, 2018, p. 565). Ultimately, the National Assembly granted its approval for the motion on November 20, 2018. Following this, the French Prime Minister, along with over 60 Senators, opted to file an appeal with the Constitutional Council, seeking a legal assessment of the motion's validity. The final text was validated by the Constitutional Council one month later (Guillaume, 2019, p. 3).

The primary goal of the new law was to overcome weaknesses of the existing French legal framework inherited from the law on the Freedom of the Press of July 1881 and to adapt it to the digital age. It ensured a higher level of overall integrity of electoral debates and voting results and reduced the risk of French citizens being tricked in exercising their vote by preventing the spread of cyber disinformation. Particular attention was paid to election campaign periods, just before and during elections. The new law made it easier to counter "fake news" during election campaigns. The online operators targeted by the law were those whose activity exceeded five million unique visitors per month, per platform, and who offered, in a professional, paid, or unpaid manner, a service of communication to the public. Therefore, it aimed at services with a broad scope of activity, including general social networks (Snapchat, Instagram, Facebook, TikTok) or specialized ones (LinkedIn), platforms for sharing of videos or audios (YouTube, Dailymotion, Twitch), online forums (Jeuxvideo.com), search engines (Bing, Google), and participative encyclopedias (Wikipedia). Social platforms were summoned to show more in the transparency of their algorithms. They had to report any sponsored content and to publish the name of their author and the amount they received. Those which exceeded a certain volume of connections per day had to have a legal representative in France. Social platforms also had to make their algorithms public.

In addition, new legal means were put in place to quickly stop the circulation of fake news. Judges in charge of such cases (juge des référés) could qualify news as "fake" based on three criteria. First, it had to be obviously false or misleading. Second, it had to be distributed massively and artificially. Finally, it had to disturb public order or threaten the validity of an election. After a case was presented to the judge, the latter would have 48 hours to implement appropriate actions to stop the dissemination of illegal content (Guillaume, 2019, p. 3). Candidates and political parties could appeal to a judge to help them to stop fake news during a period of three months before an election. Moreover, social platforms had to comply with obligations to remove manifestly false information promptly during elections. In periods without elections, they had to cooperate with French authorities and to take initiatives in combatting disinformation. They also had to communicate on their efforts. The French Media Regulatory Authority (CSA) oversaw whether social platforms cooperated in preventing disinformation, or not. The agency CSA could suspend television channels or stop programs which were suspected of being influenced by foreign states or of disseminating fake news likely to affect national French interests.

The law finally emphasized the importance of enhancing and promoting media literacy, particularly for content distributed online, within educational institutions. As Marine Guillaume recalled, the French law against the manipulation of information from 2018 introduced a strategy for countering information manipulation campaigns, primarily by bolstering the authority of reliable third parties, such as the CSA, and by urging online platforms to enhance their transparency. Additionally, the law sought to uphold the fundamental principles of freedom of speech and expression crucial to democratic societies, particularly during electoral periods when discussions and viewpoints tend to become polarized, but also outside electoral periods (Guillaume, 2019, p. 6).

In July 2019, the French National Assembly also tabled a law against online hate speech, which aimed at requiring online platforms to address content that was unlawful on grounds of race, religion, sex, sexual orientation, or disability within 24 hours. The "Avia Law" was passed on 18. May 2020, but it was ultimately found to be largely unconstitutional by France's highest court, the Constitutional Council, one month later. The council censored the obligation imposed on social networks to delete within 24 hours hateful content reported to them on Facebook, Twitter, Snapchat, or YouTube. The remaining part of the law came into force on 1st July 2020, but it was deprived of most of its substance. Finally, a rebranded "Law Reinforcing the Republican Principles" (which got the nickname "separatism law"), was passed in August 2021, which regulated social networks in France and intended to prohibit the dissemination of hateful content on social media platforms.

Foreign Cyber Electoral Interference in the French Presidential Election in 2022

The presidential election in 2022 took place in a political climate in France which had severely declined between the takeover of Macron in May 2017 and the beginning of the 2022 electoral campaign, which started for real when the Covid crisis began to ease. As the new round of election involved a president who wanted to be re-elected, a large part of the French opinion was exhausted after the yellow vest crisis, which began in October 2018 and lasted a year and a half, huge protests against the new pension reform law which was announced at the end of 2019 and the beginning of 2020 and finally by the Covid crisis that erupted in 2020 to last until 2022. The confidence of the French people in politics, in their political leadership, and in institutions had greatly eroded. Considering the state of physical and mental fatigue of the population after the Covid crisis, as well as the state of democratic fatigue that many of them felt after five years of ongoing confrontation with the power in place, one could only fear that most voters were to sink either into apathy and abstention or into anger and protest vote.

Such a political climate was a fertile ground for the spread of disinformation and polarizing narratives (Online Integrity Watch Group, 2022, p. 11). For the first time in a French presidential election, the far-right candidate Éric Zemmour used the notion of "Great Replacement" of the French population, coined by French identitarian writer Renaud Camus in the beginning of the 2010s and later picked up by American supremacists, to point out the specter of a future civil war in France. With help from supporters in cyberspace, he managed to monopolize the political debate between autumn 2021 and February 2022 to such an extent that experts talked about the "zemmourization" of the electoral campaign (Hope not Hate, 2022, p. 7). He and several candidates disseminated Russian narratives. They relied on a network of academic experts, army veterans, former spies, or journalists in mainstream newspapers, who also reported a Moscow-friendly speech (Colon, 2023, p. 282). Pro-Russian narratives and conspiracy theories about Ukraine and Covid-19 merged with pre-existing grievances in fringe online communities and were used to promote distrust of institutions and elites (Simmons, Gatewood & Fourel, 2022, p. 14).

This political and societal framework represented a far better opportunity for external actors to disrupt the presidential election than had been the case in 2017. Just days before the first round of the presidential election April 10, 2022, a study uncovered that French voters were largely permeable to fake news and that many distanced themselves from conventional information outlets. Before and during the presidential campaign, false information circulated on social media targeting candidates and the integrity of the voting system (Coleman, 2022). The topics largely reflected the concerns of the French people (Caline & Vardaxoglou, 2022, p. 2).

To prevent the risk of a repetition of the scenario from 2017, proactive measures to address foreign cyber-attacks or disinformation were taken by French authorities. They put pressure on online platforms to exercise vigilance and to deter malicious actors from participating in online influence campaigns. As early as on 18 February 2022, about five years after the Macron Leaks, campaign teams from several candidates to the upcoming presidential election met secretly to discuss how they could increase awareness concerning the protection of their cyber data. The meeting was organized by the General Secretariat of Defense and National Security (SGDSN), placed under the authority of the Prime minister. The presence of high-profile candidates' representatives at the technical and political level, showed that the protection of operational data was taken seriously by all candidates, including the National Rally, which was the new name of the National Front. From the far left to the far right, IT security officials agreed that cyber protection was a central issue for the election which – if not taken seriously this time – could alter the election outcome and undermine voters' trust in candidates, parties, and the electoral system.

ANSSI and VIGINUM, the new *Vigilance and Protection Agency against Foreign Cyber Interference*, which was established in July 2021 to track down disinformation during the election period, monitored this meeting. This agency was placed under the General Secretariat for Defense and National Security (SGDSN) and had around sixty employees, primarily data and social media experts, analysts, linguists, and computer scientists. Its mission was to safeguard the nation from foreign cyber interference, aimed at polarizing debates, disseminating false information, or promoting specific themes (VIGINUM, 2022, p. 14). The content targeted by VIGINUM had to meet four main criteria. First, it went against the interests of the Nation. Second, it conveyed false information. Third, it originated from a foreign state. Last, it was digitally and automatically amplified. The agency launched the operation dedicated to securing the French presidential election from manipulation of information in October 2021 and had two awareness-raising meetings with political parties participating in the 2022 presidential-election between October 2021 and February 2022, as well as one meeting with the main digital platforms in October 2021. Russian, and to a lesser extent Chinese, Turkish, and American conspiracists were under scrutiny.

In January 2022, the Commission on Information Manipulation overseen by Gerald Bronner submitted its recommendations to the government on how to address disinformation and conspiracy theories. Bronner was known for his work on collective beliefs and his involvement in jihadist radicalization prevention programs. The report aimed at developing a culture of digital security that would include information risk and involve all actors of the State and the government. Suggestions were made about how to counter foreign cyber electoral interference. That included the protection of the integrity of electoral processes, data sharing rules between trusted actors, the creation of an interdepartmental digital governance mechanism, the possibility to refer to the defense ethics committee on the cyber influence warfare doctrine of the ministry of the armed forces, and the creation of EU crisis management mechanism and crisis management exercises, as well as the organization of a working-group at the OECD-level (Bronner, 2022, p. 110-111).

Still, at the beginning of January 2022, the Audiovisual and Digital Communications Regulatory Authority (ARCOM) came into existence as an independent administrative agency, following the merger of the High Audiovisual Council (CSA) and the High Authority for the Distribution of Works and Protection of Rights on the Internet (HADOPI). ARCOM inherited the responsibility for overseeing audiovisual and digital communications. As France's media regulator, ARCOM contributed to raising awareness among candidates, political parties, the media, and social networks about their responsibility in addressing disinformation. As an example, ARCOM organized a meeting with representatives from Dailymotion, Google, Microsoft, LinkedIn, Meta, Snap, TikTok, Twitch, Twitter, Webedia, and Wikimedia France on January 28, 2022, to share their recommendations on transparency, request them to identify a primary point of contact for the regulator, and propose an operational procedure in the event of a crisis during the two electoral periods (ARCOM, 2022b, p. 46).

In addition, France held the presidency of the Council of the European Union at the time of the presidential election. That was a critical period for finalizing the European Digital Services Act (DSA). The DSA was the most ambitious regulation in the world in the field of the protection of the digital space against the spread of illegal content (Simmons, Gatewood & Fourel, 2022, p. 16). The protection of users' fundamental rights outlined the responsibilities of online platforms concerning illegal and problematic content. Foreign digital interference became a highly sensitive topic for major online platforms that would come under regulation (Online Integrity Watch Group, 2022, p. 28). In addition, the EU's sanctions on Kremlin-affiliated outlets from March 2022, including Russia Today and Sputnik, thwarted Russian trolls' efforts to influence voter perceptions by blocking their content across cable, satellite, internet TV, ISPs, and video-sharing apps. These measures might have motivated online platforms to comply with regulations more rapidly.

French fact-checkers also played a noticeable role in debunking hoaxes which tried to manipulate the presidential election in 2022 (Nicey, 2022). A major difference from the former election in 2017 was that many media outlets, such as France Info, Le Monde, Le Figaro or Libération, now had their own fact-checking departments. In addition, French media outlets collaborated intensively. The seven members of the International Fact-Checking Network (IFCN) in France verified a total of 169 disinformation items linked to the presidential election which circulated in France from January 1 to April 30. In early 2022, the news agency AFP and Google took the initiative to bring together 23 French media outlets and 180 journalists to create the "Objectif Désinfox" platform. This platform offered training to media editorial teams to cover both the presidential and the general election and allowed French citizens to report digital content that seemed dubious to them for fact-checking. In its 20 weeks of activity, the platform published 202 fact-checks and produced 23 videos for the public (AFP Factuel, 2022).

