esade

EsadeEcPol - Center for Economic Policy

FRIEDRICH NAUMANN FOUNDATION For Freedom.

Spain, Italy, Portugal

Combating disinformation in the ageof populism:

liberalism and civil society

FEBRUARY 2021

Abstract

Disinformation has existed for a long time. However, in the age of social media, when a tweet can reach millions of people in seconds, and in the wake of widespread populist and nativist victories, liberal democracies have become an object of concern for analysts. In this report, we present the central dilemmas regarding disinformation, its regulation, and the choices that liberalism must make to confront them. We begin by presenting a descriptive framework of the problem, including a conceptualisation and a theoretical analysis of the relevance of disinformation. We then explore the conundrum of regulation, both at international and national levels, and what steps have been taken until now. In the next section, we analyse liberalism today, in the face of advancements in cognitive psychology and our knowledge of human biases and 'bounded rationality'. Finally, we analyse cases of success in the fight against disinformation and the reasons behind those triumphs. We conclude this report with a series of policy recommendations that we hope can be used both as an effective policy making tool and as a foundation for future work on this topic

Understanding disinformation

Disinformation, old news

'The report of my death was an exaggeration'. Mark Twain took the New York Journal's 1895 announcement of his death in destitution with a sense of humour. The writer confirmed that he was not dead, that he was not even ill, and that he was not poor. Although just an anecdote, this serves as a starting point for tackling disinformation: it has always existed.

Before the concept of disinformation became popular, propaganda was mainly used by governments and institutions, especially in times of war. Information was another weapon that was used to summon national cohesion, encourage the troops, or demoralise the enemy. It was also used in a strategy of distraction and confusion when planning military campaigns.

The British newspapers The Times and the Daily Mail reported the discovery of a 'Kadaver' factory in Germany during World War I, where the Germans allegedly used the bodies of dead Allied soldiers to extract glycerine for soap and margarine. The news was fake and attributed to MI7, a branch of British military intelligence specialising in press and propaganda work.

The German Nazi propaganda machine recycled false beliefs that were used since the 12th century to justify persecution and pogroms against Jews. The most recurrent of these tales has to do with the alleged celebration of rituals in which Jews murdered and then drank the blood of Christian children.

At other times, disinformation has moved away from war toserve business. The New York Sun published in 1835 a series of sensationalist articles about the discovery ofa civilisation of winged humanoids on the moon that has gone down in history as the 'Great Moon Hoax'. The supposed news, which greatly boosted the newspaper's sales, was even accompanied by imaginative and detailed illustrations showing the way of life of this civilisation.

The invention of printing in 1440 made a market for news possible and led to the expansion of disinformation. Increases in printing and distribution multiplied the scope of news and its audience, and this also encouraged the proliferation of biased and false information, at a time when there was no legislation regarding information and responsibility.

This is an important issue: **booms in false news have historically coincided with technological revolutions that have multiplied the distribution of information and its target audience.** This was the case with the advent of the printing press and was followed by the emergence of the mass media when radio, cinema, and television all became formidable advertising and propaganda tools. So much so that in 1964 the communication theorist Marshall MacLuhan coined the famous phrase 'The medium is the message' in his book entitled 'Understanding Media: the Extensions of Man'. He explained that the nature of the medium through which a message is delivered is more important for the impact it will have on the audience than the content of the message itself.

What, who, why, and how

But what is disinformation, fake media, and a fake profile. How is it different from misinformation? Below are some operational definitions for a better understanding of the phenomenon; however, one of the main obstacles in dealing with disinformation, especially from a regulatory point of view,stems from the difficulty in establishing widely shared definitions of the phenomenon at transnational and local levels. What is true and what is false? Who decides? Howdo we draw a clear line between information and opinion? These are questions that are difficult to answer, especiallyin social media, where news, entertainment, and social content are often thrown into the same bag.

