

Why Disinformation is Here to Stay. A Socio-technical Analysis of Disinformation as a Hybrid Threat

Van Raemdonck, Nathalie; Meyer, Trisha

Published in:
Addressing Hybrid Threats: European Law and Policies

DOI:
[10.4337/9781802207408.00009](https://doi.org/10.4337/9781802207408.00009)

Publication date:
2024

Document Version:
Accepted author manuscript

[Link to publication](#)

Citation for published version (APA):
Van Raemdonck, N., & Meyer, T. (2024). Why Disinformation is Here to Stay. A Socio-technical Analysis of Disinformation as a Hybrid Threat. In L. Lonardo (Ed.), *Addressing Hybrid Threats: European Law and Policies* (pp. 57-83). Edward Elgar. <https://doi.org/10.4337/9781802207408.00009>

Copyright

No part of this publication may be reproduced or transmitted in any form, without the prior written permission of the author(s) or other rights holders to whom publication rights have been transferred, unless permitted by a license attached to the publication (a Creative Commons license or other), or unless exceptions to copyright law apply.

Take down policy

If you believe that this document infringes your copyright or other rights, please contact openaccess@vub.be, with details of the nature of the infringement. We will investigate the claim and if justified, we will take the appropriate steps.

Why Disinformation is Here to Stay. A Socio-technical Analysis of Disinformation as a Hybrid Threat

Nathalie Van Raemdonck & Trisha Meyer

Forthcoming book chapter in: Addressing Hybrid Threats: European Law and Policies, ed. Luigi Lonardo

1. Introduction	1
2. What is the disinformation threat?	2
2.1 Disinformation as a hybrid threat	2
2.2 Disinformation from a historical perspective	3
2.3 Technology as a force multiplier	4
2.4 Disinformation's incentives, diffusion and cognitive mechanisms	5
Actors (incentives)	6
Behaviour (diffusion mechanisms)	8
Content (cognitive shortcuts)	10
3. Why is disinformation a threat?	11
3.1 Confusion	11
3.2 Erosion of trust in institutions	12
3.3 Foundations of democracy: tolerance of disinformation	12
3.4 Threat of response	13
4. What are current policy responses?	14
4.1 Diplomacy responses	14
4.2 Platform responses	16
4.3 Resilience efforts	17
5. Conclusion	17

1. Introduction

The world has never been as interconnected in the history of mankind. Information and people can move at an increasing speed from any point A to point B. Globalisation and digital networks allow a free flow of information between every individual with internet access, and have brought incredible opportunities but also challenges along. While globalisation resulted in the COVID-19 virus travelling rapidly across the world in 2020, the accessibility of online sharing tools also permitted an unprecedented level of information exchange. In many cases this contributed to international collaboration in countering the coronavirus. For example, due to unprecedented data sharing the entire genetic make-up of the SARS-COV-2 virus that caused the COVID-19 pandemic was mapped within days. For comparison; it took months for the SARS virus to be mapped in 2003.¹ Similarly, the internet allowed many office employees to work from home and reduce the spread of the virus while keeping essential services that do not require physical presence running. International data exchange also hastened vaccine development, whose speed has been the source of much rumours,

¹ Ian Le Guillou, 'Covid-19: How unprecedented data sharing has led to faster-than-ever outbreak research' (2020) Horizon <https://ec.europa.eu/research-and-innovation/en/horizon-magazine/covid-19-how-unprecedented-data-sharing-has-led-faster-ever-outbreak-research> accessed 6 December 2021.

vaccine hesitancy and disinformation. Such a time of uncertainty has been exploited by actors who want to destabilise democracies by pushing or nurturing anti-vaccination sentiments and conspiracy beliefs.² Without necessarily attributing this to disinformation operations, anti-vaccine sentiments and conspiracy beliefs led to protests against coronavirus measures around the world. A recent survey at a protest in Belgium showed that a large majority (89%) of the protesters against COVID-19 measures were not convinced of the safety of COVID-19 vaccines, with 87% expressing a deep distrust in authorities, believing health organisations were not fully honest about the COVID-19 vaccines.³

The spread of disinformation has been much debated and attributed to the omnipresence of social media: Vosoughi, Roy and Aral found that falsehoods would diffuse six times faster and deeper on Twitter than regular content.⁴ However as we will show in this chapter, the hybrid threat of disinformation is about more than technology. It threatens civil discourse and unravels a shared sense of reality. Disinformation not only harms democracy, but is most effective when democracy is in peril and public trust in institutions is low.

We first define ‘disinformation as a hybrid threat’, its historic background, and the role of technology. We highlight the incentives of actors, how disinformation gets diffused and which cognitive mechanisms are at play for disinformation to take root. Next we expound why disinformation is a hybrid threat, how it challenges democracy to its core, and why there is also a threat in response to disinformation. Lastly, we provide an overview of the hybrid solutions to this hybrid threat by looking at current policy responses that use diplomatic means, responses that regulate platforms’ role in the diffusion of disinformation, and responses that aim to strengthen the resilience of citizens to resist disinformation content.

2. What is the disinformation threat?

2.1 Disinformation as a hybrid threat

Unfortunately, examples of disinformation are not difficult to find. In the European Union (EU), policy-makers started to pay attention to disinformation during the Crimean crisis in 2014. Then, during the 2016 US presidential elections, it became evident that disinformation had become part of Russia’s foreign policy arsenal. Further and sadly, the failure of social media platforms to stop the spread of online disinformation contributed to the genocide of the Rohingya muslim minority in Myanmar. As a final example, during the health pandemic and the 2020 US presidential elections, online disinformation on COVID-19, vaccines and elections seemed rampant, fueled not least by a rise in popularity of conspiracy theories.

The motives of the culprits are ugly, usually to gain political power, whether domestic or foreign, through spreading of inaccurate information, as a strategy to sow doubt and destabilize truths. In this chapter, we build on the definitions of disinformation, as presented by Wardle & Derakhshan in their Council of Europe study and by the High Level Expert Group on Disinformation in their European Commission final report. Disinformation is “information that is false and deliberately created to harm

² Ariel Bogle and Albert Zhang, ‘Chinese and Russian influence campaigns risk undermining Covid-19 vaccination programs’ (2021) ASPI <https://www.aspistrategist.org.au/chinese-and-russian-influence-campaigns-risk-undermining-covid-19-vaccination-programs/> accessed 6 December 2021.

³ Ruud Wouters, Michiel De Vydt and Luna Staes, ‘Coronabetogers tegen het licht gehouden’ [2022] *Media Middenveld & Politiek*, Universiteit Antwerpen 24.

⁴ Soroush Vosoughi, Deb Roy, and Sinan Aral. “The Spread of True and False News Online.” *Science* 359, no. 6380 (March 9, 2018): 1146–51. <https://doi.org/10.1126/science.aap9559>.

a person, social group, organization or country”⁵, it has the *function* to mislead.⁶ In the European Commission's definition, disinformation is defined as “false, inaccurate or misleading information that is designed, presented and promoted to intentionally cause public harm or gain profit”⁷. Disinformation differs from misinformation, which only refers to the inadvertent sharing of false information.⁸ Disinformation and misinformation can also be distinguished from propaganda, which still has a factual basis but presents facts in a tendentious way to instill a particular attitude or response.⁹

The intentionality of disinformation is an important characteristic when it comes to its categorization of a hybrid threat. Hybrid threats are an umbrella concept to describe new forms of conflict and warfare that are a combination of actions that undermine democratic state systems. Hybrid threats are identified as unacceptable foreign interference in sovereign states’ internal affairs and space.¹⁰ Disinformation needs very little foreign interference to take root, but as it has the potential to undermine fair election processes and endanger public health, it qualifies as a hybrid threat. While we attempt to mainly focus on disinformation for the purposes of this chapter, we are aware that the information pollution of our collective online ecosystems has devastating impacts, regardless of the motives. Disinformation is unwittingly copied and even produced by well-meaning citizens, which would technically make this misinformation, yet such practices also pose challenges to democracy that do not even require foreign interference.¹¹

The hybrid threat of disinformation forces Western democracies to contemplate an existential question: *how to counter the erosion of core values of Western democratic societies, when these values, such as open and free communication, might be the means of undermining their own existence?* This catch-22 that democratic societies find themselves in means the level playing field for disinformation is operationally uneven. Authoritarian regimes have more control over their national information spheres, giving them a defensive advantage, while liberal democracies offer open information spheres and spaces for public debate, which makes them vulnerable targets.¹²

2.2 Disinformation from a historical perspective

Disinformation has significant historic precedent. Jonathan Swift famously said in his essay on Political Lying in 1710 that “[f]alsehood flies, and the Truth comes limping after it”.¹³ Indeed organised disinformation to advance political goals certainly is not new. For instance, in ancient Roman times around 33 BC, Octavian and Marcus Antonius launched disinformation campaigns on each other to gain the favour of the senate and the population. One of them was the revelation of a will that Antonius

⁵ Claire Wardle and Hossein Derakhshan, ‘Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making’ (Council of Europe 2017).