Conclusion

The French presidential election in 2022 did not unfold as feared by many experts and foreign attempts to disrupt it resulted in limited impact. In its first activity report published in November 2022, VIGINUM stated that it had detected 60 inauthentic phenomena on digital platforms during operations aiming at protecting the digital public debate surrounding the presidential and general elections in 2022. Just five were finally characterized as meeting the criteria defining foreign cyber electoral interference (VIGINUM, 2022, p. 16). In its report about the French presidential election in April and the general election later in June, ARCOM reported on its site that platform operators noted calm electoral campaigns on their services compared to the presidential and general elections in 2017.

French authorities took actions between 2017 and 2022 and passed new laws to combat disinformation more efficiently. The enactment of the law against the manipulation of information in 2018, significantly contributed to the establishment of institutional capabilities for combatting online information manipulation (ARCOM, 2022a, p. 4). Online platforms came under pressure to be more vigilant and to deter foreign actors from engaging in cyber interference. France also created national agencies which aimed at combatting foreign cyber interference and disinformation. VIGINUM was launched in the lead-up to the presidential election, while France's media regulator ARCOM gained strength at the same time and played a central role in raising awareness among candidates, political parties, the media, and social networks about cyber challenges and their responsibility in countering disinformation. French media and fact-checkers played their part in acting separately and grouped to debunk hoaxes generated by domestic and foreign actors trying to corrupt the election. IFCN-factcheckers couldn't find any significant evidence of foreign interference in the production or propagation of disinformation related to this election (Henin, 2022). They could rely on French readers and viewers who generally were knowledgeable, capable of critical thought, and who had access to a great variety of information sources.

Other factors have been put forward to explain the relative success of the presidential election in 2022. One of them was the voting conditions in France with paper ballots, single ballot, and no online voting for this election on the one hand, as well as the disparities in the results between the first and second round of the presidential election. The few contents aiming to challenge the legitimacy of the election did not go viral (ARCOM, 2022b, p. 52). Another factor was the level of certainty surrounding the re-election of president Macron according to all sociological polls conducted throughout the presidential race. His approval rating remained consistently at around 25%, ensuring his advancement from the first round of the election on April 10, 2022, to the second round two weeks later, and to an expected victory against far-right contester Marine le Pen. A third decisive factor was Russia's invasion of Ukraine, which discredited candidates who advocated a pro-Moscow stance (Chabal, Behrent, & Van Renterghem, 2023, p. 75-76). Worried French citizens who feared an increasing inflation rate affecting their personal economy and an explosive international situation rallied behind their president as a reassuring and protective figure in pure Napoleonic style (Hazareesingh, 2021).

Without the war, Russian propaganda might have had a much greater impact on disabused voters. Macron lost the general election later in June 2022 after all. This time, voters just got scared. Russia's war efforts mobilized at the same time enormous financial, technical, and human resources which could not be spent to harm France. In addition, the sanctions from the European Union on Kremlin-backed news outlets made it harder for Russian trolls to interfere and manipulate voters (Kayali & Goujard, 2022). When the trauma of the Ukrainian war vanishes, one can just assume that Moscow's strategy of confusing foreign cyber electoral interference with more common influence policies will resume with more vigor (Tanguy & Le Grip, 2023, p. 89). This reminds us that even if national and international measures to counter foreign cyber electoral interference are crucial in maintaining electoral integrity, they alone cannot rectify the deep-seated mistrust in political figures and institutions that can jeopardize democratic dialogue. Strategies to counter foreign cyber electoral interference are simply no stand-in for a sound political environment which bolsters democratic resilience.

References

Aaltola, M. (2018). Democracy's Eleventh Hour: Safeguarding Democratic Elections Against Cyber-Enabled Aurocratic Meddling. *FIIA Briefing Paper,* nr 226. https://www.fiia.fi/wp-content/uploads/2017/11/bp226_democracys_eleventh_hour.pdf

AFP Factuel (2022, June 30). Elections 2022: les vérifications de la coalition Objectif Désinfox. https://factuel.afp.com/elections-2022-les-verifications-de-la-coalition-objectif-desinfox

Annan, K. (2020). *Protecting Electoral Integrity in The digital age*. Kofi Annan Foundation. https://www.kofiannanfoundation.org/app/uploads/2020/05/85ef4e5d-kaf-kacedda-report_2020_english.pdf

ARCOM (2022a). The fight against the manipulation of information on online platforms. A look back at three years of implementation of the French law of December 22nd 2018. https://www.arcom.fr/sites/default/files/2022-12/Manipulation%20of%20information%20-%20Implementation%20review%202018-2022.pdf

ARCOM (2022b). Rapport sur les campagnes électorales. Élection à la présidence de la République (10 avril–24 avril 2022) Élections législatives (12 juin–19 juin 2022). November 2022. https://www.arcom.fr/nos-ressources/etudes-et-donnees/mediatheque/rapport-sur-les-campagnes-electorales-2022-election-la-presidence-de-la-republique-et-elections-legislatives

Ayrault, J.-M. (2017, February 17). Déclaration de M. Jean-Marc Ayrault, ministre des affaires étrangères et du développement international, en réponse à une question sur la cybersécurité, à l'Assemblée nationale le 15 février 2017. https://www.vie-publique.fr/discours/202180-jean-marc-ayrault-15022017-question-assemblee-nationale-cybersecurite

Brattberg, E., & Maurer, T. (2018, May 23). Russian election interference: Europe's counter to fake news and cyber-attacks. *Carnegie Endowment for International Peace.* https://carnegieendowment.org/2018/05/23/russian-election-interference-europe-s-counter-to-fake-news-and-cyber-attacks-pub-76435

Bronner, G. (2022). *Les Lumières à l'ère numérique*. Rapport de la Commission Bronner. https://pro.bpi.fr/content/uploads/sites/3/2022/01/les-lumieres-a-lere-du-numerique-rapport-de-la-commission-janvier-2022-2.pdf

Caline, G., & Vardaxoglou, L. (2022, April 8). Présidentielle 2022: comment Mesurer l'impact des "fake news" sur les électeurs? *Fondation Jaurès,* 6 p. https://www.jean-jaures.org/publication/presidentielle-2022-comment-mesurer-limpact-des-fake-news-sur-les-electeurs/?post_id=32929&export_pdf=1

Capron, A. (2017, May 5th). How we debunked rumours that Macron has an offshore account. *France 24.* https://observers.france24.com/en/20170505-france-elections-macron-lepen-offshore-bahamas-debunked

Chabal, A., Behrent, M. C., & Van Renterghem, M. (2023). Historical perspectives on the 2022 electoral cycle in France. *French History,* 37-1, 71–84. https://doi.org/10.1093/fh/crac070

Coleman, A. (2022, April 23). French election: Misinformation targets candidates and voting system. *BBC News.* https://www.bbc.com/news/61179620

Colon, D. (2023). *La guerre de l'information: Les Etats a la conquête de nos esprits.* Tallandier.

Couzigou, I. (2021). The French Legislation against digital information manipulation in electoral campaigns: a scope limited by freedom of expression. *Election Law Journal, 20*(1), 98–115. https://doi.org/10.1089/elj.2021.0001

De Bellescize, D. (2018). Fake news: une loi polémique, qui pose plus de questions qu'elle n'en résout. *Constitutions. Revue de droit constitutionnel appliqué,* 4, 559-570.

Delerue, F. (2021). Chapitre 15. Les menaces numériques en période électorale. In C. Marangé, & M. Quessard (2021). *Les guerres de l'information à l'ère numéri-*

que, Presses universitaires de France, 347–363. https://doi.org/10.3917/puf.maran.2021.01.0347

Gaumont N., Panahi M., & Chavalarias D. (2018) Reconstruction of the socio-semantic dynamics of political activist Twitter networks—Method and application to the 2017 French presidential election. *PLoS ONE* 13(9), https://doi.org/10.1371/journal.pone.0201879

Guillaume, M. (2019). Combating the manipulation of information – a French case. *Hybrid CoE,* May 2019. https://www.hybridcoe.fi/wp-content/uploads/2020/07/HybridCoE_SA_16_manipulation-of-information_.pdf

Hacquebord, F. (2017) Two Years of Pawn Storm: Examining an Increasingly Relevant Threat. *Trend Micro.* https://documents.trendmicro.com/assets/wp/wp-two-years-of-pawn-storm.pdf

Harkinson, J. (2017, May 3). Inside marine Le pen's "foreign legion" of American alt-right trolls. *Mother Jones.* https://www.motherjones.com/politics/2017/05/marine-le-pen-alt-right-american-trolls/

Hazareesingh, S. (2021, July 9). Une part de nous: Emmanuel Macron's admiration for Napoleon. LS. *Times Literary Supplement,* no. 6171.

Henin, N. (2022, June 28). What did disinformation look like during the 2022 French presidential election? An overview based on fact-checking articles. EU Disinfo Lab. https://www.disinfo.eu/publications/what-did-disinformation-look-like-during-the-2022-french-presidential-election-an-overview-based-on-fact-checking-articles/

Hope Not Hate (2022). *Eric Zemmour and his rise on social media platforms.* Hope Not Hate. March 2022. https://hopenothate.org.uk/wp-content/uploads/2022/04/eric-zemmour-report-2022-5-final1.pdf

House of Commons (2019). *Disinformation and 'fake news'*: Final Report. Digital, Culture, Media and Sport Committee. Eighth Report of Session 2017–19. https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf

Jeangène Vilmer, J.-B., Escorcia, A., Guillaume, M., & Herrera, J. (2018) Les Manipulations de l'information : un défi pour nos démocraties. Rapport. *Centre d'analyse, de prévision et de stratégie du Ministère de l'Europe et des Affaires étrangères et Institut de recherche stratégique de l'École militaire* (IRSEM). https://www.diplomatie.gouv.fr/IMG/pdf/les_manipulations_de_l_information_2__cle04b2b6.pdf

Kayali, L., & Goujard, C. (2022, March 2). EU officially boots Russia's RT, sputnik outlets. *Politico.* https://www.politico.eu/article/russia-rt-sputnik-illegal-europe/

Le Figaro (2017, January 8). Le ministère de la Défense a déjoué 24.000 cyberattaques en 2016. https://www.lefigaro.fr/international/2017/01/08/01003-20170108ARTFIG00050-le-ministere-de-la-defense-a-dejoue-24000-cyberattaques-en-2016.php

Le Monde (2017, May 6). En marche ! dénonce un piratage « massif et coordonné » de la campagne de Macron. https://www.lemonde.fr/election-presidentielle-2017/article/2017/05/06/l-equipe-d-en-marche-fait-etat-d-une-action-de-piratage-massive-et-coordonnee_5123310_4854003.html

Leicester, J. (2016, December 15). After US election hacks, France girds against cyberattacks. *Phys.org.* https://phys.org/news/2016-12-election-hacks-france-girds-cyber-attacks.html

Liebetrau, T. (2022). Organizing cyber capability across military and intelligence entities: collaboration, separation, or centralization, *Policy Design and Practice,* 6(2), 131-145, https://doi.org/10.1080/25741292.2022.2127551

Lorenz-Spreen, P., Oswald, L., Lewandowsky, S., & Hertwig, R. (2021, November 22). A Systematic Review of Worldwide Causal and Correlational Evidence on Digital Media and Democracy. *Nat Hum Behav* 7, 74-101 https://doi.org/10.31235/osf.io/p3z9v

Macron, E. (2017, May 29) Conférence de presse conjointe de MM. Emmanuel Macron, Président de la République, et Vladimir Poutine, Président de la Fédération de Russie, sur les relations franco-russes et sur les conflits en Syrie et en Ukraine. Vie publique. https://www.vie-publique.fr/discours/203301-conference-de-presse-conjointe-de-mm-emmanuel-macron-president-de-la-r

McGaughey, E. (2018). Could Brexit be void? *King's Law Journal,* 29(3), 331-343. https://doi.org/10.1080/09615768.2018.1555881

Mohan, V., & Wall, A. (2019). Foreign Electoral Interference: Past, Present, and Future. *Georgetown Journal of International Affairs,* 20, 110-116. https://doi.org/10.1353/gia.2019.0019

Nicey, J. (2022). Le fact-checking en France, une réponse en condensé du journalisme face aux transformations numériques des années 2000 et 2010. *Les Enjeux de l'information et de la communication,* 23(1), 67-79. https://doi.org/10.3917/enic.032.0067

Office of the Director of National Intelligence (2017, January 6). Background to "Assessing Russian activities and intentions in recent US elections": The analytic process and cyber incident attribution. UNT Digital Library. https://digital.library.unt.edu/ark:/67531/metadc949269/

Online Integrity Watch Group (2022). T*he French information ecosystem put to the test.* https://checkfirst.network/wp-content/uploads/2022/06/Elections-2022_EN_Report.pdf

Pajot, B. (2021). Chapitre 13. La France: quelle stratégie de résilience face aux manipulations de l'information? In C. Marangé (Ed.), *Les guerres de l'information à l'ère numérique,* 309-319. Presses Universitaires de France.