WHAT:

- **Disinformation:** false information intentionally spread for political or economic purposes.
- **Misinformation:** wrong information transmitted in an unintentional way. Information that is not necessarily illegal but has a negative social impact. Behind unintentionally transmitted false information there may be someone who has launched it precisely with the aim of confusing others into transmitting it as true.
- **Malinformation:** Information based on reality but used in a biased way to harm a person, social group, organisation, or country.

- **Fake media:** websites that gather disinformative material for viralised distribution through social networks.
- Fake profile: a social network account for a non-existent person, organisation, or company. Sometimes these profiles are used to viralise disinformative content for economic or political purposes.
- **Bot:** a computer program that performs repetitive tasks on digital platforms through a chain of commands. They are often used by disinformation actors to viralise content through fake profiles.

WHO:

Disinformation actors may be nationals or foreigners, they may be state, non-state, or non-state groups serving or sponsored by a state. They may operate nationally, or in a broader context.

WHY:

Disinformation can be based on the strategies of individuals or groups who promote a certain world view, who share an economic interest, or who simply wish to harm or cause instability for countries or organisations.

HOW:

Disinformation actors use digital platforms to massively disseminate content. They often raise armies dedicated to digital guerrilla warfare. Some states employ hundreds of thousands of people to perform these tasks. Disinformation also uses fake profiles run by bots to viralise content. Networks have an enormous capacity for viralisation, and platform algorithms incentivise users with content that matches their biases and so holds their attention. This causes many social network users to unwittingly become promoters of disinformation. When the actors of disinformation succeed in having their campaign viralised by users who are not part of their strategy, then disinformation multiplies its scope and becomes more difficult to combat.

The digital revolution

As we have seen, disinformation has always existed, and moments of technological change tend to accelerate its proliferation. Nearly six centuries after the invention of the printing press and only a few decades beyond the golden age of radioand television, we are living immersed in a new technological revolution that began to take shape at the end of the 1950s and has transformed the information landscape created by analogue media.

Concern about disinformation has increased in the last five years. However, it is worth pausing to analyse this new technological revolution and its social and political implications.

THE DISINFORMATION AGE

Technological revolutions are disruptive and generate conflicts and social changes to which adaptation is necessary. The digital revolution is no different, but it has one characteristic that we must remember: it has led to an **acceleration of technological change** – and this has impacted on the ability of societies to adapt to the pace imposed by technology.

Digitisation at the same time has opened the doors to what has become known as the information age. A few decades ago, information-age theorists discussed a future marked by a democratisation of information and the universalisation of access to knowledge. Philosophers such as Jürgen Habermas predicted the birth of a kind of meritocratic digital agora, in which ideas would compete in open competition until the best arguments were distilled. But it has gradually become clear that the breadth of possibilities opened by the internet has not necessarily benefited knowledge and the structure of contemporary societies.

Digitisation has led to **over-information,** an over-production of content in which it is difficult to filter and discard what is irrelevant or fake, discern valuable information, and know how to direct our attention towards it. We are subject to a constant bombardment of stimuli. Some experts already consider that attention span (focus) is the new intelligence quotient (Newport, 2016). The result is a paradox: in the age of information and knowledge, we are not necessarily better informed or more knowledgeable. The first to warn of this was the Nobel Prize winner and pioneer of artificial intelligence, Herbert A. Simon, in 1971: 'In an information-rich world, the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. What information consumes is rather obvious: the attention of its recipients'. It is unsurprising that digital platforms have become industries that **monetise attention.**

FROM A DEMOCRACY OF AUDIENCES TO DEMOCRACY ON DEMAND

The digitalisation of debate has produced a **breakdown in mediation** that has important political and social consequences. Political scientist Bernard Manin published a book in the 1990s entitled the *Principles of Representative Government* in which he analysed the evolution of democracies. Manin referred to the political moment of that time as a 'democracy of audiences', in which the mass media, especially television, played a central role in the relationship between representatives and the represented. This central position influenced the selection and promotion of political candidates, so much so that Manin spoke of a government of 'media experts'.