⁶ Don Fallis, ‘What Is Disinformation?’ (2015) 63 Library Trends 401.

⁷ European Commission High Level Expert Group on Fake News and Online Disinformation, ‘A Multi-Dimensional Approach to Disinformation - Report of the Independent High Level Group on Fake News and Online Disinformation’ (European Commission DG Connect 2018) <<https://digital-strategy.ec.europa.eu/en/library/final-report-high-level-expert-group-fake-news-and-online-disinformation>> accessed 7 December 2021.

⁸ Wardle & Derakhshan (n 5).

⁹ As defined by the Oxford English Dictionary, 2nd ed., 1989

¹⁰ Georgios Giannopoulos and Hannah Smith, ‘The Landscape of Hybrid Threats: A Conceptual Model’ (European Commission Joint Research Centre and European Centre of Excellence for Countering Hybrid Threats (Hybrid CoE) 2021) <<https://data.europa.eu/doi/10.2760/44985>> accessed 7 December 2021.

¹¹ Whitney Phillips and Ryan M Milner, *You Are Here: A Field Guide for Navigating Polarized Speech Conspiracy Theories, and Our Polluted Media Landscape* (The MIT Press 2021).

¹² Louk Faesen and others, ‘Red Lines and Baselines. Towards a European Multistakeholder Approach to Counter Disinformation’ (The Hague Centre for Strategic Studies 2021) <<https://hcss.nl/report/red-lines-baselines/>> accessed 7 December 2021.

¹³ Jonathan Swift, ‘Essay upon the Art of Political Lying’ (1710) 15 *The Examiner*.

purportedly made for Cleopatra, in which he would leave Eastern Roman territories to Cleopatra and be buried in Egypt. The authenticity of the will is to this day still disputed.¹⁴ Disinformation has also been used in conjunction with military operations, so-called “military deception”.

The practice of disinformation for strategic purposes saw an unfolding and a perfecting in the 20th century. The early years of the Soviet Union illustrate this, as the union employed disinformation for its initial survival. Thomas Rid describes an elaborate ruse by the Soviets after the 1920s Bolshevik revolution, in an attempt to suppress a counterrevolution.¹⁵ The authority of the new Soviet regime was fragile and challenged by monarchists exiled in European countries, so they launched ‘operation Trust’ a disinformation campaign to convince the exiled monarchists that a counterrevolution was being prepared internally. The strategy was wildly successful as it undercut the efforts of the exiled counterrevolutionaries who were of the belief that an internal resistance was brewing. It rendered their political and military capabilities to the point of insignificance. At the same time, the Soviets planted disinformation in Western intelligence agencies that the Russian military was more powerful than it really was. The success of ‘Operation Trust’ emboldened the Soviets and triggered the creation of the first dedicated disinformation unit in the world.¹⁶ Such military deception was also applied by Allied forces during World War II. “Operation Bodyguard”, a disinformation campaign intended to deceive the Germans on the planned location of the D-Day invasion, sent out fake radio transmissions and created fraudulent military reports on the landing spot.¹⁷

2.3 Technology as a force multiplier

Information technology has historically been a force multiplier for disinformation as it opened up new opportunities for false information to reach targets. The impact of these technological changes is extensively described in Kavanagh and Rich observation of ‘Truth Decay’, a process of information disorder where the line between fact and opinion was increasingly blurred, and trust in respected sources of information declined during several periods of US history.¹⁸ First the advent of the newspaper in the late eighteen-hundreds, then the introduction of radio which reached an audience wider than any newspaper ever had, later the introduction of television and photojournalism. Changes in the mass communication system contribute to Truth Decay because new information technologies increased the volume of opinion relative to fact-based stories, to the point where the line between them became blurred. This is not to equate opinion with disinformation, nor to ignore the democratizing role that these means of mass communication have played, but to point to shifts and abuse that take place in subtle and manifest ways. Information technology is a tool to be wielded carefully.

Social media platforms’ role in the spread of disinformation has been particularly prominent in the last few years. They have been designed to prioritise virality over veracity, providing algorithmic reinforcement to disinformation¹⁹ and have increased the volume of information and the speed at which

¹⁴ John Robert Johnson, ‘The Authenticity and Validity of Antony’s Will’ (1978) 47 L’Antiquité Classique 494; Frank A Sirianni, ‘Was Antony’s Will Partially Forged?’ (1984) 53 L’Antiquité Classique 236.

¹⁵ Thomas Rid, *Active Measures: The Secret History of Disinformation and Political Warfare* (Farrar, Straus and Giroux 2020).

¹⁶ *ibid.*

¹⁷ Michael Farquhar, *A Treasury of Deception: Liars, Misleaders, Hoodwinkers, and the Extraordinary True Stories of History’s Greatest Hoaxes, Fakes, and Frauds* (Penguin Books 2005).

¹⁸ Jennifer Kavanagh and Michael Rich, ‘Truth Decay: An Initial Exploration of the Diminishing Role of Facts and Analysis in American Public Life’ (RAND Corporation 2018) <<https://doi.org/10.7249/RR2314>> accessed 7 December 2021.

¹⁹ Heidi Tworek, ‘Social Media Platforms and the Upside of Ignorance’ [2019] *Centre for International Governance Innovation* <<https://www.cigionline.org/articles/social-media-platforms-and-upside-ignorance>> accessed 7 March

it can be accessed to an unprecedented level. Indeed, social media have substantially changed the level playing field of disinformation as a hybrid threat. While the era of social media is similar to the previous periods of Truth Decay, Kavanagh and Rich observe a trend that was previously less prominent: an increasing disagreement over objective facts and their interpretation, where people lay claim to their ‘own’ set of facts.²⁰ There is currently no consensus over the causes of such fundamental rupture in factuality. Researchers have tried to explain this parallel reality of facts as ‘echo chambers’.²¹ These are spaces offered or even created by social media and its algorithms where individuals are partitioned off with other like-minded people who conform to a similar pre-existing attitude and bias.²² Garrett however argues that echo chambers and audience fragmentation are not the real problem of disinformation, as there is strong empirical evidence that most individuals encounter a range of political viewpoints and have healthy media diets.²³ Guess found similar results; relatively few people could be seen as ‘trapped in an echo chamber’.²⁴ There is however still reason for concern: Guess also found that the few individuals in echo chambers appear far more politically active in driving traffic to what he terms ‘partisan outlets’ and thereby wield a disproportionate influence and visibility in society, contributing more significantly to information pollution.

Whether ‘echo chambers’ exist and have a big impact or not, it remains important not to fall in the trap of tech determinism. The examples of ‘Truth Decay’ that Kavanagh and Rich find throughout history may coincide with information technology innovation, but they also have a very strong social component. As we explore further on in this book chapter, trust in public institutions and the media has overall declined because of economic and political uncertainties. Society is more than the technology that supports its communication processes. If we only focus on the technology, we lose sight of key actors and vectors at play in the production and promotion of disinformation.