Parly, F. (2019, January 18). Déclaration de Mme Florence Parly, ministre des armées, sur le volet de la cyberdéfense des armées. https://www.vie-publique.fr/discours/269137-florence-parly-18012019-strategie-cyber-des-armees-cyberdefense

Porter, T., & Tan, N. (2022). Introduction: Democracy and digitization. *Globalizations,* 20(1), 132-136. https://doi.org/10.1080/14747731.2023.2156033

Simmons, C., Gatewood, C. & Fourel, Z. (2022). *Monitoring Disinformation and Influence Campaigns in the 2022 French Elections: Lessons Learnt.* Institute for Strategic Dialogue. https://www.isdglobal.org/wp-content/uploads/2022/12/Monitoring-Disinformation-and-Influence-Campaigns-in-the-2022-French-Elections_Lessons-Learnt.pdf

Strand Larsen, E. (2019). The Limits of Truth – A Case Study of Faktisk and Cross-Check. In F. Orban & E.S. Larsen (Eds.), *Discussing borders, escaping traps: Transdisciplinary and Transspatial approaches,* 47-66, Waxmann. https://www.waxmann.com/index.php?eID=download&buchnr=4045

Tanguy, J.-P., & Le Grip, C. (2023). Rapport fait au nom de la commission d'enquête relative aux ingérences politiques, économiques et financières de puissances étrangères – États, organisations, entreprises, groupes d'intérêts, personnes privées – visant à influencer ou corrompre des relais d'opinion, des dirigeants ou des partis politiques français. *Assemblée Nationale.* https://www.assemblee-nationale.fr/dyn/16/rapports/ceingeren/l16b1311-t1_rapport-enquete.pdf

Tenove, C., Buffie, J., McKay, S., & Moscrop, D. (2018). Digital threats to democratic elections: How foreign actors use digital techniques to undermine democracy. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3235819

Toucas B. (2017, May 30). *The Macron leaks: The defeat of informational warfare*. Center for Strategic and International Studies. https://www.csis.org/analysis/macron-leaks-defeat-informational-warfare

Vanderbiest, N. (2018). Les institutions démocratiques: l'influence des réseaux sociaux durant une élection présidentielle. In S. Taillat (Ed.), *La Cyberdéfense: Politique de l'espace numérique,* 181-188, Armand Colin. https://doi.org/10.3917/arco.danet.2018.01.0181

VIGINUM (2022). VIGINUM year#1. *General Secretariat for Defense and National Security.* https://www.sgdsn.gouv.fr/files/files/Publications/RA-Viginum-Annee1-32p-V20_EN_LQP-1.pdf

Political Trust in Scotland: 2014 Independence Referendum and Digital Political Participation

Laila Berg

Introduction

This chapter concerns the link between political trust and digital technology in contemporary Scotland, and in particular how this link became manifested in digital participation in the Scottish independence referendum which took place on September 14, 2014. The background for this exploration is the combination of considerable political and discursive developments in Scotland such as the increase in devolved political powers and the growing salience of Scottish national identity and nationalism, and the digitalisation of Scottish society. It is argued here that the combination of political change and digitalisation of political action and participation has resulted in a greater public involvement in the period leading up to the 2014 referendum and thereafter. Building on statistical data from the Scottish Social Attitudes (SSA) and the author's own qualitative, open-ended interviews with 40 informants during a 7-month fieldwork in Scotland in 2018, this chapter examines the relationship between recent political changes and developments in Scotland, and national levels of political trust. Scotland is indeed an interesting case because of its position as a nation but not a state, and as the SSA, the annual survey tracking the Scottish public's social, political and moral attitudes since its establishment in 1999, has shown, political trust remains high in the Scottish government but low in the UK government, suggesting the relevance of national identity in affecting levels of political trust (Scholes et al., 2022).

Since the (re)establishment of the Scottish Parliament in 1999, the political and constitutional context of Scotland has been one of consistent and extensive change, leading to a stronger emphasis on Scottish national identity and engaging a large proportion of the Scottish public in political matters. Following the (re)establishment of the Scottish Parliament, there was a sequence of transformative political developments in quick succession that has restructured and altered the Scottish and British political landscapes, their positions to each other and to the wider European community at large. Since 1999, Scotland has been given devolved powers from Westminster twice, first in 2011 and then in 2016, and currently holds power over a substantial number of matters including among others, the economy, transport and taxation, and education (Scottish Government, 2024). Moreover, Scotland has witnessed two referendums within a two-year interval: the 14 September 2014 referendum on Scottish independence, and shortly thereafter the UK-wide 23 June 2016 EU referendum. This has led to significant shifts in political voting and party support in Scotland, and has revitalised Scottish nationalism, making Scottish national identity more salient and politically connected (Jackson, 2020).

The public's participation in Scottish politics and in particular in the issue of Scottish autonomy has increasingly taken place in the digital realm, using social media to engage, spread and mobilise political opinion. Digitalisation has transformed communication on both the individual and societal level, giving rise to novel digital tools and platforms for the reception, participation and distribution of political thought, action and debate. As such, digital technology has become a permanent part of political life, and while it can disrupt and destabilise political systems by spreading distrust and misinformation and thus contribute to political fragmentation, it can also have democratic transformative power to enable civic participation and give voice and agency to groups that traditionally have stood outside of policy making (Wike et al., 2022). The Scottish independence referendum is argued here to be an example of the latter.

Political Trust and National Identity in a Digital Age

It is widely recognised that political trust is an important element in democratic, political systems which ensures their well-functioning and stability. Research has shown that trust impacts a number of social and political issues, attitudes and practices (Newton, 2014, p. 19–20). As explicated by Ken Newton, trust is associated with longevity and life satisfaction (Halliwell & Putnam, 2004), tolerance and willingness to help vulnerable groups within societies (Uslaner, 2002; Marien & Hooghe, 2011), civic participation (Uslaner & Brown, 2005) and low levels of crime (Putnam, 1993) (Newton, 2014, p. 19). Political trust is also a requirement for a well-functioning economy, the creation of jobs, and for political governments to be able to "plan and execute policies and deliver services." (United Nations, 2021, p. 1). In short, high levels of political trust safeguard against political volatility, instability, suspicion and disengagement among the public. A decline in political trust will therefore have adverse effects on the stability and well-being of democratic political systems.

This chapter understands political trust as the faith people hold towards their government (Abrams & Travaglino, 2018) and the "belief or working assumption that political actors or institutions will act in accordance with your interests or preferences even if you do not enforce it (Easton, 1975)" (Quilter-Pinner et al., 2021, p.3). What follows from these understandings of political trust, is that political trust is a highly relational concept as it implies a link between the political system and the electorate. As such, political trust can be defined as the confidence the public as a collective have in their political institutions and political leaders (Norris, 2011).

Meanwhile, national identity is an important factor of whether or not the public even perceives itself to constitute a meaningful group with common interests and values. National identity, the conviction and belief of the particularity and uniqueness of an imagined political community, in other words, a nation, that binds the members to each other in intricate webs, is thus a factor which can affect political trust (Anderson, 1983). National identity implies both a personal and collective connection and identification with the nation, it provides people with meaningful concepts of who they are and who they are not by situating them physically, legally, socially and emotionally to a nation or "a homeland" (Billig, 1995, p. 8; Cohen, 1996). As the case in Scotland shows, the rise of Scottish nationalism has brought the idea of the Scottish nation's distinction and uniqueness as well as its salience to the forefront. This has implications for how the Scottish public perceive and relate to their two respective governments, the UK government and the Scottish government because the two are often positioned in opposition to each other in nationalist discourse (Glass & MacKenzie, 2015; Jackson, 2020; McCrone, Kendrick & Straw, 1989). The influence of national identity can therefore manifest itself in stronger trust in the regional or national governments and a weaker support and trust in the broader UK government (Quilter et al., 2021). As argued in a report by the Institute for Public Policy Research (IPPR), national identity plays an important role in political trust:

> "In Britain's devolved nations, strong sub-state identity is associated with lower satisfaction with the way democracy works in the UK. […] in Scotland, those that feel more Scottish than British are substantially less satisfied with the state of UK democracy than those who feel more British than Scottish." (Quilter-Pinner et al., 2021, p. 20).

In the information age, the issue of political trust is further complicated by the fact that information has become commodified, and communication and information is today created and disseminated faster and wider than ever before. Power over communication and information has thus been widened, and the scope of different, often competing interpretations of political and social issues is currently vast. As mentioned in the introduction, digital technology is a double-edged phenomenon. On the one hand, it can provide new means for public political participation, engage a wide range of individuals and reach groups which are traditionally politically inactive, but it can also contribute to political incivility and tear down trust in the political establishment (Van't Riet & Van Stekelenburg, 2022). Digital technology in relation to political awareness, participation and in contributing to levels of political trust, is thus both a constructive and a destructive force (Wike et al., 2022, p. 5). The digital era of the unprecedented flow of information and misinformation, contrasting and conflicting narratives, political division, instability and volatility, poses a number of challenges to democracy, all of which have a direct impact on political trust, contributing to its decline (Eder et al., 2014, p. 1). Political trust in the digital age is therefore more precarious and more susceptible to challenges and threats, of which the UK is an apt example.

British and Scottish Political Trust

Levels of political trust in the UK are presently low, and lower than the OEC average (Policy Institute, 2023). A report from the Policy Institute shows that the British public's trust in parliament has halved since the 1990s when people's perceptions of the British parliament were "at their most positive" with 46% having confidence in parliament; by 2022 this had fallen to "a historic low of 23%" (Policy Institute, 2023, p. 4). Many agree that there has been a steady decline of trust and confidence in the British political system, including trust in politicians and institutions, a development beginning in the 1970s and accelerating in the 1990s (Curtice & Scholes, 2021; Perry, 2021). The British context is critical as levels of political trust in the UK are remarkably low, as argued by Quilter-Pinner et al.: "The evidence is clear that political trust has been declining over time in the UK and in many other countries, though the UK has lower levels of trust than many." (2021, p. 3).