The democracy of audiences enabled direct communication between political leaders and the public. The candidates could address their potential voters without intermediation and speak to them through the screen. This encouraged a progressive personalisation of politics, whereby the candidate gained weight to the detriment of the party (which had been the protagonist in the previous stage). Individual personality traits began to be important in electoral programmes.

Since Manin published his work in 1995, several things have changed in audience democracy, to the extent that it might be appropriate to define a **new stage in the evolution of representative government.** The mass media, especially television, continue playing a key role in political strategies. However, **the consolidation of the internet has transformed the nature of the relationship between representatives and the people they represent.** Television made direct communication between candidates and their electorates possible, but this relationship was one-way. The voter was largely a passive subject of communication initiated by the leader. The internet changes this relationship by enabling a **two-way relationship** between representatives and those represented through channels such as social networks. The internet also revolutionises the time needed for communication, which becomes instantaneous. Leaders can communicate with their followers at any time and followers can respond in real time: impact and feedback are immediate.

This technical possibility has added to the trend towards a personalisation of politics that began with television and has made the **breakdown of mediation** more pronounced. This breakdown of mediation extends beyond politics and into all areas of our lives. The technological revolution has extended the public capacity for decision-making without intermediaries and has taken it to places we did not imagine just 20 years ago. Now the cinema is our living room, we can choose our taxi driver with an application, book a holiday without going through an agency, open a bank account without going into a bank branch, buy shoes without being served by a shop assistant, or have dinner from our favourite restaurant without leaving home. And the pandemic lockdowns imposed during 2020 and 2021 have intensified this process.

We can have almost anything right now. It is no coincidence that this technological change has arrived at the same time as a crisis in the representative model of liberal democracies and the rise of **populism**. In contrast to systems of parliamentary representation in which parties mediate between political action and the will of voters, populism claims that the elites have captured the intermediaries and so demands that power is exercised directly by the people. And with populism, **polarisation** increases, as the populist discourse divides society into the morally virtuous and the enemy who must be expelled from the system.

From a representation crisis to an information crisis

If technological transformations are behind the crisis of representation, then the consequences are felt beyond politics.

The breakdown in the traditional information model has had various consequences. Information pluralism has increased: the internet has made it possible to overcome the classic media ecosystem, often dominated at a national level by an elite that resembled an oligopoly. An increase in digital pluralism has enabled the representation of social sectors who previously had no voice in the media.

However, the fragmentation of the media space and the replacement of traditional media (which was subject to legislation regulating the right to information) by much less accountable platforms and social networks has opened the door to undesirable practices. In the digital environment, there are no gatekeepers to monitor the veracity and quality of information, and so the internet has become a space for spreading disinformation and conspiracy theories that lead to social mistrust and undermine democracy.

Moreover, the major digital platforms have become mega-corporations with global reach, and this had led to the emergence of quasi-monopolistic regimes: there is scarcely any competition for Facebook, Twitter, or Google in their various areas. And the traditional press has retreated in the face of an unstoppable advance. Most citizens already get their news through these digital platforms (Shearer, 2018), a trend that has increased during the pandemic. The crisis in journalism has now spread to all Western countries, where newspapers have had to raise pay walls because of the inability of online advertising to compensate for the collapse in print sales.

A study published by the London School of Economics in 2019 described 'the five giant evils of the information crisis' as follows:

- **Confusion:** citizens have difficulty distinguishing what is true and what is fake.
- **Cynicism:** because of the above confusion: the public loses confidence in information sources.
- **Fragmentation:** information available on the internet is potentially infinite, yet this space is becoming increasingly narrow due to a cyber-balkanisation (Van Alstyne and Brynjolfsson, 1997) that divides citizens and walls them into echo chambers with parallel realities and narratives.
- **Irresponsibility:** traditional media are weakened and cannot control the information market, which has moved to digital platforms and websites that are intrinsically not media and are not legally responsible like traditional media. These platforms distribute a huge amount of information, but they accept no responsibility and claim to be like bulletin boards.
- **Apathy.** As a result, citizens become emotionally disengaged from liberal institutions, experience a loss of confidence in democracy, and this favours polarisation (LSE Commission, 2019).