2.4 Disinformation’s incentives, diffusion and cognitive mechanisms

To better understand disinformation in the information ecosystem, we can zoom in on Actors, Behaviour and Content, as Camille François lays it out in her ‘ABC’ framework.²⁵ The framework was intended to reconcile different approaches to deal with disinformation. For example, states and diplomatic efforts primarily seek to attribute disinformation campaigns to actors; whereas platforms have focused mostly on detecting coordinated inauthentic behaviour and hindering its diffusion mechanisms; meanwhile policymakers are preoccupied with the question of what constitutes harmful content and the ways in which we can limit availability thereof and can alert citizens to deceptive content. All three dimensions play a role in disinformation campaigns and need to be expanded on

2022; Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (PublicAffairs 2019).

²⁰ Kavanagh & Rich (n18).

²¹ Seth Flaxman, Sharad Goel and Justin M Rao, ‘Filter Bubbles, Echo Chambers, and Online News Consumption’ (2016) 80 *Public Opinion Quarterly* 298; Fabiana Zollo and Walter Quattrociocchi, ‘Misinformation Spreading on Facebook’ [2017] arXiv <arXiv:1706.09494v2>; Walter Quattrociocchi, Antonio Scala and Cass Sunstein, *Echo Chambers on Facebook* (SSRN Electronic Journal 2019).

²² Stephan Lewandowsky, Ullrich KH Ecker and John Cook, ‘Beyond Misinformation: Understanding and Coping with the “Post-Truth” Era’ (2017) 6 *Journal of Applied Research in Memory and Cognition* 353.

²³ R Kelly Garrett, ‘The “Echo Chamber” Distraction: Disinformation Campaigns Are the Problem, Not Audience Fragmentation’ (2017) 6 *Journal of Applied Research in Memory and Cognition* 370.

²⁴ Andrew M Guess, ‘(Almost) Everything in Moderation: New Evidence on Americans’ Online Media Diets’ (2021) 65 *American Journal of Political Science* 1007.

²⁵ Camille François, ‘Actors, Behaviors, Content: A Disinformation ABC. Highlighting Three Vectors of Viral Deception to Guide Industry & Regulatory Responses’ (Working Paper of the Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression 2019) <https://science.house.gov/imo/media/doc/Francois%20Addendum%20to%20Testimony%20-%20ABC%20Framework%2019_Sept_2019.pdf> accessed 7 December 2021.

before we dive into why this is such a threat. First we consider ‘actors’ by explaining the incentives that drive those creating and spreading disinformation, then we look at ‘behaviour’ and the diffusion strategies to amplify disinformation, and lastly we focus on ‘content’ and the cognitive mechanisms that make disinformation content appealing.

Actors (incentives)

Lauren Hamm of EU DisinfoLab²⁶ identified four types of incentives underlying a disinformation campaign that reflect well the type of actors behind such a campaign:

- Foreign influence, where disinformation is used to disrupt societies to push a certain geopolitical agenda
- Political influence, where disinformation is used to undermine political adversaries and push a domestic agenda
- Lucrative intentions, where disinformation is used to make a financial profit
- Issue-based intentions, where disinformation is used to serve an ideological and/or normative goal.

Most disinformation campaigns can be found at the crossroads of several intentions, drawing on different actors who benefit from the disinformation.

Foreign influence campaigns usually target electoral processes. Some of the most prominent interferences in recent referenda and elections in EU and NATO countries can be tied directly or indirectly to Russia. This was the case for the 2016 Brexit referendum, the 2016 Dutch Ukraine referendum²⁷, the 2017 Catalonia referendum²⁸, the US 2016 presidential election, the French 2017 presidential election and the German 2021 federal election²⁹. The Russian Federation is however not the only state actor employing disinformation for foreign influence. China has primarily focused its information operations on painting the state in a positive light, and casting doubts on events that reflected poorly on China, for example the denial of the existence of Uighur detention camps.³⁰ Since the COVID-19 pandemic, China also promoted conspiracy theories painting other countries such as the US as responsible for the virus, which researchers of the Alliance for Securing Democracy have found to be reinforced by US adversaries Iran and Russia.³¹ Preceding the digital age, the United States’ Central Intelligence Agency (CIA) also regularly used disinformation as a Cold War tactic to stop popular leftist political movements from taking root in many Latin-American countries.³²

²⁶ Lauren Hamm ‘The Few Faces of Disinformation’ (2019) EUDisinfoLab <https://www.disinfo.eu/publications/the-few-faces-of-disinformation> accessed 7 December 2021.

²⁷ Andrew Higgins, ‘Fake News, Fake Ukrainians: How a Group of Russians Tilted a Dutch Vote’ (2017) The New York Times <https://www.nytimes.com/2017/02/16/world/europe/russia-ukraine-fake-news-dutch-vote.html> accessed 6 December 2021.

²⁸ Olga Lautman, ‘Catalonia: Where There’s Trouble There’s Russia’ (2021) Center for European Policy Analysis (CEPA) <https://cepa.org/catalonia-where-theres-trouble-theres-russia/> accessed 6 December 2021.

²⁹ Gabrielle Roncone and others, ‘UNC1151 Assessed with High Confidence to have Links to Belarus, Ghostwriter Campaign Aligned with Belarusian Government Interests’ (2021) Mandiant <https://www.mandiant.com/resources/unc1151-linked-to-belarus-government> accessed 6 December 2021.

³⁰ Clint Watts, ‘Triad of Disinformation: How Russia, Iran, & China Ally in a Messaging War against America’ (2020) Alliance for Securing Democracy <https://securingdemocracy.gmfus.org/triad-of-disinformation-how-russia-iran-china-ally-in-a-messaging-war-against-america/> accessed 6 December 2021.

³¹ *ibid*

³² Grace Livingstone, *America’s Backyard: The United States and Latin America from the Monroe Doctrine to the War on Terror* (Zed Books 2009).

Foreign influence may very well move away from coordinated campaigns, to promoting and amplifying existing campaigns by homegrown populist, far-right and anti-establishment groups.³³ Russian interference has seen years of disinformation campaigns aimed at reducing trust in the European Union, its member states and democratic institutions and undermine liberal-democratic values, as the EU Stratcom division describes it.³⁴ According to EU Stratcom, the tactic has nourished local Eurosceptic and populist voices who are now emboldened to create their own disinformation campaigns. The push away from foreign-born disinformation campaigns was also a finding by the Oxford Internet Institute³⁵, which observed that the disinformation threat is originating more and more from domestic actors around the world. In 2020, they found evidence in sixty-one countries, among which fourteen EU and NATO members³⁶, of politicians and political parties to have used *computational propaganda*. Computational propaganda is defined in this case as information used to manipulate public opinion online, going broader than mere disinformation, as defined in this chapter.

Private companies also provide disinformation as a service to a variety of actors. The Oxford Internet Institute found forty-eight instances of private companies in 2020 deploying computational propaganda on behalf of a political actor. Since 2018 there have been more than sixty-five firms offering computational propaganda as a service. In total, the Global Inventory of Organized Social Media Manipulation found that almost US \$60 million has been spent on hiring these firms since 2009.³⁷ Indeed, disinformation is also created and spread by actors who run a profit from those engaging with it. There is a profit to be made in clickbait disinformation, like the Macedonian fake news operations that became notorious during the 2016 US elections. The websites would run articles full of whatever falsehood got them the most clicks to lure people there and generate revenue from ads.³⁸ Research by the Centre for Countering Digital Hate also uncovered the profit models of anti-vaccine entrepreneurs and conspiracy professionals. Some of the leading figures make six-figure salaries from their followers.³⁹

Finally, online subcultures also spread disinformation for trolling or ideological purposes.⁴⁰ For example the QAnon movement, which became a virulent conspiracy theories network spread across at least seventy-one countries⁴¹, primarily originated on 4chan. The platform is known for its toxic use of irony, and most users admitted to be merely ‘LARPing’⁴² and not taking the conspiracy theories seriously. It then underwent a process of what De Zeeuw et al. call a process of ‘normie-fication’ where

³³ Michael Birnbaum, ‘Europe was worried Russia would mess with its elections. Now it has other fears.’ (2019) The Washington Post https://www.washingtonpost.com/world/europe/europe-was-worried-russia-would-mess-with-its-elections-now-it-has-other-fears/2019/05/20/d0c18552-77f1-11e9-a7bf-c8a43b84ee31_story.html accessed 6 December 2021.