Today, a large proportion of the British public does not have faith that their government and politicians have the public's best interest at heart. Instead, many distrust their politicians to be seeking personal gain. Recent developments such as Brexit and the Covid-crisis have added additional contextual layers to the decline of political trust in the British political system (Curtice & Scholes, 2021; Butt et al., 2022). According to the report by IPPR, only 6% of the 8000 respondents in the research survey hold full trust in the British political system (Quilter-Pinner et al., 2021). These figures are historically low and are a signal of a considerable change in the British citizen's perceptions and attitudes to their political system, establishing a new norm of distrust in politics (Quilter-Pinner et al., 2021, p. 3). In this context, populist parties and leaders have gained more support and power at the expense of established political parties. In the UK, this can be seen in the uprise of UKIP, but also in the increased support for Scottish independence as shown in the Scottish referendum of 2014. Disintegration of the UK seems to be a continual trend in which British citizens are not only more politically divided, but where regionalism and nationalism pose a serious challenge to the unity of the British state.

The development of decline in political trust is driven by a myriad of different factors which combined have an adverse effect on political trust. Factors undermining political trust are therefore many and often mutually reinforcing. The IPPR report has identified five main challenges currently threatening the UK. These five challenges are *populism*, *polarisation*, *participation*, *dissatisfaction* and *de-alignment* (Quilter-Pinner et al., 2021, p. 22). Although no single challenge alone can be said to be the reason for the overall decline in political trust in the UK, and in the West at large, they all play their part in destabilising and damaging political trust. For instance, populism, "tends to reflect a deep suspicion – or distrust – of the prevailing establishment." (Quilter-Pinner et al., 2021, p. 22). Polarisation's link to political trust is that it leads to general distrust in the governing political parties, making political consensus more difficult to achieve and creating obstacles for the enactment of policies (Quilter-Pinner et al., 2021, p. 22).

The factor of participation is the most contested among scholars around the question of whether low levels of trust give rise to political participation for anti-establishment parties, or whether low levels of trust are manifested in an overall low political participation, whether for establishment parties or anti-establishment. Either way, participation increases political distrust (Quilter-Pinner et al., 2021, p. 22). Meanwhile, dissatisfaction people feel towards their political systems and politicians lowers political trust by creating “an environment in which it is harder for politicians to succeed” (Quilter-Pinner et al., 2021, p. 22). Finally, low levels of political trust give rise to the challenge of de-alignment, which “undermines the formation of stable party preferences and thereby stimulates volatility” as well as contributing to the electorate to change their party loyalty and preference (Quilter-Pinner et al., 2020, p. 22). A decline in political trust is significant for two main reasons. Firstly, decline in trust affects and leads to political disengagement by the public, manifesting itself in low voting turnouts and in pushing voters towards populist political parties and alternative sources of political power, all of which weaken democracy (Quilter-Pinner et al., 2021, p. 3). Secondly, the decline in trust affects the government’s ability to deliver policies supporting and strengthening social progress (Quilter-Pinner et al., 2021, p. 3).

While political trust is in decline in the UK as a whole, in Scotland, trust in the Scottish government has been relatively stable for almost a decade (Reid et al., 2020; Scholes et al., 2022). The Scottish case is interesting for a number of reasons, not least because of the tension between autonomy and restriction that characterises Scotland’s constitutional position and public discourse. Scotland, a nation but not a state, and which is a part of the UK, has in recent years gained more political powers that have contributed to a wider political divergence between Scotland and the UK on a wide number of issues, including that of political trust. This materialises in popular convictions of a political divide and general divergence between Scotland and England, as the two nations increasingly adopt different approaches to societal issues such as seen in the handling of the Covid19 pandemic, but also in different, at times conflicting attitudes to key political issues, such as the UK’s exit from the EU as shown in the voting differences between the two countries in the EU referendum, as Scotland voted 62% in favour of remaining, and England voted 53.4% in favour of leaving (Butt et al., 2022, p. 1; BBC, 2016; Curtice et al., 2021, p. 3).

Latest findings from the SSA reveal that the Scottish public has higher levels of trust in the Scottish Government compared to levels of trust they hold towards the UK Government. The 2021–2022 findings show that 66% of the survey respondents answered positively to trusting the Scottish government either “just about all the time” or “most of the time” to act in Scotland’s best interests (Scholes et al., 2022, p. 13). In contrast, 22% answered positively when asked the same question (to act in the best interests of Scotland) about the UK government (Scholes et al., 2022, p. 13). 46% of the respondents said that they “almost never” trust the UK government to act in Scotland’s best interests, compared with 11% who “almost never” trust the Scottish government to do the same (Scholes et al., 2022, p. 13). On the question of trusting the UK and the Scottish governments, in addition to local governments to make fair decisions, the same tendency of higher trust in the Scottish government than the UK government can be seen. The number of respondents who either trust the Scottish Government "a great deal" or "quite a lot" on making fair decisions was 48%, compared to only 15% who trust the UK Government to do the same (Scholes et al., 2022, p. 13). 19% showed low trust in the Scottish government to make fair decisions, while 56% showed low trust in the UK government.

Political trust, as discussed earlier, is a relational concept in which the citizens and their government are in a mutual relationship of trust. Therefore, the question the SSA posed to the respondents of whether they felt listened to by their Scottish and UK governments is a significant marker of the level of trust in Scotland. To this question, 58% answered that the Scottish Government was either "very" or "quite good" at listening to the public before making political decisions, compared to 18% who answered the same in regard to the UK government (Scholes et al., 2022, p. 18). Another key finding in the SSA 2021-2022 was that the Scottish Government was perceived favourably in terms of "giving Scotland a stronger voice in the UK" with 68% answering affirmatively that the Scottish government gives Scotland a stronger voice in the UK, 28% answering "no difference" and 7% stating that the Scottish government gives Scotland a weaker voice in the UK (Scholes et al., 2022, p. 21). Perhaps not surprisingly then, that on the question of who should run Scotland, the Scottish Government or the UK Government, 75% answered that they want the Scottish Government to have the most influence in how Scotland is run compared with 14% who expressed that the UK Government ought to have the most influence in how Scotland is run (Sholes et al., 2022, p. 23).

The SSA findings are clear in that the position of the Scottish government is currently strong in Scotland, and that political trust is considerably higher in the Scottish government than it is in the UK government among the Scottish respondents. The respondents are of course a varied group whom the SSA authors have analysed in order to find correlations along several variables, including support for Scottish independence, political affiliation, and dominant values of the respondents. The people in support of Scottish independence and the political parties, the SNP and the Scottish Green Party, showed the highest level of support towards the Scottish government, while the supporters of Scotland's membership in the UK and political affiliation with the Conservative Party showed the least level of support for the Scottish Government but the highest for the UK Government (Scholes et al., 2022, p. 14). These statistical findings are an important source of knowledge of the political situation in Scotland and they offer a unique insight into the public's levels of trust in the Scottish political system, however, they should be seen in relation to an in-depth analysis of the political shifts and developments in Scotland which have led to this position today.

Scottish Referendum as "an Electoral Shock" and Empirical Data

A study discussing the recent political changes in the UK, and in particular the democratic challenge of de-alignment as presented by Quilter-Pinner et al. (2021), have similarly explored what the authors call the "five electoral shocks" that are argued to have changed British politics dramatically (Fieldhouse et al., 2020). The five shocks in recent British politics are 1) rise in immigration post 2004, 2) the financial crisis of 2007–2008, 3) the coalition government between the Conservatives and the Liberal Democrats (2010-2015), 4) the Scottish independence referendum of 2014, and 5) the EU referendum of 2016 or Brexit (Fieldhouse et al., 2020, p. 1). An electoral shock is defined as "an abrupt change to the status quo. […] They represent a significant and often unanticipated change." (Fieldhouse et al., 2020, p. 2). Electoral shocks are, furthermore, noticeable for all, are difficult to ignore, and what is more, "have the potential to change how parties are perceived and therefore to reshape the party system." (Fieldhouse et al., 2020, p. 2). The ability to change political systems, influence and reach a wide range of the electorate are the defining features of electoral shocks (Fieldhouse et al., 2020). As such, electoral shocks are not short-lived abruptions which end as quickly as they arrive, and after which, things go back to normal, and the status quo is unaltered but rather restored. To the contrary, what makes electoral shocks politically salient is their longevity and long-lasting impact on the society and political landscape (Fieldhouse et al., 2020).

The Scottish independence referendum is such an electoral shock which has significantly changed the political landscape of Scotland. The referendum on Scottish independence held on 14 September 2014 has been described as a decisive turning point in Scottish politics, marking a new era in Scottish society (Jackson 2020). Michael Keating described the referendum as, "[…] a momentous event, which engaged the political class and the general public like few political events before it." (2017, p. 1). Indeed, the consequences of the referendum have been wide-reaching, as the referendum has destabilised the status quo and offered alternative political alignments and attachments to the Scottish public (Fieldhouse et al., 2020). Many of my informants during the 2018 fieldwork in Scotland[1] have expressed views that underscore the importance of the referendum, as conveyed by one informant who said the following of its impact:

> "Like, I suppose that people got involved not just in campaigning for the referendum, but they got involved in social activism; you're going leafleting in working class areas, getting people to vote for… then you start talking about other issues and people got involved in a whole host of stuff, so really it was a political awakening for a lot of people. And being Scottish and being a campaigner for Scottish independence also meant that you are a political activist in a whole host of other things." (Berg, 2021, p. 260).

1 Fieldwork was conducted primarily in Glasgow with shorter stays in Aberdeen and Inverness.

The referendum was presented by many informants as a unique opportunity to bring forth radical societal and political change, particularly among the younger generations, as reflected on by one middle-aged informant: “People became energised, especially the younger age group – they were around 70% in favour [of Scottish independence].” (Berg, 2021, p. 261). An informant in his late 20s expressed a similar sentiment and a first-hand experience of what the younger generation desires:

> “And so for me, for my generation, I think we’re quite sick of it, we want to start setting our own course. We’re the generation that voted 60% in favour of Europe, maybe even more, 65% in favour of independence. That’s the future we want, we want our future to be progressive, to be guided by fairness and an equitable society.” (Berg, 2021, p. 260–261).

As an electoral shock with extensive, far-reaching and long-lasting impact, the referendum on Scottish independence was an unparalleled event for political participation. The turnout rate for the referendum reached “the highest ever recorded number for a vote in Scotland” with 84.6% (Fieldhouse et al., 2020, p. 139; Electoral Management Board for Scotland, 2014). This historic high turnout is not only due to the core issue of Scottish independence being proposed as a viable possibility, but also because, during the period leading up to the referendum, ordinary people got involved in the Yes-campaign. Ordinary people engaged in the political question of Scottish independence and thus in politics itself in a form that has later been described as a “democratic revolution” (MacWhirter, 2015). The Scottish journalist Ian Macwhirter commented on the effects of the high public participation in the 2014 referendum and reflected that despite the fact that independence supporters did not win the vote, the political participation itself had sustained post-referendum:

> “The civic engagement that fuelled the referendum campaign and motivated 97% of Scots to register to vote was still there. Scottish voters had been told after the No vote that they had to get back in their box, pack away the festival of democracy, and return to the normality of boring responsible politics. But the people refused to get back in their box […] People are still coming out in their thousands to talk politics in pubs, town halls, theatres, and book festivals. I’ve spoken to many of them myself, and it is difficult not to be infected with their enthusiasm. Armed with the Internet, seized by a sense of communal purpose, the people of Scotland refuse to believe that a better society is impossible.” (MacWhirter, 2015, p. 24).