Why we should care: a call to arms

We know that fdisinformation has always existed: however, it becomes more relevant during periods of technological change because these periods tend to be socially and politically disruptive. We have seen that we are immersed in a technological revolution that is changing the media landscape that was shaped in the last century. This is not the first time such change has happened, but this does not mean that we should ignore the risks of disinformation.

From a quantitative point of view, the technological revolution has made it possible for the scope of information (and therefore disinformation) to be more massive than at any other time in history. At the same time, the same technological revolution has caused a progressive process of economic convergence and social integration that we call globalisation. This process has transcended national borders and led to a growing global interdependence – so that a local threat can become a global risk.

From the qualitative point of view, an analysis of what is at stake with disinformation is necessary. We have seen how the technological change in which this phenomenon is embedded has conflicting political and social consequences. Disinformation contributes to the erosion of public trust in information sources, and ultimately, undermines the link between the public and state institutions. It is therefore legitimate to say that **democracy** itself is at stake.

Evidence of this is that the **actors who promote disinformation tend to aim their actions against democratic institutions,** and especially, electoral processes. Their objective is to weaken and destabilise liberal democracies. The campaign for the Brexit referendum in 2016, and the subsequent American election that same year, were marked by disinformation – especially from Russia and China. Russia also acted during the campaign for the illegal independence referendum held in Catalonia in 2017, which sparked a constitutional crisis in Spain. We thus see how disinformation is often sponsored by state actors who represent illiberal models and oppose Western democracy.

European Commission vice-president Věra Jourová directly pointed to Russia and China as active spreaders of disinformation that endangers our democracy and said that we cannot remain indifferent: 'This is not even a wake-up call – this is a call to arms' (Jourová, 2020). A call to arms which for the European Union means the implementation of strategies and regulations to combat disinformation and protect democratic structures.

The regulatory structure

From Russia without love

Interference by state actors with interests in destabilising liberal democracies (especially major electoral processes) during the last decade has awakened Western governments to the need to regulate the flow of information on digital platforms.

The political conflicts of the 21st century are often fought on the internet. We know that China has an army of two million internet users whose activity is centred on disinformation and propaganda. This activity fulfils a double objective, both inward and outward: inwardly, it encourages national cohesion and targets pockets of dissidence that may destabilise the regime; and outwardly, it launches attacks against nations that are ideological or economic rivals.

However, Russia is the state that most uses disinformation. It is estimated that inside Russia half of all online conversations are generated by 'bots', i.e., robots that massively spread disinformation content (Cembrero, 2020).

We have already mentioned the widespread activities of users and bots in disinformation and propaganda work – mainly originating in Russia and China. Their activity has been reported following important electoral processes, such as the referendum that produced Brexit or the 2016 election of Donald Trump. The health crisis produced by the pandemic has also been accompanied by disinformation and conspiracy theories about the origin of the virus or a cure. This has been accelerated by the increased use of digital platforms as a means of information during lockdowns.

Covid-19 has brought about an **'infodemic'**, which the World Health Organization defines as 'an excessive amount of information about a problem, which makes it difficult to identify a solution. They can spread misinformation, disinformation, and rumours during a health emergency. Infodemics can hamper an effective public health response and create confusion and distrust among people'.

Some of this false information seeks to add to the climate of social and institutional distrust we have mentioned. Misinformation about a health threat as great as the one we now face is a major worry, because it can damage people's health. Especially when false information is spread about measures to prevent or cure disease. We have seen the president of a nation like the United States recommending the consumption of bleach to overcome covid-19.

On other occasions, disinformation has fanned xenophobia against the citizens of China, where the epidemic was first declared. And China has also launched propaganda and disinformation campaigns to improve its own image and tarnish that of other countries.