³⁴ EUvsDisinfo, ‘EU Elections Update: Reaping What Was Sown’ (2019) EUvsDisinfo <https://euvsdisinfo.eu/eu-elections-update-reaping-what-was-sown/> accessed 6 December 2021.

³⁵ Samantha Bradshaw, Hannah Baily and Philip N Howard, ‘Industrialized Disinformation 2020 Global Inventory of Organized Social Media Manipulation’ (Oxford Internet Institute 2021).

³⁶ EU and NATO states: Austria, Czech Republic, Germany, Greece, Hungary, Italy, Malta, Netherlands, Poland, Spain, Sweden, Turkey, United Kingdom and United States

³⁷ Bradshaw, Baily and Howard (n 35).

³⁸ CNN.com, ‘The Fake News Machinery: inside a town gearing up for 2020’ (2017) <https://money.cnn.com/interactive/media/the-macedonia-story/> accessed 7 December 2021.

³⁹ Centre for Countering Digital Hate, ‘Pandemic Profiteers. The Business of Anti-Vaxx’ (Centre for Countering Digital Hate 2020) <<https://www.counterhate.com/pandemicprofiteers>> accessed 7 December 2021.

⁴⁰ Alice Marwick and Rebecca Lewis, ‘Media Manipulation and Disinformation Online’ [2017] Data & Society <https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf> accessed 6 December 2021.

⁴¹ Masood Farivar, ‘How the QAnon Conspiracy Theory Went Global’ (2020) VOA News <https://www.voanews.com/a/usa-how-qanon-conspiracy-theory-went-global/6194444.html> accessed 6 December 2021

⁴² LARP stands for Live Action Role-Play, a type of fantasy game that on 4chan is used to indicate the practice of “playing” at doing politics.

the QAnon conspiracy theories gradually spread to networks unfamiliar with its original (sub)cultural context, where users started following and spreading the conspiracy for ideological purposes.⁴³ Some Qanon adherents also took part in the January 6 attack on the US Capitol.⁴⁴

Behaviour (diffusion mechanisms)

Disinformation by its very definition needs to be observed by people in order for it to cause harm and be considered a hybrid threat. Some campaigns were found, such as the Russian-attributed Secondary Infektion, to have had a lot of resources spent on it, but not actually reach a very wide audience.⁴⁵ They are therefore not considered effective disinformation campaigns. For disinformation to reach a broad audience, it needs to find its way into the information ecosystem. Media outlets are a popular and effective avenue in this regard, so are influencers, and to a lesser extent bots and trolls. The coordination of accounts and techniques used to spread false content is commonly referred to as ‘coordinated inauthentic behaviour’ (CIB) even though a standardized definition is still missing on what exactly constitutes CIB.⁴⁶ Most often a combination of diffusion methods and vectors is used.

Polarised, biased or factional media can have a great impact in spreading disinformation. Researchers of the Berkman Klein Center found that the right-wing media ecosystem, primarily Fox News and talk radio, reinforced and spread the narrative that there was widespread voter fraud on the mail-in voting ballots of the US 2020 presidential election, more so than Russian trolls or Facebook clickbait seekers.⁴⁷ The Kremlin-backed news outlet Russia Today is also recognized as an important instrument of Russia’s information warfare, which serves to erode a positive image of Western countries and mix valid criticism with conspiracy theories.⁴⁸ Further, established news brands can become disinformation spreaders by no longer abiding by journalistic integrity standards. This was the case with the French outlet FranceSoir. EU DisinfoLab reported how the number of disinformation, conspiracy theories and anti-vaccination content increased on the formerly reputable outlet by the end of 2019 after it landed under new management and most of its journalists were fired.⁴⁹ The switch had lucrative effects, as the website went from thousands to millions of monthly visitors during the COVID-19 pandemic.⁵⁰ Political actors can also create alternative media outlets to pose as credible sources, or even recruit unwitting journalists from reputable news outlets. These can amplify a piece of disinformation so it will

⁴³ Daniel De Zeeuw and others, ‘Tracing Normification’ [2020] First Monday
<https://journals.uic.edu/ojs/index.php/fm/article/view/10643> accessed 6 December 2021.

⁴⁴ Michael Jensen and Sheehan Kane ‘Profiles of Individual Radicalization in the United States (PIRUS)’ (2021) START Research Brief
https://www.start.umd.edu/sites/default/files/publications/local_attachments/START_PIRUS_QAnon_Mar2021.pdf last accessed 21 February 2022

⁴⁵ Lily Hay Newman, ‘The Russian Disinfo Operation You Never Heard About’ (2020) Wired
<https://www.wired.com/story/russia-secondary-infektion-disinformation/> accessed 6 December 2021.

⁴⁶ Antoine Grégoire, ‘CIB Detection Tree: Third Branch’ (2021) EU DisinfoLab
<https://www.disinfo.eu/publications/cib-detection-tree-third-branch/> accessed 6 December 2021.

⁴⁷ Bruce Etling and others, ‘Mail-In Voter Fraud: Anatomy of a Disinformation Campaign’ (2020) Berkman Center Research Publication No. 2020-6 https://cyber.harvard.edu/publication/2020/Mail-in-Voter-Fraud-Disinformation-2020?mc_cid=f067cbe3ca3 accessed 6 December 2021/

⁴⁸ Mona Elswah and Philip N Howard, ‘“Anything That Causes Chaos”: The Organizational Behavior of Russia Today (RT)’ (2020) 70 Journal of Communication 623; Ilya Yablokov, ‘Conspiracy Theories as a Russian Public Diplomacy Tool: The Case of Russia Today (RT)’ (2015) 35 Politics 301.

⁴⁹ EU DisinfoLab, ‘The role of “media” in producing and spreading disinformation campaigns’ (2021) EU DisinfoLab
<https://www.disinfo.eu/publications/the-role-of-media-in-producing-and-spreading-disinformation-campaigns/> accessed 6 December 2021.

⁵⁰ SC avec AFP, ‘FranceSoir, le complotisme pour moteur?’ (2021) Stratégies.fr
<https://www.strategies.fr/actualites/médias/4053018W/francesoir-le-complotisme-pour-moteur-.html> accessed 7 December 2021.

get picked up by other news outlets. This is called the ‘trading up the chain’ strategy, as first reported by Marwick and Lewis⁵¹.

Media outlets can amplify disinformation on their own, but such outlets can also serve as platforms from which content is amplified by fake accounts on social media. So-called ‘troll factories’ such as the widely covered Russian-linked Internet Research Agency⁵² or the Chinese-attributed ‘fifty-cent army’⁵³ operate by creating fake social media accounts that push disinformation narratives and disinformation sources into the online ecosystem. To have an impact online, most of these accounts try to become ‘digital influencers’. A keen understanding of the culture and the political sentiment of the population whose attention they are trying to gain is needed in order for this to succeed.⁵⁴

Actors sometimes also make use of established influencers who have a public persona and serve as opinion leaders, paying them to spread disinformation. An example of such an attempt emerged during the COVID-19 pandemic, when German and French YouTubers received offers to spread disinformation on the Pfizer-BioNTech vaccine.⁵⁵

The importance of repetition is not to be underestimated either. Social media studies found ‘familiarity’ to be a powerful and persuasive factor.⁵⁶ Pennycook, Cannon & Rand describe how frequent exposure increases the accessibility of disinformation in memory, creating a sort of sleeper effect where the information but not the credibility of the source sticks.⁵⁷ As Paul and Matthews discuss in their 2016 paper on the methods by which Russia effectively creates a ‘firehose of falsehood’, repetition is one of the most effective techniques for getting people to accept disinformation.⁵⁸ To increase this repetition and also make disinformation content visible in the first place, CIB campaigns spam messages in social media comment sections or as replies to trending topics by using bots⁵⁹ or script-based disinformation workers.⁶⁰

In addition, many campaigns count on reaching a broad audience through organic spread of the disinformation. By creating majority illusions and ‘illusions of engagement’⁶¹ with the initial inauthentic accounts, regular users can become enthusiastic and convinced of the disinformation and amplify it themselves. Ong & Cabañes speak of a ‘porous border’ between paid/unpaid or worker/fan

⁵¹ Marwick and Lewis (n 40).