My empirical findings also show that many in Scotland felt that the referendum was reinvigorating the Scottish public by offering them hope and agency, and an opportunity to directly impact their political system (Berg, 2021). An opportunity which was expressed as strongly welcome after a long period of feeling political disenchantment, as one informant put it: “The Scottish Government is trying to engage the public and they are doing that in a way where they are reducing the feeling of worthlessness after the Thatcher era.” (Berg, 2021, p. 261). Many of my informants made statements that described the referendum as having dramatically altered not only Scottish politics but also Scottish nationalism:

> "Scotland feels very much like a new nation now… that it's been reborn. I think so. Well, yeah… I think it's a forward-looking country, whereas it used to be perhaps quite constrained, backwards-looking country, but now I think… there is a desire and particularly among the young groups, particularly with things like independence, they make this a very forward-looking nation." (Berg, 2021, p. 260).

Digital media and digital tools have played a key part in enabling and spreading this sense of hope and agency. People took to social media to participate in political discourse and some formed organisations and groups supporting the Scottish independence movement. Examples of such groups and organisations include Bella Caledonia and Common Weal, as well as the establishment of the pro-independence national newspaper called *the National* in November 2014, two months after the referendum took place. In addition, many more individuals were active in the Yes-campaign on social media, particularly on Twitter (now X). The grassroots and bottom-up structure of these groups and organisations were a substantial element in creating a mass social movement in Scotland in favour of independence, and they were a significant part of the Yes-campaign. Although much is written on the negative consequences of digital technology and social media on political trust and established politics, digital platforms have also the ability to create political participation and engagement, and thus a stronger democracy.

The 2014 Scottish independence referendum is thus one of the five electoral shocks in British politics of the 2010s (Fieldhouse et al., 2020). Its effects are widespread, wide reaching, profound and unprecedented. Specifically, the 2014 referendum changed political alignments and loyalties in Scotland, creating a new political scene, which in many ways marks a break with the previous order. Examples of the political change caused by the referendum are the role and status of the Labour party in Scotland. Once the majority party and undisputed political force, Scottish Labour was nearly wiped out of the Scottish political landscape in the 2015 general election which occurred a year after the referendum: "The referendum campaign seemed to damage Labour's popularity, which continued to erode right through to the 2015 General Election, which delivered Labour's worst defeat in Scotland since 1918." (Fieldhouse et al., 2020, p. 140) In the 2015 general election, the Scottish National Party (SNP) won its landslide victory winning 50% of the vote and securing 56 out of 59 Scottish seats (Scottish Parliament, 2015). This was an increase in representation at Westminster from 6 seats to 56 seats, a historical rise (Scottish Parliament, 2015). It was also the first time a single-issue party and a non-establishment party had ever achieved this level of support in British political history. The SNP has historically been seen as a fringe party and an unlikely candidate for a majority party governing Scotland (Hassan, 2009). Indeed, as Gerry Hassan notes:

> "The SNP has, in the last 40 years, moved from being a marginal force often ridiculed, patronised and caricatured by opponents to a force which is both respected and feared, and which has defined and reshaped Scottish politics, brought the Scottish dimension centre stage and forced other political parties to respond on their terms." (2009, p. 1)

The SNP's success was also in part due to perfect timing, as the general election was held a year after the Scottish independence referendum, and offered a chance for the party's supporters and those voters who were undecided in the referendum to cast their vote anew. The effects were radical, as Scotland secured its position as the ruling party of Scotland and had swayed voters from Labour, a party which has historically been strong in Scotland. By destabilising party loyalties and restructuring the landscape of political parties in Scotland, the referendum of 2014 created new possible political identities and correspondingly, created new alternative voting patterns and behaviours (Fieldhouse et al., 2020). This effect of reshaping and restructuring political identities and loyalties can be seen as part of the wider UK trend of de-alignment where voters no longer stay as loyal to political parties and are much more likely to change support (Quilter-Pinner et al., 2021). Interestingly, the de-alignment in Scotland has created a more stable new landscape of political support where previous Labour voters have generally switched over to support the SNP.

Digital Tools and the 2014 Referendum

New opportunities to participate politically are emerging due to digital developments. Social media platforms were widely used in the run-up to the 2014 Scottish independence referendum and have arguably played a key role in mobilising the two fractions: the Yes-campaign in favour of independence and the No-campaign entitled *Better Together* in favour of Scotland's membership in the UK. Particularly the Yes-side were eager users of digital tools, and the digital platforms were seen as strengthening the Yes-campaign: "[…] digital media were an important platform in harnessing and providing a means of expression of the energy and enthusiasm of Yes supporters during the campaign and beyond it." (Langer et al., 2019). Political campaigns are increasingly using digital tools by combining traditional modes of campaigning with novel, digital arenas, and contemporary electoral campaigns are, as argued by Langer et al. (2019), "shifting towards hybridity, a combination of professional co-ordination of activity from above with greater bottom-up, decentralized participation from a range of loosely connected non-elite actors." The degree to which and the ways political campaigns combine the traditional modes of campaigning, that is, the bottom-up/decentralised model, or the top-down/command and control model, varies greatly. Investigating the differences between the two sides in the 2014 Scottish referendum, Langer et al. (2019) have found differences in approaches to, and the use of, digital technology in the two campaigns. The main difference is that the Yes-campaign had to a larger extent combined traditional, controlled campaigning with the use of digital tools in their campaign. The Yes-campaign was widely recognised for being a civic movement with participation by non-elite actors such as ordinary people, groups and organisations, gathered around the issue of advocating in favour of Scottish independence. In contrast, the No-campaign used digital tools to a much lesser extent and instead relied on the traditional top-down approach where the campaign was largely constructed and led by the politicians and officially approved actors (Langer et al., 2019).

Digital platforms and digital communication have both a democratic and an anti-democratic potential and effect. In theory, digital communication and particularly the role of social media can provide access to political engagement and involvement for people traditionally outside of the political establishment and as such provide an invaluable transformative potential of political action and democracy (Langer et al., 2019). More people can express their views, engage with others, influence policy makers, and provide a necessary critical voice to the political elite. On the other hand, social media and other digital platforms can have an adversary effect on democracy by dividing the public, making opinions more polarised and feeding individuals with tailored and biased content, thus creating an echo chamber effect in which "a person encounters only beliefs or opinions that coincide with their own, so that their existing views are reinforced and alternative ideas are not considered." (Oxford English Dictionary, 2023). However, digital tools can likewise have the potential of mobilising and inspiring people to take control over their political situation and directly participate in political, social and economic issues.

Conclusion

The years around the referendum on Scottish independence in 2014 are characterised by political momentum, novel political participation enabled by digital tools and social media in particular, and as importantly, by a widespread sense of collective hope and agency to change Scotland for the better. Social media and digital platforms have been used to organise, inspire, connect and engage the public in political issues by both grassroots actors such as the bottom-up organisations, individuals on social media as well as official actors such as the SNP and the Scottish government. The purpose has been to use social media and digital platforms to gain the Scottish public's support for the issue of Scottish independence and to increase the public's levels of trust in an alternative vision for Scotland (Berg, 2021). Although the results of the 2014 Scottish independence referendum yielded a no result for the independence supporters, the politically active period around the referendum seems to have had a long-lasting effect in positive perceptions of the Scottish Government's ability to successfully govern Scotland, as figures from the SSA reveal (Scholes et al., 2022). The grassroots organisations that sprung up in support of the Yes-campaign in the run-up to the independence referendum in 2014 have been presented as examples of *true de-*

mocracy, that is the active participation of the people in political issues directly affecting and regarding them. These organisations can contribute not only to increasing political agency among the public but also in rebuilding political trust, as these organisations are not associated with the political establishment which many in Scotland and the rest of the UK have grown suspicious of:

> "National Collective, the Radical Independence Campaign, Common Weal, Bella Caledonia and many more local under-the-radar groups, groups that are self-organised and crowd-financed through online donations and memberships – these groups sit outside the traditional political setup, yet it can be argued have had a growing influence on the feel and framing of the debate. Presenting Yes as much more than an elite project while helping to back No into that top-down corner." (Sullivan, 2014, p. 98).

The reason why these self-sprung and self-run organisations and groups were considered important was that they expressed a feeling of hope in the political future of Scotland as much as they supported the idea that the Scottish public has power to influence the direction in which Scotland is headed. A number of my informants saw a clear divide between the years before and after the Scottish referendum in regard to the Scottish public's optimism for the political future of Scotland and their own agency to affect important political, social, cultural and economic issues (Berg, 2021). Post-referendum, it seems that many more people have gained a renewed sense of hope and agency that they have the capacity and ability to influence and even change Scotland's political and constitutional position. As one of my informants said: "I think that's a… I suppose if you're grown up during the referendum, there's this whole new way to be Scottish, you know what I mean." (Berg, 2021, p. 260). Arguably, this "new way" of being Scottish is a shift in perceptions from feeling politically alienated and apathetic towards feelings of reinvigoration and hopefulness. Many of my informants expressed discontent with the UK government but felt that the Scottish government was much more able to uphold dominant Scottish values of social justice, egalitarianism and progressive political views. As one informant commented:

> "We're governed by a group that meet in a town that is so far away and has completely different values. And I just think we have different values! And I think perhaps we would [in the case of Scottish independence] end up less well-off because we would end up paying higher taxes and things, but to think that our tax money would be better spent […]." (Berg, 2021, p. 211).

While the UK government is perceived to have political and social values that differ from collective Scottish values, the Scottish government is, as has been shown by the SSA findings, largely perceived to act in accordance with these values. Perception seems to be as important as actual political performance and actions. This means that perceptions of politicians influence people as much as politicians' policies and actions. "In other words, it is not only important that politicians and institutions do the right thing, deliver on their promises, and perform well: they must also be perceived to do so. […] When it comes to determining levels of trust, identities, and perceptions matter." (Quilter-Pinner et al., 2020, p. 9). What does this mean for Scotland? If people feel that they have agency and impact on their political systems, then they are more likely to have trust in these political systems and the people who run the country. Participation can be seen as an antidote to political disenchantment and can in turn contribute to increasing political trust.

References

Abrams, D. & Travaglino, G.A. (2018). Immigration, political trust and Brexit – Testing an aversion amplification hypothesis. *British Journal of Social Psychology,* 57, 310-326. https://doi.org/10.1111/bjso.12233

Anderson, B. (1983). *Imagined Communities: Reflections on the Origin and Spread of Nationalism.* Verso.

BBC. (2016). *EU Referendum. Results:* https://www.bbc.co.uk/news/politics/eu_referendum/results

Berg, L. (2021). *If they can do it, why can't we? Popular perceptions of the Nordic region in Scotland.* [PhD. Thesis, University of Oslo]. https://www.duo.uio.no/handle/10852/93316

Billig, M. (1995). *Banal Nationalism.* Sage.

Butt, S., Clery, E., & Curtice, J. (Eds.). (2022). *British Social Attitudes 39: Broken Britain?* National Centre for Social Research. https://natcen.ac.uk/publications/british-social-attitudes-39-broken-britain

Cohen, A. (1996). Personal Nationalism: A Scottish View of some Rites, Rights, and Wrongs. *American Ethnologist,* 23 (4), 802–815. https://doi.org/10.1525/ae.1996.23.4.02a00070

Curtice, J., & Shoules, A. (2021). *British Social Attitudes 38: Democracy.* The National Centre for Social Research. https://natcen.ac.uk/publications/british-social-attitudes-38

Devine, D., & Valgarđsson, V. O. (2023). Stability and change in political trust: Evidence and implications from six panel studies. *European Journal of Political Research:* https://doi.org/10.1111/1475-6765.12606

Easton, D. (1975). A re-assessment of the concept of political support. *British Journal of Political Science,* 5(4), 435-457. https://doi.org/10.1017/S0007123400008309

Eder, C., Mochmann, I., & Quandt, M. (2015). *Political trust and disenchantment with politics: International perspectives.* Brill. https://doi.org/10.1163/9789004276062

Electoral Management Board for Scotland. (2014). *Result of the Scottish Independence Referendum 2014.* https://www.emb.scot/scottish-independence-referendum-2014-3/scottish-independence-referendum-18-september-2014

Fieldhouse, E., Green, J., Evans, G., Mellon, J., Prosser, C., Schmitt, H., & van der Eijik, C. (2020). *Electoral shocks.* Oxford University Press. https://doi.org/10.1093/oso/9780198800583.001.0001

Glass, B. S., & MacKenzie, J. M. (Eds.). (2015). *Scotland, Empire and Decolonisation in the Twentieth Century.* Manchester University Press. https://doi.org/10.7765/9781784992255

Halliwell, J.F., & Putnam, R.D. (2004). The social context of well-being. *Philosophical Transactions of the Royal Society B*. http://dx.doi.org/10.1098/rstb.2004.1522

Hassan, G. (Ed.). (2009). *The modern SNP: From protest to power*. Edinburgh University Press. https://doi.org/10.1515/9780748642113

Jackson, B. (2020). *The case for Scottish independence: A history of nationalist political thought in modern Scotland*. Cambridge University Press. https://doi.org/10.1017/9781108883733

Kalogeropoulos, A., Rori, L., & Dimitrakopoulou, D. (2021). 'Social media help me distinguish between truth and lies': News consumption in the polarised and low-trust media landscape of Greece. *South European Society and Politics,* 26, 109-132. https://doi.org/10.1080/13608746.2021.1980941

Keating, M. (Ed.). (2017). *A wealthier, fairer Scotland*. Edinburgh University Press. https://doi.org/10.1515/9781474416443

Langer, A.I., Comerford, M., & McNulty, D. (2019). *Top-down or bottom-up? Campaigns, social media, and the Scottish independence referendum*. London School of Economics and Political Science: https://blogs.lse.ac.uk/politicsandpolicy/social-media-campaigns-for-indyref/

Macwhirter, I. (2015). *Tsunami: Scotland's democratic revolution*. Freight Books.

Marien, S., & Hooghe, M. (2011). Does political trust matter? An empirical investigation into the relation between political trust and support for law compliance. *European Journal of Political Research*. https://doi.org/10.1111/j.1475-6765.2010.01930.x

McCrone, D., Kendrick, S., & Straw, P. (1989). *The making of Scotland: Nation, culture and social change*. Edinburgh University Press.

National. (2023). Nicola Sturgeon: Social media is damaging the political debate. https://www.thenational.scot/news/23399994.nicola-sturgeon-social-media-damaging-political-debate/

Newton, K. (2014). Trust and political disenchantment. In Eder, C., Monchmann, I.C., & Quandt, M. (Eds.). *Political trust and disenchantment with politics: International perspectives,* (pp. 19-30). Brill. https://doi.org/10.1163/9789004276062

Norris, P. (2011). *Democratic deficit: Critical citizens revisited*. Cambridge University Press. https://doi.org/10.1017/CBO9780511973383

Oxford English Dictionary. (2023). *Echo Chamber*. https://www.oed.com/dictionary/echo-chamber_n?tab=factsheet#995958804

Perry, J. (2021). *Trust in public institutions: Trends and implications for economic development*. United Nations. https://doi.org/10.18356/27081990-108

Policy Institute. (2023). *Trust in trouble? UK and international confidence in institutions*. The Policy Institute, Kings College London. https://doi.org/10.18742/pub01-126

Putnam, R. D. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton University Press. https://doi.org/10.2307/j.ctt7s8r7

Quilter-Pinner, H., Statham, R., Jennings, W., & Valgarđsson, V. O. (2021). *Trust Issues: dealing with distrust in politics*. Institute for Public Policy Research. https://www.ippr.org/articles/trust-issues

Reid, S., Montagu, I., & Scholes, A. (2020). *Scottish Social Attitudes Survey 2019: Attitudes to government, the economy and the health service, and political engagement in Scotland*. ScotCen Social Research/Scottish Government: https://www.gov.scot/publications/scottish-social-attitudes-2019-attitudes-government-political-engagement/

Scholes, A., Montagu, I., Whitford, R., Daniels-Creasey, A., & Dean, L. (2022). *Scottish Social Attitudes 2021/2022*. Scotcen Social Research & Scottish Government. https://www.gov.scot/news/scottish-social-attitudes-survey-2021-2022/

Scottish Government. (2024). *What the Scottish Government does:* https://www.gov.scot/about/what-the-government-does/.

Scottish Parliament. (2015). *How did Scotland Vote – UK General Election 2015.* https://archive2021.parliament.scot/ResearchBriefingsAndFactsheets/S4/How_did_Scotland_Vote_-_UK_General_Election_2015.pdf

Stein, J., Broderstad, S. T., & Bjørnå, H. (2022). Territorial reforms, mobilisation, and political trust: a case study from Norway. In *Local Government Studies*. https://doi.org/10.1080/03003930.2021.2025360

Sullivan, W. (2014). *The Missing Scotland: Why over a million Scots choose not to vote and what it means for our democracy.* Luath Press.

United Nations. (2021). Trust in public institutions: Trends and implications for economic development: https://www.un.org/development/desa/dspd/2021/07/trust-public-institutions/

Uslaner, E. M. (2004). *The moral foundations of trust.* Cambridge University Press. https://doi.org/10.1017/CBO9780511614934

Uslaner, E. M., & Brown, M. (2005). Inequality, trust, and civic engagement. *American Politics Research,* 33 (6), 868-894. https://doi.org/10.1177/1532673X04271903

Van't Riet, J., & Van Stekelenburg, A. (2022). The effects of political incivility on political trust and political participation: A meta-analysis of experimental research. In *Human Communication Research,* 48, 203-229. https://doi.org/10.1093/hcr/hqab022

Wike, R., Silver, L., Fetterolf, J., Huang, C., Austin, S., Clancy, L., & Gubbala, S. (2022). Social media seen as mostly good for democracy across many nations, but US is a major outlier. Pew Research Center. https://www.pewresearch.org/global/2022/12/06/social-media-seen-as-mostly-good-for-democracy-across-many-nations-but-u-s-is-a-major-outlier/

Trust and Betrayal in the Virtual and the Real World: An Approach to *Nefando* (2016)

Wladimir Chávez

"What is the point of technology if not to narrate our horrors?"
(Ojeda 107).

Introduction

This study analyzes the novel *Nefando* [Heinous], written by Ecuadorian Mónica Ojeda, and focuses on the way dark places and disgraceful actions could be portrayed both in reality and in the digital society, in addition to evaluating the levels of trust that the characters feel towards authority and institutions. The novel revolves around an online video game from the deep web: *Nefando, viaje a las entrañas de una habitación* [Heinous, Journey into the Bowels of a Room], well-known to a restricted circle of people and extremely disturbing for its controversial plot. The behavior of the protagonists and the places mentioned in *Nefando,* both in the virtual and the real world, are scrutinized mainly from the perspectives of scholars like Sue Tait, who analyzes the fascination for violence in the digital era; Tahmina Ferdous and Chowdhury Abdullah Al-Hossienie, specialists in topics such as trust and distrust in government; Kimberly Christopherson, whose work highlights the consequences of anonymity and autonomy on the internet; and Robert W. Gehl, who reflects over the concept of trust in the deep web through the notion of legitimacy.

Nefando – A Context

Author Mónica Ojeda was born in Guayaquil (Ecuador) in 1988. Her first novel, *La desfiguración Silva* [The Silva Disfigurement], won the international Alba Narrativa Award in 2014. Three years later, Ojeda was included in the "Bogotá39" list, a catalogue of 39 writers under 39 years old who allegedly best represent the current Latin American narrative. *Nefando*, her second novel, was published in 2016 by Spanish publishing house Candaya.[1]

Nefando portrays the lives of six young people who share an apartment in Barcelona. Three of them are siblings from Ecuador: Irene, Cecilia and Emilio Terán. They are registered at the university, although they never attend classes. Also, in that apartment lives Kiki Ortega, a Mexican writer who is in Spain thanks to a scholarship; "Cuco" Martínez, a hacker who works with a group of thieves who

1 *Nefando* was translated into English by Sarah Booker and published by Coffee House Press on October 24, 2023. However, at the time of writing this article, the translation was not yet available. Therefore, all translations of the novel are mine.

steal wallets from tourists; and Iván Herrera, who is studying for a master's degree in creative writing at the university. Each of the characters has experienced different forms of violence, but the story focuses on the Terán siblings. With the help of Cuco Martínez, who in addition to being a hacker is a designer and programmer, the Terán siblings create a video game to post on the deep web: *Nefando, a Journey into the Bowels of a Room*. The game is disturbing, and those who take the risk of playing it and are patient enough to reach upper levels can see videos of animal torture, images of people mutilating themselves, and lastly, videos of the rapes suffered by the Terán siblings themselves: their own father was responsible for this crime.

Spanish writer and journalist Alberto Olmos has pointed out that *Nefando* is "a brilliant and sick novel about the 'deep web', [2] pedophilia and the passion for literature" (Olmos, 2016. All translations are mine). Argentine writer Mariana Enríquez has said that this literary work is "risky and at times unbearable (in a good way)" (Abdala, 2019). Regarding her own novel, Ojeda has pointed out that she was terrified of entering the world of *Nefando* and child pornography because the book "works on the horror that is outside of you, the horror of what human beings are capable of doing to each other. *Nefando* implies total darkness and I was there groping my way" (Arenas, 2019).

Endless Violence: The Virtual World as an Extension of the Real One

The Terán siblings arrive in Spain with a scholarship to study literature at the University of Barcelona and during their stay, they become interested in reading science fiction books. People don't usually see them as individuals with their own characteristics, rather as a single group: "For me they didn't exist individually. I know it sounds scary, I know, but deep down I think they didn't like seeing themselves separately either" (Ojeda, 2016, p. 68). However, their neighbors gradually get to know them. Irene Terán is the older sister. She talks a lot and expresses herself openly on difficult issues. Emilio Terán is a bit shyer; he writes constantly, it is a habit that he has maintained since he was a teenager. Cecilia Terán, on the other hand, is the most introverted of the three, takes anxiolytics and has suffered crises in which she harms herself. The brothers carry on their shoulders the weight of a terrible memory: the constant rapes committed by their father.

The Terán siblings turn to "Cuco" Martínez,[3] who studied video game design at the university but never finished his degree. "Cuco" Martínez, a hacker expert

---

2 Deep Web is "is an umbrella term for parts of the internet not fully accessible using standard search engines such as Google, Bing and Yahoo. The contents of the deep web range from pages that were not indexed by search engines, paywalled sites, private databases and the dark web" (Patrizio, n.d.).