⁵² Andrew Dawson and Martin Innes, ‘How Russia’s Internet Research Agency Built Its Disinformation Campaign’ (2019) 90 *The Political Quarterly* 245; Jessikka Aro, ‘Yle Kioski Traces the Origins of Russian Social Media Propaganda – Never-before- Seen Material from the Troll Factory’ (Kioski, 2015) <<http://kioski.yle.fi/omat/at-the-origins-of-russian-propaganda>> accessed 9 December 2021.

⁵³ Rongbin Han, ‘Manufacturing Consent in Cyberspace: China’s “Fifty-Cent Army”’ (2015) 44 *Journal of Current Chinese Affairs* 105.

⁵⁴ Jonathan Corpus Ong and Jason Vincent A Cabañes, ‘Architects of Networked Disinformation. Behind the Scenes of Troll Accounts and Fake News Production in the Philippines’ (The Newtom Tech4Dec Network 2018) <<http://newtontechfordev.com/wp-content/uploads/2018/02/Architects-of-Networked-Disinformation-Executive-Summary-Final.pdf>> accessed 8 December 2021.

⁵⁵ Charlie Haynes and Flora Carmichael, ‘The YouTubers who blew the whistle on an anti-vax plot’ (2021) BBC <<https://www.bbc.com/news/blogs-trending-57928647>> accessed 6 December 2021

⁵⁶ Gordon Pennycook, Tyrone D Cannon and David G Rand, ‘Prior Exposure Increases Perceived Accuracy of Fake News’ (2018) 147 *1865*.

⁵⁷ *ibid.*

⁵⁸ Christopher Paul and Miriam Matthews, ‘The Russian “Firehose of Falsehood” Propaganda Model: Why It Might Work and Options to Counter It’ (RAND Corporation 2016) <<https://www.rand.org/pubs/perspectives/PE198.html>> accessed 8 December 2021.

⁵⁹ Anna Reynolds, ‘Social Media as a Tool of Hybrid Warfare’ (NATO Strategic Communications Centre of Excellence 2016) <<http://www.stratcomcoe.org/social-media-tool-hybrid-warfare>> accessed 11 October 2021.

⁶⁰ Ong and Cabañes (n 54).

⁶¹ *ibid.*

in the disinformation practice, which also illustrates the difficulty in drawing a border between disinformation and misinformation. Often the spread of political disinformation arises from the complicity of ordinary people who were enticed by the deception.⁶²

Content (cognitive shortcuts)

Academic research on disinformation strongly suggests that belief in false or misleading information is driven more by individual emotional and cognitive responses — amplified by macro social, political and cultural trends — than specific information technologies⁶³. While it is important that a piece of disinformation content actually reaches an individual, the cognitive mechanisms play an essential part in how the content is received and whether it sticks. When deciding what is true, people are often influenced by their pre-existing beliefs, and tend to rely on intuitions instead of resorting to analytical thinking.⁶⁴ Kuo and Marwick point out the importance of context in understanding how content resonates, and warn scholars not to treat disinformation as a mysterious toxin that just infects users.⁶⁵ This warning for oversimplification is also heeded by Simon & Camargo, who pointed out that the metaphor of the ‘infodemic’ that was often used during the pandemic to compare the information pollution with the spread of the virus is misleading.⁶⁶

Unlike a virus or a toxin that has the power to infect everyone equally, some members of the population are more susceptible to disinformation. Research has found that disinformation narratives often successfully build on pre-existing ideologies, frequently involving race and inequality.⁶⁷ In this regard, Nisbet & Kamenchuk noted that disinformation campaigns use ‘identity grievances’.⁶⁸ These are frustrations about political, economic, religious or cultural wrongs and/or low institutional trust. Identity grievance disinformation campaigns capitalize on two psychological mechanisms: motivated reasoning and affective polarisation. Motivated reasoning is the desire to avoid dissonant information or beliefs. This can lead to confirmation bias, where users are more prone to believe information that is more consistent with their own ideology and social identity, or opinions of which they are already convinced.⁶⁹ Such identity grievance campaigns can amplify closely held social identities, and increase negative feelings towards out-groups with which one does not identify itself. This builds on affective

⁶² *ibid* p.28.

⁶³ Erik C Nisbet and Olga Kamenchuk, ‘The Psychology of State-Sponsored Disinformation Campaigns and Implications for Public Diplomacy’ (2019) 14 *The Hague Journal of Diplomacy* 18.

⁶⁴ Ullrich KH Ecker and others, ‘The Psychological Drivers of Misinformation Belief and Its Resistance to Correction’ (2022) 1 *Nature Reviews Psychology* 13.

⁶⁵ Rachel Kuo and Alice Marwick, ‘Critical Disinformation Studies: History, Power, and Politics’ [2021] *Harvard Kennedy School Misinformation Review* <<https://misinforeview.hks.harvard.edu/article/critical-disinformation-studies-history-power-and-politics/>> accessed 26 November 2021.

⁶⁶ Felix M Simon and Chico Q Camargo, ‘Autopsy of a Metaphor: The Origins, Use and Blind Spots of the “Infodemic”’ [2021] *New Media & Society* pp.1-22.

⁶⁷ Deen Freelon and others, ‘Black Trolls Matter: Racial and Ideological Asymmetries in Social Media Disinformation’ [2020] *Social Science Computer Review* 0894439320914853; Mutale Nkonde and others, ‘Disinformation Creep: ADOS and the Strategic Weaponization of Breaking News’ [2021] *Harvard Kennedy School Misinformation Review* <<https://misinforeview.hks.harvard.edu/article/disinformation-creep-adoss-and-the-strategic-weaponization-of-breaking-news/>> accessed 8 December 2021; Jonathan Corpus Ong, ‘The Contagion of Stigmatization: Racism and Discrimination in the “Infodemic” Moment’ (*Mediawell SSRC*, 4 February 2021) <<https://mediawell.ssrc.org/literature-reviews/the-contagion-of-stigmatization-racism-and-discrimination-in-the-infodemic-moment/versions/1-0/>> accessed 8 December 2021.

⁶⁸ Nisbet and Kamenchuk (n 63).

⁶⁹ Amos Tversky and Daniel Kahneman, ‘Judgment under Uncertainty: Heuristics and Biases’ (1974) 185 *Science* 1124.

polarisation, where greater negative affect about another group will increase the likelihood of adopting disinformation about such a group.⁷⁰

3. Why is disinformation a threat?

As the previous section illustrated, we cannot separate disinformation from the incentives, technical drivers and socio-psychological factors that are present in these hyperconnected times. Abuses of power, dysfunctional political systems, inequalities and exclusion are breeding grounds for disinformation. The UN Special Rapporteur on Freedom of Opinion and Expression, Irene Khan, eloquently stated that, “[d]isinformation is not the cause but the consequence of societal crises and the breakdown of public trust in institutions. Strategies to address disinformation are unlikely to succeed without more attention being paid to these underlying factors”.⁷¹ In addition, our current out of kilter skepticism leads to distrust in pretty much anything, from politics and media to science and religion. Disinformation thrives in our post-trust, post-truth society.⁷² In the section below, we explore why disinformation should be considered a hybrid threat.