3 "Cuco" is a nickname related to a mythical creature. "El Cuco is used to frighten children. The stories of El Cuco are used in hopes of keeping the young ones off the streets late at night or to make them go to sleep". (The Guide to the Colonial Zone and the Dominican Republic, n.d.).

in the deep web, considers himself a free and creative person. He helps the Terán siblings to design the video game because "It is not possible to say 'no' to victims like them" (Ojeda, 2016, p. 22). At this point, however, it is legitimate to ask: what do they obtain by posting these images? Who will be the spectators and what kind of effect can they cause? For researcher Sue Tait, there are several consumers of videos and images[4] like this, although traditionally only two large communities have been considered for studies: first, those who like and feel aroused by pornography. In this context, pornography does not refer exclusively to images of a sexual nature, but involves unedited graphic illustrations of any kind, disturbing or violent. Pornography here is a term that "has been mobilized to express a range of concerns around the corruption of empathy occasioned by atrocity and its representation" (Tait, 2008, p. 95). The second type of viewer wants these images to be seen in all their rawness in order to achieve social change. According to this perspective, the mainstream media unnecessarily sanitizes photos and films where bodies suffer or are shown torn to pieces. If a TV program talks about war, for instance, it should show graphically the real consequences of war, and because of that "images of atrocity are understood to make a call to conscience, to enable the viewer to bear witness to scenes cleansed from mainstream media through repressive standards of taste and decency" (Tait, 2008, p. 93).

At the same time, the choice to use the deep web, and specifically the dark web[5], to release a type of material impossible to find in the mainstream media, is not a coincidence: "The term Dark Web very likely evokes some decidedly illegitimate associations: drug markets, unregulated guns for sale, child exploitation images, stolen credit cards for sale, or phishing attacks" (Gehl, 2018, p. 2). In the imagined idea of what the internet is, the dark web is considered a virtual paradise of vices frequented by criminals.[6]

Nefando's book cover shows two aspects that will be developed during the novel's plot: the lack of empathy and an overwhelming feeling of suffocation. The illustration was made by artist Federico Pineda and shows the partial face of a girl who, with her eyes closed, appears to be sinking into a pool. The photograph is related to a passage in the book in which the father Terán wants to teach his daughter Irene, who was 8 years old at the time, to swim. The father throws her into the pool, so she starts to splash around on her own. As she begins to sink, he pulls her out, makes her vomit up the swallowed water, and throws her back into the pool. "The father laughed to encourage her to learn through ridicule" (Ojeda, 2016, p. 74). That feeling of drowning is experienced by Irene and her siblings

4 In addition to taking in consideration the two more common groups (pornographers and social change seekers), Tait proposes other viewers of body horror: "I have identified amoral, vulnerable, entitled, and responsive gazes" (Tait, 2008, p. 107).

5 Dark web "is an encrypted portion of the internet that is not indexed by search engines and requires specific configuration or authorization to access" (Terrell Hanna, n.d.).

6 In reality, the dark web is much more complex: "In more generous interpretations, many Dark Web sites might be judged as valuable forums of personal and political expression, allowing political dissidents to express their views without fear of government reprisal, or enabling people to socialize without fear of corporate surveillance" (Gehl, 2018, p. 7).

every time their father is near them, no matter if they are inside or outside the pool: "Perhaps swallowing water was necessary to conquer the fear; after all, to avoid truly drowning, one must first pretend to drown" (Ojeda, p. 76). That sensation of physical and psychological suffocation, in the case of the Terán brothers, began in their own home.

In *Nefando* the defects of the real world extend to the virtual one. In neither of them does it seem possible to find a balance or "breathe without drowning", as is implied in the illustration of *Nefando*'s book cover. In the real world, Kiki Ortega thinks that Barcelona is "a dump, just like Mexico D.F." (Ojeda, 2016, p. 9), and her own "refuge-room" is like a basement with a smell of humidity and sweat that makes her close her eyes (Ojeda, pp. 10-11). One day "Cuco" Martínez was beaten up in the street and one of the Latin American characters says "I thought that these things did not happen here because it is the first world and everyone is happy" (Ojeda, p. 92). But the virtual world is no different in its flaws either. Scholar Robert W. Gehl states that "we have to attend to questions of government, corporate, and social power practices as we consider any networked technology" (Gehl, 2018, p. 3). Both on the dark web and on the World Wide Web, *Nefando*'s characters only encounter a chaotic world, just like in everyday life. On one occasion, "Cuco" Martínez himself reflected on the coexistence of real space and the virtual one:

The internet that we know is full of places, languages, territories, and is, in itself, an alternate world. The funny thing is that, deep down, we do not reinvent anything in this new world. We have this powerful tool, this parallel space that, in principle, should be ideal as long as it is completely controlled by us, its creators, but it has the same flaws as the physical world, or if you prefer, the real one. (Ojeda, 2016, p. 70).

"Cuco" Martínez himself perpetuates crime in real life and in the digital world. He is a member of the criminal band Unión Xenófoba Antituristas [Xenophobic Anti-Tourist Union], which robs foreigners walking along the Barceloneta beach or La Rambla. "Cuco" Martínez takes the credit cards to clone them, transfer money or sell the data and information in the dark web. However, he is also a complex character. After he breaks up with his girlfriend Lola, he is contacted by a television reality show that "uses people's private lives to record a dramatic show in which they make everyone believe that it's cool to turn the problems of a person into an amusement park" (Ojeda, 2016, p. 103). The TV show wants to do an episode about him and his ex-girlfriend. In the end, el "Cuco" refuses to participate. It would seem strange that "Cuco" distances himself from a reality show, but agrees to post videos on the dark web. However, the difference lies in the nature of the footage. Tait states that "The subjects of 'mainstream' pornography generally 'consent' to their appearance on screen. Their roles are performed for the camera. The same cannot be said of the victims depicted in documentary footage of death and suffering" (Tait, 2008, p. 96). In any case, "Cuco" does not have a pornographic perspective. For him, the "anti-censorship position organizes the desire to see the 'real' as rational" (Ojeda, 2016, p. 104). He seeks a social, anarchist and anti-system change, therefore he helps the Terán siblings: the videos show a personal and social problem.

The Role of the State: A Betrayal of the Authority Figure

Criminality activity occurs both on the internet and in real life, and sometimes the roles of victims and perpetrators are confused. The characters in *Nefando* make decisions that may have legal consequences. For example, the Teráns were going to be investigated by the police, although they flee Spain before they can be questioned. The police are interested in knowing the origins of the images in the videogame, but above all they want to arrest the Terán siblings because they maintain a web page called *Proyecto Cratos* [Project Kratos[7]], virtual space where they share pirated versions of movies and scanned books. Thanks to *Proyecto Cratos*, the Teráns earned a lot of money. Authorities are more focused on punishing copyright violations than on discovering a sexual predator.

"Cuco" Martinez can't quite understand why the brothers have not yet reported to authorities that their father is a rapist. He supposes that they punish the father with blackmail, because they ask him for money very often. Iván Herrera, on the other hand, believes that the Terán siblings "learned to deal with the past in a peculiar way and that they did not consider themselves victims […] They also didn't hate their parents.[8] They despised them and used them, but they did not hate them. Or maybe they knew how to hide all their hatred very well and live outside of it, which is good" (Ojeda, 2016, p. 134). But there is a third, much more likely option: the Teráns distrust the authorities.

In general, *Nefando*'s characters show little trust in the government and its institutions: "To inculpate government has become a recent trend worldwide due to the eroding degree of confidence in representatives and public sector institutions" (Ferdous & Al-Hossienie, 2019, p. 1). The fact that *Nefando*'s characters resort to violence to express themselves or to obtain money shows their lack of conformity with a social organization that seems to have forgotten them. After losing faith in the authorities, they take alternative paths of personal development that may clash with the law. Lack of trust can have different consequences:

> Trust is a social, economic and political binding agent. A vast research literature on trust and "social capital" documents the connections between trust and person-

---

7 According to Greek mythology, Kratos was "a personification of brute Strength or Power" (Greekmythology.com, n.d.).

8 Despite the fact that it is mainly the father's criminal and hypocritical attitude which is denounced by the siblings in *Nefando* – the father makes documentaries and receives film awards while he regularly rapes his children – the mother is also condemned for her passivity; She suspects that something strange is happening but does nothing: "My mother always looked at us from a sharp corner. She knew what daddy did to us. She knew how to read the traces" (Ojeda, 2016, p. 126). In fact, the Terán siblings do not believe that their mother loves them (Ojeda, 2016, p. 123).

> al happiness, trust and other measures of well-being, trust and collective problem solving, trust and economic development and trust and social cohesion. Trust is the lifeblood of friendship and caregiving. When trust is absent, all kinds of societal woes unfold – including violence, social chaos and paralyzing risk-aversion (Rainie & Anderson, 2017).

Since the Teráns have little trust in the institutions, they never sought help from either the Ecuadorian or the Spanish authorities. "Distrust is considered healthy because posing a skeptical attitude towards the governance might alert the government to bring rectification and reformation in public institutions and administration" (Ferdous & Al-Hossienie, 2019, p. 14). However, this is not a "healthy distrust": the characters in the novel do not believe in the very functioning of institutions. They have good reasons for that. As soon as the Spanish police realize that the victims in the videos of the rapes are the Teráns themselves, the investigation is archived. The only action taken by police was to make sure that *Nefando, viaje a las entrañas de una habitación* stops circulating on the internet.[9] The disturbing videos, those of animal torture and of people mutilating themselves, have protagonists who cannot be identified, in addition to the fact that they have been on the deep web for years. That material will remain accessible because police are unwilling or unable to do anything about it. The videos of the rapes of the Terán siblings are also old, so the authorities do not even contact Ecuadorian police.

There is a weak image of the civil authority in other passages of *Nefando*: police have no control over certain neighborhoods of Barcelona nor over what happens on the internet. This is a process in which "[…] distrust may drive the citizen discontent and apathetic to the political process entails abating confidence to the regime, to the public institution and to the government eventually" (Ferdous & Al-Hossienie, 2019, p. 2). The figure of the state, incapable of guaranteeing the rights of citizens on the streets, proves much less effective in the World Wide Web. Researcher Robert W. Gehl points out that the dark web has "fraught relationships with the state, the legitimated holder of the monopoly on violence" (2018, p. 3). As a result, the dark web can offer trusted audiences for those who do not believe in the role of state institutions.

Trust on the Internet: The Legitimacy of the Peers

Since the Terán siblings could not trust[10] their parents or the rule of law, their options were few. "Trust is less relevant when there is no need to develop loyalty because there are no alternatives" (Rainie & Anderson, 2017). After authority figures have failed in their tasks, the Terán siblings seek refuge in the anonymity offered

9 According to Gehl, "Dark Web sites are notoriously ephemeral, appearing online for a few months and disappearing without a trace" (2018, p. 18).

10 Trust and distrust are important issues in *Nefando*. "Cuco" Martínez states that: " (…) for some of us the hostility in the world started at home. Do you know what I tell you? There are small first devastations that configure you. I do not know. The moment you lose trust." (Ojeda, 2016, p. 201).

by the internet.[11] For the kind of story they want to share, it is important that the victims are not easily identifiable.

> Though many studies have focused on the negative effects of anonymity, such as increases in aggressive behavior (Zimbardo, 1969), encouraging suicidal individuals to follow through with their threats (Mann, 1981), and other anti-normative and potentially anti-social behaviors, there has also been research examining the positive effects of anonymity. The most notable means by which anonymity can be positive is the importance for privacy on psychological well-being. Privacy refers to the ability to control the amount of contact one has with others (Christopherson, 2007, p. 3040).

The Terán siblings control the narrative of their own story when they interact with others on the dark web. It would be different if they had asked for help from the legal authorities. As they expect nothing from the democratic institutions nor from the state, the Terán siblings find comfort among online users on the dark web. They try to gain the trust and legitimacy of their peers, which are important concepts:

Indeed, legitimacy can be a powerful window into the Dark Web (…) All use variations on the word "legitimacy" to describe Dark Web sites, practices, and technologies. They call certain things legitimate and others illegitimate (or, to use the parlance of the Tor- and I2P-based Hidden Answers site, "legit or sh!t"). At the core of this discernment is a trial of legitimacy, where the Dark Web's uses and meanings are under intense scrutiny by a range of social groups. This trial is more complex than a stark illegitimate/ legitimate dichotomy: "legitimacy" is a highly context-dependent term, with many shades of meaning and interpretation. Along the way, declarations of something being "legit," in contrast to other, illegitimate things, mark moments of power practices (Gehl, 2018, pp. 2-3).

These power relations seem to interest the Teráns, who show online the moments of terror experienced by them. Their narrative is accepted as horrendous, but legitimate. Because of that, it seems to produce in them the positive benefit of "psychological well-being", as previously noted by scholar Kimberly Christopherson. As expected, such a positive effect would not have occurred if they had reported the crime to the authorities. In addition, the advantage of anonymity "could be used strategically for other purposes (e.g., to engage in inappropriate social behavior without fear of reprisal)" (Christopherson, 2007, pp. 3052-53). Certainly, the act of posting a video on the internet with explicit sexual violence results not only in an extreme form of social denunciation, but also in an inappropriate social behavior. At the same time, this act shows a level of deep autonomy:

---

11 In this regard, scholars Lee Rainie and Janna Anderson states that certain digital interactions "have also coincided with a sharp decline in trust for major institutions, such as government (and Congress and the presidency), the news media, public schools, the church and banks" (Rainie & Anderson, 2017).

> The final factor that was associated with anonymity is autonomy. Autonomy involves the chance to experiment with new behaviors without fear of social consequences. Individuals can use their anonymity to almost become a different person without fear of being identified and negatively evaluated by those they know. This factor may lead to an extreme sense of freedom for the individual and allow him or her to engage in behaviors typically disapproved of by others without fear of the consequences that may ensue as a result (Christopherson, 2007, p. 3041)

The Teráns engaged in an inappropriate social behavior thanks to the help of anonymity. Only after they posted evidence on the dark web of the abuse they suffered, was their experience understood as "legit". It is important to emphasize that their decision produced more horror than empathy among the online community. In any case, a fact considered "legit" on the internet does not necessarily have to arouse sympathy. Rather, something "legit" is a sort of guarantee given by users that the event is "real". Scholar Sue Tait has stated that images of atrocity could be a call to conscience to achieve social change. The denunciation of a horrendous and "legit" crime, certainly, could help achieve that goal. Due to the lack of interest of the state and its institutions, the Teráns are forced to tell their personal story in the virtual world, a place where they found anonymous peers who paid more attention than the police or authorities themselves.

Conclusion

Nefando portrays a very violent society. Such a violence goes from the unsafe and dirty streets of Barcelona and infects the virtual world. The Terán siblings have a horrific story to tell, but they distrust the authorities and their ability to deliver justice. "Such distrust [in citizen's mind] may drive the citizen discontent and apathetic to the political process entails abating confidence to the regime, to the public institution and to the government eventually" (Ferdous & Al-Hossienie, 2019, p. 2). They posted a disturbing videogame that caused different reactions in the viewers. According to Tait, "Body horror sites decontextualize the body in pain: the biography of the suffering subject is occulted and trauma is transfigured into imagery which stimulates, fascinates, or repulses the viewer" (Tait, 2008, p. 107). The repulsion generated by images of raw violence is a form of social denunciation as soon as the audience knows the story behind them. And the space for that complaint is not a police station, but the dark web, where the Teráns can find a public for their purposes.

In *Nefando* it seems that the digital society has taken the place of the real one, without necessarily becoming an improved version. "Cuco" Martínez remembers the mythological story of Cronus[12] when he thinks about the Teráns' father: "He was like Cronus, you know? By swallowing his children, he swallowed the notion

12 According to Greek mythology, Titan Cronus is warned that one of his children is going to dethrone him, so he devours them to prevent the prophecy from being fulfilled.

of tomorrow and all the days that followed with the only goal of perpetuating himself. Swallowing one's parents, on the other hand, is natural: it is not parricide, but communion; it is not the search to perpetuate oneself, but the only possible way to be born" (Ojeda, 2016, p. 205). In this context, "swallowing one's parents" is a metaphor that refers to cutting ties, abandoning the tutelage of authority and forging a path of one's own.

Although the Cronus' metaphor in *Nefando* could be understood as the digital society "swallowing" and adopting everything bad from the real one, perpetuating injustice and pain, there are also reasons for online users to remain cautiously optimistic. The internet and the dark web still have places which one can use for reporting criminal offenses without censorship, thanks to the advantages granted by anonymity. However, those virtual places will always be at risk without an adequate response from the state and its institutions, as scholar Bhaskar Chakravorti states: "Especially in the less-trustworthy regions, [Internet] users will need governments to enact strong digital policies to protect people from fake news and fraudulent scams, as well as regulatory oversight to protect consumers' data privacy and human rights" (Chakravorti, 2018).

References

Abdala, V. (2019). *Las autoras argentinas premiadas en el mundo eligen sus libros del año.* Retrieved from https://www.clarin.com/cultura/autoras-argentinas-premiadas-mundo-eligen-libros-ano_0_GDwIHbUz.html

Arenas, I. (2019). *Mónica Ojeda y la experiencia sensual con el lenguaje.* Retrieved from https://www.elespectador.com/el-magazin-cultural/monica-ojeda-y-la-experiencia-sensual-con-el-lenguaje-article-845986/

Chakravorti, B. (2018). *Trust in digital technology will be the internet's next frontier, for 2018 and beyond.* Retrieved from https://theconversation.com/trust-in-digital-technology-will-be-the-internets-next-frontier-for-2018-and-beyond-87566

Christopherson, K. (2007). The positive and negative implications of anonymity in Internet social interactions: "On the Internet, Nobody Knows You're a Dog". *Computers in Human Behavior,* vol. 23, no. 6, pp. 3038–3056. https://doi.org/10.1016/j.chb.2006.09.001

Ferdous, T. & Al-Hossienie, C. (2019). Trust in Government: Factors Affecting Public Trust and Distrust. *Jahangirnagar Journal of Administrative Studies,* no. 12, pp. 49–63.

Gehl, R. W. (2018). *Weaving the Dark Web.* Cambridge: MIT Press. *https://doi.org/10.7551/mitpress/11266.001.0001*

Greekmythology.com (n.d.). *Kratos.* Retrieved from https://www.greekmythology.com/Other_Gods/Kratos/kratos.html

Ojeda, M. (2016). *Nefando.* Barcelona: Candaya.

Olmos, A. (2016). *Mala fama*. Retrieved from https://blogs.elconfidencial.com/cultura/mala-fama/2016-10-12/nefando-libro-miedo-asco-peli-japonesa_1273495/

Patrizio, A. (n.d.). *Deep web.* Retrieved from https://www.techtarget.com/whatis/definition/deep-Web

Rainie, L., & Anderson, J. (2017). The Fate of Online Trust in the Next Decade. *Pew Research Center.* Retrieved from https://www.pewresearch.org/internet/2017/08/10/the-fate-of-online-trust-in-the-next-decade/

Tait, S. (2008). Pornographies of Violence? Internet Spectatorship on Body Horror. *Critical Studies in Media Communication,* vol. 25, no. 1, pp. 91–111. https://doi.org/10.1080/15295030701851148

Terrell Hanna, K. (n.d.). *Dark web* (darknet). Retrieved from https://www.techtarget.com/whatis/definition/dark-web

The Guide to the Colonial Zone and the Dominican Republic (n.d.). *El Cuco.* Retrieved from https://www.colonialzone-dr.com/el-cuco/

Authors

This book has been conceived and produced by AreaS,
a research group based at Østfold University College in Halden,
Norway, and AreaS' partners.

Further information can be found on the AreaS website:
https://www.hiof.no/oss/forskning/grupper/areas/areas-in-english/

NELSON BAHATI is a Master's student at Oslo Metropolitan University, Norway, with a background in Media and communication studies.
Nelsonbahati50@gmail.com

LAILA BERG is Associate professor at Østfold University College, Norway, with a background in social anthropology and British area studies.
laila.berg@hiof.no

HARALD BORGEBUND is Associate Professor at Østfold University College, Norway, with a background in Political Science.
harald.borgebund@hiof.no

MARIE CHANDELIER is a Postdoctoral Researcher at Côte d'Azur University, Nice, France, with a background in Corpus Linguistics and Discourse Analysis.
mariechandelier16@gmail.com

WLADIMIR CHÁVEZ is Associate Professor at Østfold University College, Norway, with a background in Spanish and Latin American literature and history.
wladimir.a.vaca@hiof.no

KIRSTI K. COLE is the Co-Director of the Campus Writing and Speaking Program and Professor of English at North Carolina State University, with a background in Rhetoric, Composition, and Writing Studies.
kkcole@gmail.com

SAMBA DIALIMPA BADJI is a Research Fellow with DDMAC Project and a Ph.D. Candidate at Oslo Metropolitan University, Norway, with a background in journalism and fact-checking.
sambadia@oslomet.no

CLAIRE DOQUET is Full Professor at the University of Bordeaux, France, with a background in French Linguistics.
doquet.claire@orange.fr

HÅVARD FRIIS NILSEN is a historian and Professor of social science at Østfold University College.
havard.friis.nilsen@hiof.no

ØYVIND GJERSTAD is Associate Professor at the University of Bergen, Norway, with a background in French Linguistics.
Oyvind.Gjerstad@uib.no

KAI A. HEIDEMANN is Associate Professor at Maastricht University, The Netherlands, with a background in Sociology.
kai.heidemann@maastrichtuniversity.nl

SONDRE LINDAHL is Associate Professor at Østfold University College, Norway, with a background in Political Science.
sondre.lindahl@hiof.no

ANJE MÜLLER GJESDAL is Associate Professor at Østfold University College, Norway, with a background in corpus-assisted discourse analysis.
anje.m.gjesdal@hiof.no

FRANCK ORBAN is Associate Professor at Østfold University College, Norway, with a background in Historical Studies.
franck.orban@hiof.no

CÉLINE POUDAT is Associate Professor at the University Côte d'Azur, France, with a background in Discourse Analysis and Corpus Linguistics.
celine.poudat@unice.fr

ELIN STRAND LARSEN is Associate Professor at Østfold University College, Norway, with a background in Media and Communication Studies.
elin.strand.larsen@hiof.no

MARIE STØREN JAREID is a PhD candidate in Sociology at the University of Oslo, Norway, with a background in Political Science.
marie.jareid@hiof.no

JOHANNA M. WAGNER is Professor at Østfold University College, Norway, with a background in English literature, and gender and critical theory.
johanna.m.wagner@hiof.no