3.1 Confusion

Disinformation causes confusion and loss of shared social reality, and loss of trust in the information ecosystem.⁷³ This technique is also called ‘informational gaslighting’, where audiences are overwhelmed by the widespread pollution of the information environment to the point where it becomes difficult to discern truth from fiction. In such a situation, audiences are prone to ‘informational learned helplessness’ as Nisbet & Kamenchuck describe it.⁷⁴ They accept the situation as a given, instead of attempting to fix or avoid the aversive situation. Audiences become less deliberative in evaluating messages, and become more open to persuasive instead of factual messages. Hannah Arendt makes the argument in ‘the origins of totalitarianism’ that people for whom the distinction between fact and fiction and the distinction between true and false (i.e., the standards of thought) no longer exist, are ideal subjects of totalitarianism. They will follow a charismatic leader whose narratives are more comforting and provide a sense of structure. According to Roberts, these forms of Friction and Flooding are more effective in censoring speech than Fear and Control.⁷⁵ Such disinformation flooding tactics are a new form of speech control according to Tim Wu.⁷⁶ They target speakers indirectly by undermining them with a flood of fake news, drowning out their speech and harassing those who counter the disinformation with troll armies. This strategy impacts activists, political actors and investigative journalists who base their efforts on verifiable sources, but whose information is accused of being “fake” by their adversaries, or gets mixed up with disinformation. This in turn creates confusion for supporters, who have limited resources to investigate what is and is not disinformation. Such tactics undermine efforts

⁷⁰ Magdalena Wojcieszak and R Kelly Garrett, ‘Social Identity, Selective Exposure, and Affective Polarization: How Priming National Identity Shapes Attitudes Toward Immigrants Via News Selection’ (2018) 44 Human Communication Research 247.

⁷¹ UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, ‘Disinformation and Freedom of Opinion and Expression, A/HRC/47/25’ (2021) <<https://undocs.org/A/HRC/47/25>> accessed 8 December 2021.

⁷² Julian Baggini, *A Short History of Truth: Consolations for a Post-Truth World* (Quercus 2017).

⁷³ Seva Gunitsky, ‘Democracies Can’t Blame Putin for Their Disinformation Problem’ [2020] *Foreign Policy* 7.

⁷⁴ Nisbet and Kamenchuk (n 63).

⁷⁵ Margaret E Roberts, *Censored: Distraction and Diversion Inside China’s Great Firewall* (Princeton University Press 2018).

⁷⁶ Tim Wu, ‘Is the First Amendment Obsolete?’ [2017] Knight First Amendment Institute at Columbia University <<https://knightcolumbia.org/content/tim-wu-first-amendment-obsolete>> accessed 8 December 2021.

to tackle the societal crises and injustices that were partially responsible for the breakdown of trust in institutions in the first place, hindering efforts to address disinformation in a systematic way.

3.2 Erosion of trust in institutions

Disinformation is primarily an existential threat to liberal democracies, which is what hybrid threats have in common. As Thomas Rid points out in his book on Active Measures, disinformation is an attack against “a political system that places its trust in essential custodians of factual authority”.⁷⁷ McKay & Tenove highlight for these purposes the use of ‘corrosive falsehoods’, which is disinformation that undermines sources of higher epistemic quality, and ‘epistemic cynicism’, where citizens start believing that truth claims, even made by experts, follow a political agenda.⁷⁸ Disinformation essentially erodes trust in the institutions that make up the pillars of liberal democracies, which require trust in moments of high uncertainty and political fragility, such as election times. Disinformation breaks down the authority of evidence, allowing that gap to be filled with emotions.

We highlighted the shift in balance between fact and opinion in legacy media above. The erosion of journalistic standards is evident in many democracies, driven not least by a need to maintain audiences in a competitive (online) media environment.⁷⁹ Similarly, political parties experience pressure to appeal to a fragmented electorate, in a context where personal freedoms have been emphasised and social ties consequently eroded.⁸⁰

Disinformation thus thrives in and exacerbates situations of low political trust. It weakens state legitimacy, to the extent that the communication of media and democratic institutions is no longer accepted as trustworthy.⁸¹ We need a nuanced conversation on how we determine what is a fact. This starts with providing explanation and transparency on the processes underlying media reporting, science or political decisions. It also requires acknowledging and engaging with other viewpoints, while insisting that rigour and transparency is applied equally to allow scrutiny of the credibility and veracity of statements. This is easier said than done in an online environment where politically motivated actors can simply ‘flood the zone with shit’.⁸²

3.3 Foundations of democracy: tolerance of disinformation

Democracies tolerate and encourage opposing and even outlandish viewpoints, as pluralism and disagreements are an integral part of open democratic societies. As political philosopher Chantal Mouffe posits with her agonistic perspective, the political and societal status quo needs to be able to be

⁷⁷ Rid (n 15).

⁷⁸ Spencer McKay and Chris Tenove, ‘Disinformation as a Threat to Deliberative Democracy’ (2021) 74 Political Research Quarterly 703.

⁷⁹ Reuters Institute, ‘Digital News Report 2021’ (University of Oxford 2021) <<https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021>> accessed 8 December 2021.

⁸⁰ Stefaan Walgrave and others, ‘How Issue Salience Pushes Voters to the Left or to the Right’ (2020) 2 Politics of the Low Countries; Jonathan Sacks, *Morality: Restoring the Common Good in Divided Times* (Basic Books 2020).

⁸¹ Flemming Splidsboel Hansen, ‘Russian Hybrid Warfare: A Study of Disinformation’ (Danish Institute for International Studies (DIIS) 2017) <http://pure.diis.dk/ws/files/950041/DIIS_RP_2017_6_web.pdf> accessed 28 September 2021.

⁸² This was famously said by Steve Bannon in a 2018 interview with Michael Lewis, with the full quote going “The real opposition is the media. And the way to deal with them is to flood the zone with shit” Michael Lewis, ‘Has Anyone Seen the President?’ (2018) Bloomberg <https://www.bloomberg.com/opinion/articles/2018-02-09/has-anyone-seen-the-president> accessed 12 February 2022

challenged at all points in time, which sometimes requires conflict.⁸³ According to Mouffe, conflict is healthy for democracy, as long as it remains focused on political conflict, and does not devolve into ‘othering’ or affective polarisation. Article 19 of the UN Declaration of Human Rights also reads that: “Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.”⁸⁴ This includes the right to tell falsehoods, disinform and lie, as long as this does not fall within restrictions to freedom of expression. These restrictions to freedom of opinion and expression must be reserved for very exceptional circumstances and be “provided by law; undertaken to respect the right or reputations of others; protect national security or public order, or protect public health or morals; and necessary and proportionate to achieve a legitimate objective”⁸⁵. Whether democracies should allow lies in democratic deliberation is therefore a normative question. According to philosopher Karl Popper, there should be certain limits of tolerance to outlandish viewpoints to maintain a liberal democracy. Popper describes this as a paradox of tolerance, where a society that is tolerant without limits, will eventually be seized or destroyed by the intolerant. Such limits of tolerance also concern factuality. As Hannah Arendt posited it in her book ‘Between Past and Future’, opinions are allowed to run in many different directions, but must respect the factual reality. There is no democratic deliberation possible if empirically undeniable facts are up for discussion.⁸⁶ McKay & Tenove also warned for ‘unjustified inclusions of falsehoods’, where democracies that give space to falsehoods displace and devalue the contributions of legitimate members of the public.⁸⁷

As we stated at the beginning of this chapter, democracies are inherently vulnerable to disinformation and this hybrid threat forces Western democracies to some existential reflections. Disinformation takes advantage of and abuses democratic values such as the free flow of information and the fostering of multiple competing narratives, as it is often not tolerant of other viewpoints.⁸⁸ Malcontent voices that react to real or perceived failings of the democratic project and out-of-touch elites rightly deserve to be heard in the public sphere, and disinformation can be hard to distinguish from political speech and opinion.⁸⁹ The issue with (foreign) state-supported domestic disruption is that it plants and nourishes seeds of subversion that rejects democratic processes and institutions.⁹⁰ When these take root successfully, they thrive independently and are able to destroy democracy from the inside out.

3.4 Threat of response

In the digital domain disinformation operations have become less measured and even more opaque. They can originate organically, need very little amplification, and transform on their own. Although as stated above, the erosion of trust should not be underestimated, the impact of (foreign) disinformation tactics should not be overestimated either. State actor disinformation was unlikely to have swayed 2016 US presidential elections, nor in a similar vein, could the UK Information Commissioner’s Office find decisive impact of Cambridge Analytica on Brexit.⁹¹

⁸³ Chantal Mouffe ‘Deliberative Democracy or Agonistic Pluralism’ (1999) *Social Research*, 66(3), pp. 745-758

⁸⁴ United Nations Declaration of Human Rights 1948, Article 19. See also the similarly worded Article 19 of the International Covenant on Civil and Political Rights

⁸⁵ International Covenant on Civil and Political Rights 1966, Article 19(3).

⁸⁶ Hannah Arendt, ‘Between Past and Future’ (1961) New York: Viking p.136

⁸⁷ McKay & Tenove (n 78).

⁸⁸ Gunitsky (n 73).

⁸⁹ Paul M Barrett, ‘Tackling Domestic Disinformation: What the Social Media Companies Need to Do’ (NYU Stern Centre for Business and Human Rights 2019).

⁹⁰ EUvsDisinfo, ‘EU Elections Update. Reaping What Was Sown’ (2019) 150 <<https://euvsdisinfo.eu/eu-elections-update-reaping-what-was-sown/>> accessed 7 December 2021.

⁹¹ Philip N Howard and others, ‘The IRA, Social Media and Political Polarization in the United States, 2012-2018’ (Oxford Internet Institute 2018); Gabrielle Lim, ‘The Risks of Exaggerating Foreign Influence Operations and

As we showed in employing the actor-behaviour-content (ABC) framework, a purely actor-based focus on disinformation is not desirable. There is even a danger in shifting the blame of information disorder solely on foreign actors. The narrative that innocent democracy is being subverted by outside forces reduces responsibility and only focuses on inoculation efforts and informational distancing.⁹²

When policymakers respond to disinformation with more control, gatekeeping and censorship, it is seen as a welcome byproduct of disinformation operations by the actors who are not proponents of a free and open internet. Indeed as Giles states, “reflexive control can lead the adversary to make a series of decisions that successively discard options that would improve their position, until they are finally faced with a choice between bad and worse”.⁹³ Heavy-handed interventions against disinformation end up threatening liberal societies’ credibility even further.⁹⁴

4. What are current policy responses?

As the previous two sections show, the hybrid threat of disinformation is a multidimensional problem, where potential responses can cause more harm to democracy than it solves the threats to democracy. Hybrid threats therefore call for hybrid solutions. When it comes to protecting public values on social media platforms, this should not just be a concern of platforms. Helberger, Pierson & Poell speak of a cooperative responsibility of platforms, public institutions, and users.⁹⁵ They propose the creation of (legal) frameworks for shared responsibilities, where platforms and users are not just subjects of regulation, but partners in co-creating and executing such frameworks. This type of co-regulation can create mechanisms of oversight and accountability, which have been largely absent in the self-regulatory efforts of platforms so far.⁹⁶ In this section, we take a brief look at what potential policy responses are drawing on the ABC-framework. We look at diplomacy responses, which signal democratic values to foreign actors; platform responses, how the diffusion of disinformation in the public debate can be regulated; and resilience, strengthening the cognitive abilities and self-determination of citizens to resist disinformation content.

4.1 Diplomacy responses

There is a big demand for accountability from actors to not deliberately spread falsehoods into the information ecosystem. On the international level, Faesen et al. suggest focusing on establishing government-to-government norms against disinformation that are based on the principle of non-intervention.⁹⁷ Given the deep divisions in the United Nations on state behaviour in cyberspace, this may indeed be a more viable option than attributing disinformation operations to specific actors.

Disinformation’ (Centre for International Governance Innovation, 2020) <https://www.cigionline.org/articles/risks-exaggerating-foreign-influence-operations-and-disinformation/> accessed 6 December 2021.

⁹² Gunitsky (n 73).

⁹³ Keir Giles, ‘Russian Information Warfare. Construct and Purpose’ in Timothy Clack and Robert Johnson (eds), *The World Information War* (1st edn, Routledge 2016).

⁹⁴ Rid (n 15).

⁹⁵ Natali Helberger, Jo Pierson and Thomas Poell, ‘Governing Online Platforms: From Contested to Cooperative Responsibility’ (2018) 34 *The Information Society* 1.

⁹⁶ Faesen and others (n 12); Trisha Meyer and Alexandre Alaphilippe, ‘Platform (Un)Accountability. Reviewing Platform Responses to the Global Disinfodemic One Year Onward’ in Judit Bayer and others (eds), *Perspectives on Platform Regulation: Concepts and Models of Social Media Governance Across the Globe* (1st edn, Nomos Verlagsgesellschaft mbH & Co KG 2021) <<https://doi.org/10.5771/9783748929789-507>> accessed 12 February 2021.

⁹⁷ Faesen and others (n 12)

This does not however mean the EU and NATO member states are not exploring punitive options against disinformation campaigns. As High Representative/Vice-President Josep Borrell said in the 2020 Joint communication on Tackling COVID-19 disinformation, “[d]isinformation in times of the coronavirus can kill. We have a duty to protect our citizens by making them aware of false information, and expose the actors responsible for engaging in such practices.” Borell promised to create a toolbox for countering foreign influence operations and interference, including new instruments that allow for the imposing of costs on perpetrators.⁹⁸ At the time of writing, in 2021, the EU has not employed any sanctions against disinformation actors, although it has threatened sanctions over the ‘ghostwriter’ cyber activities around the German elections that had a disinformation component.⁹⁹ Ghostwriter has been associated with the Russian state by the EU High Representative, although other researchers found signs that point to Belarus as the responsible actor.¹⁰⁰ The United States has taken the step of imposing sanctions against specific members of the Russian intelligence services and entities involved in the 2016 election interference¹⁰¹, and the 2020 election¹⁰².

Other diplomacy endeavours focus on correcting disinformation in the public domain. The European External Action Service (EEAS) created the East StratCom Task Force in 2015 to address pro-Kremlin disinformation campaigns.¹⁰³ Their flagship project EUvsDisinfo attempts to raise awareness and expose disinformation activities that affect the EU, its member states and countries in the shared neighborhood. It employs public diplomacy by compiling disinformation cases originating in pro-Kremlin media and increasing public understanding of Russian-backed disinformation operations. NATO also attempts to launch counternarratives when disinformation is spread about NATO allied troops and operations. It describes its approach against disinformation as a two-step process of UNDERSTAND and ENGAGE: first to understand the information environment by tracking, evaluating and analysing the information environment surrounding its operations; then to embed those insights in its strategic communication.¹⁰⁴

⁹⁸ EEAS, ‘The Essential Fight against Disinformation and Manipulation’ (27 December 2020)

<https://eeas.europa.eu/headquarters/headquarters-homepage/91038/essential-fight-against-disinformation-and-manipulation_enhttps://eeas.europa.eu/headquarters/headquarters-homepage/91038/essential-fight-against-disinformation-and-manipulation_en> accessed 9 December 2021.

⁹⁹ Council of the EU, ‘Declaration by the High Representative on Behalf of the European Union on Respect for the EU’s Democratic Processes’ (24 September 2021) <<https://www.consilium.europa.eu/en/press/press-releases/2021/09/24/declaration-by-the-high-representative-on-behalf-of-the-european-union-on-respect-for-the-eu-s-democratic-processes/>> accessed 9 December 2021.

¹⁰⁰ Roncone and others (n 29)

¹⁰¹ The first sanctions imposed by the Obama administration in 2016 mostly referred to ‘cyber-enabled activities that caused a misappropriation of information with the purpose or effect of interfering with or undermining election processes or institutions’, US Department of the Treasury, ‘Treasury Sanctions Two Individuals for Malicious Cyber-Enabled Activities’ <<https://home.treasury.gov/news/press-releases/j10693>> accessed 9 December 2021. Then by the Trump administration in 2018 had similar wording, US Department of the Treasury, ‘Treasury Sanctions Russian Cyber Actors for Interference with the 2016 U.S. Elections and Malicious Cyber-Attacks’ <<https://home.treasury.gov/news/press-releases/sm0312>> accessed 9 December 2021.

¹⁰² The 2020 election sanctions explicitly mention the use of disinformation outlets and the spread of misinformation about US political candidates and US election processes and institutions, US Department of the Treasury, ‘Treasury Escalates Sanctions Against the Russian Government’s Attempts to Influence U.S. Elections’ <<https://home.treasury.gov/news/press-releases/jy0126https://home.treasury.gov/news/press-releases/jy0126>> accessed 9 December 2021.

¹⁰³ European Council, ‘European Council Meeting (19 and 20 March 2015) Conclusions EUCO 11/15’ <<https://www.consilium.europa.eu/media/21888/european-council-conclusions-19-20-march-2015-en.pdf>> accessed 9 December 2021.

¹⁰⁴ NATO, ‘NATO’s Approach to Countering Disinformation: A Focus on COVID-19’ (17 July 2020) <<https://www.nato.int/cps/en/natohq/177273.htm#fn1>> accessed 9 December 2021.

4.2 Platform responses

Social media platforms carry a large responsibility when it comes to disinformation. They host disinformation and serve as amplifiers in the spread of disinformation. We need their technological fixes, but at the same time, their efforts are currently largely uncoordinated.

Online platforms tackle disinformation primarily through their community guidelines and editorial policies. They label, deprioritize, block or take down content, using both automated and human moderators. Platforms also support the work of third-party fact checkers and increasingly rely on promoting reliable and trustworthy content. Users and fact-checkers are able to report and flag content but are not always able to appeal decisions made. Importantly, in the platforms' rules, disinformation is defined more broadly than strictly illegal content, also including undesirable and harmful content. In other words, editorial policies tend to be stricter than legally required (at least in the jurisdiction of legal registration of the platforms). Further, platforms also target disinformation at an account level. They have developed robust mechanisms for detecting manipulation and coordinated inauthentic behaviour (digital fingerprinting). This can in some part be attributed to the large amounts of metadata the platforms take from a user, which does raise privacy concerns and the need for safeguards against personal data breaches. Social media platforms block or remove users who repeatedly break platform rules, thereby rooting out coordinated inauthentic behavior.

Platform responsibility has been a central tenet in the EU's efforts against disinformation. The EU recognizes the role that platforms play as modern soapboxes in the public square. This has been evident in its efforts to conclude a Code of Practice on Disinformation to encourage tackling and monitoring of disinformation campaigns online, primarily at platform level, in the EU in 2018. The Code of Practice has undergone evaluation, revision and expansion in 2021, with new signatories from the advertising industry, fact-checking and disinformation research community having joined the ranks.¹⁰⁵ The self-regulatory Code of Practice is however not recognised to result in sufficient platform action.¹⁰⁶ The proposed EU Digital Services Act will serve as a regulatory backstop and address platform accountability, including (most relevant for disinformation) transparency reporting, access to platform data for researchers, mitigation of systemic risks, oversight structures, and notification, review and appeal mechanisms for actions taken.¹⁰⁷ Another complementary set of legislation, the proposed regulation on political advertising and targeting, would tackle transparency in political campaigning, online and offline, in the run up to the 2024 EU elections.¹⁰⁸

¹⁰⁵ EU Code of Practice on Disinformation 2018; European Commission, 'Revision of the Code of Practice: The Strengthened Code Expected by March 2022' (2 December 2021) <<https://digital-strategy.ec.europa.eu/en/news/revision-code-practice-strengthened-code-expected-march-2022>> accessed 9 December 2021.

¹⁰⁶ European Commission, 'Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. European Commission Guidance on Strengthening the Code of Practice on Disinformation' (2021) COM(2021) 262 final <<https://digital-strategy.ec.europa.eu/en/library/guidance-strengthening-code-practice-disinformation>> accessed 9 December 2021; European Regulatory Group for Audiovisual Media Services, 'ERGA Report on Disinformation: Assessment of the Implementation of the Code of Practice' (2021) <<https://erga-online.eu/wp-content/uploads/2020/05/ERGA-2019-report-published-2020-LQ.pdf>> accessed 9 December 2021.

¹⁰⁷ European Commission, Proposal For a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC 2020 [COM/2020/825 final].

¹⁰⁸ European Commission, 'European Democracy: Commission Sets out New Laws on Political Advertising, Electoral Rights and Party Funding' (25 November 2021) <https://ec.europa.eu/commission/presscorner/detail/en/ip_21_6118> accessed 9 December 2021.

4.3 Resilience efforts

The EU also tackles online disinformation through resilience, meaning citizens become more aware of disinformation campaigns, are able to recognize it and are empowered to adequately respond to its spread. This was emphasized in the EU Communication on increasing resilience and bolstering capabilities to address hybrid threats in 2018¹⁰⁹, highlighting the EU's multi-pronged approach towards disinformation. The highly relevant 2020 European Democracy Action Plan¹¹⁰ seeks to strengthen societal and democratic foundations through some already mentioned initiatives on election integrity and platform accountability, but also strengthen civic participation, media freedom and pluralism. In this context, the European Commission also supported the creation of the European Digital Media Observatory (EDMO) as well as national and regional hubs, to foster networks of fact checkers, researchers and media literacy professionals tackling disinformation in Europe.¹¹¹ Related, the Audiovisual and Media Action Plan aims, through financial and organizational support measures, to open the way to a pluralistic and high-quality media and information supply. Similar to the European Democracy Action Plan, it also includes measures to raise media literacy.¹¹² Finally, in addition to the work of the East Stratcom Task Force, the European External Action Service also invests in building resilience against foreign disinformation, information manipulation and interference, through capacity building with governments, media and civil society of neighbouring countries.¹¹³ These measures demonstrate the EU's whole of society approach to disinformation.

5. Conclusion

The threat of disinformation is hybrid. In this chapter, we briefly discussed the actors, diffusion methods and cognitive drivers of disinformation spread. We argued that disinformation is more than a foreign threat, it is an inherent vulnerability of a democratic society. Disinformation abuses the free flow of information and the fostering of multiple competing narratives.

The solution to disinformation is therefore necessarily hybrid. Much attention has gone to foreign actor intervention and platform regulation, but other measures focused on societal resilience are equally important. Importantly, we cannot and should not count on social media platforms' content moderation tools to make the disinformation challenge disappear. Our moral code needs to be rebuilt and each citizen has a role to play.

¹⁰⁹ European Commission and EU High Representative of the Union for Foreign Affairs and Security Policy, 'Joint Communication to the European Parliament, the European Council and the Council. Increasing Resilience and Bolstering Capabilities to Address Hybrid Threats' (2018) JOIN(2018) 16 final <https://eeas.europa.eu/sites/default/files/joint_communication_increasing_resilience_and_bolstering_capabilities_to_address_hybrid_threats.pdf> accessed 9 December 2021.

¹¹⁰ European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on the European Democracy Action Plan 2020 [COM/2020/790 final].

¹¹¹ European Digital Media Observatory, 'United Against Disinformation' <<https://edmo.eu/>> accessed 9 December 2021.

¹¹² European Commission, Communication From the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Europe's Media in the Digital Decade: an Action Plan to Support Recovery and Transformation 2020 [COM/2020/784 final].

¹¹³ EEAS, 'Tackling Disinformation: Information on the Work of the EEAS Strategic Communication Division and Its Task Forces (SG.STRAT.2)' (12 October 2021) <[https://eeas.europa.eu/headquarters/headquarters-homepage_en/105460/Tackling%20disinformation:%20Information%20on%20the%20work%20of%20the%20EEAS%20Strategic%20Communication%20division%20and%20its%20task%20forces%20\(SG.STRAT.2\)](https://eeas.europa.eu/headquarters/headquarters-homepage_en/105460/Tackling%20disinformation:%20Information%20on%20the%20work%20of%20the%20EEAS%20Strategic%20Communication%20division%20and%20its%20task%20forces%20(SG.STRAT.2))> accessed 9 December 2021.

Democracies must pay attention to the underlying factors driving disinformation. To quote Julian Baggini's *Consolations for a Post-Truth World*, "[e]stablishing the truth requires 'epistemic virtues' like embracing modesty, skepticism, openness to other perspectives, a spirit of collective inquiry, a readiness to confront power, a desire to create better truths and a willingness to let our morals be guided by the facts."¹¹⁴ We must seek to be sincere, accurate and other-centred in our interactions and attitudes towards others. In the end disinformation is tackled first and foremost in citizens' attitudes and interactions with others.

¹¹⁴ Julian Baggini, *A Short History of Truth: Consolations for a Post-Truth World* (Quercus 2017).