These state and non-state actors have exploited the holes in the system to discredit liberal democracies – and combating them is a major challenge. However, this is a challenge that cannot be tackled without first addressing the regulation of the digital public space and the responsibility of internet platforms.

Are digital platforms just bulletin boards?

This brings us back to the information crisis we discussed above. The transition of the information model towards digital platforms has several drawbacks. These platforms were presented as mere bulletin boards where anyone could share information. As with **bulletin boards,** it was understood that the only person responsible for what was published was the person who posted the content. The platform simply provided a medium.

However, digital platforms are not simple bulletin boards. They work with **algorithms** designed to maximise user attention in a segmented and targeted manner, prioritising content that arouses the most interest to increase the amount of time that the user is on the platform. As a consequence of this segmentation, reality is perceived differently by members of the same community and this creates **social fragmentation.**

Moreover, use of these hyper-focus capabilities by dishonest actors with the aim of harming individuals or organisations represents a threat to the whole system. Finally, we know that bad news, news that appeals to indignation and the emotions, and news that contains conspiracy elements, is the news most likely to be viralised. This results in the emergence of perverse economic rewards for platforms that contribute to the spread of disinformation.

Therefore, the intervention of platforms in the content that is published makes them active rather than neutral mediators between the content and network users: and this forces us to discard the claim that they are mere bulletin boards. Platforms edit, moderate, prioritise, set the conditions of publication, and control the dissemination of content. The **business model** of the digital platforms cannot be ignored when we tackle disinformation. This opens the door to demanding greater responsibility from the platforms for the content on their networks.

Europe vs disinformation

The question of the responsibility of digital platforms for the content on their networks led the European Commission to form an expert group to examine misinformation in 2017. The group's conclusions were set out in the 2018 Communication **'Tackling Online Disinformation: a European Approach'** (European Commission, 2018). These recommendations were then included in the **Code of Practice on Disinformation** (European Commission, 2018). The agreement adopting the code was a milestone in that it brought the industry, for the first time, to voluntarily accept self-regulatory standards to combat disinformation.

At the end of that same year, the Commission and the European External Action Service launched an **Action Plan Against Disinformation** (European Commission, 2019) with a strategy focussed on the 2019 European elections and which aimed to strengthen cooperation between member states and European institutions in the area of disinformation. This effort was joined in March 2019 by the **Rapid Alert System:** a new mechanism which has since facilitated collaboration with other international partners such as the G7 and NATO, as well as with digital platforms, to identify and prevent the spread of disinformation. Finally, in March 2020, the Commission launched a website countering disinformation stories about covid-19.

The agreement between the platforms and the Commission reveals the extent to which internet companies have become aware in recent years of their impact and social relevance. The platforms have since introduced self-regulatory measures that have contributed to a safer digital space and have regularly sent the Commission reports detailing actions taken to combat misinformation. The actions are specifically designed to ensure the neutrality of political campaigns, and focus on the scrutiny of advertising, the removal of false information, and on limiting the visibility of sites considered to promote disinformation. This collaboration has produced results and, unlike what happened three or four years ago, the most recent major electoral processes, from the European Parliament elections in 2019 to the recent US elections, took place without major intrusions from state actors with destabilising interests.

Companies such as Facebook, Google, Twitter, Mozilla, Microsoft, and, more recently, TikTok, have developed protocols to monitor content on their networks and remove fake content, in collaboration with fact-checker agencies that claim to be independent, as well as with the collaboration of academics and experts in the field.

The Facebook strategy, for example, is based on three pillars:

* **Remove** information that is illegal or may harm people.

* **Reduce** the visibility of information which, although it does not violate any law or cause direct physical harm, has a significant impact on democracy and people. At the same time, the user is provided with verified information about the content in question that explains why a posting is probably fake.

* **Inform.** Ensure that users have sufficient content and context to form an informed opinion and provide users with relevant information from trustworthy sources.

However, as action is taken to counter misinformation, the actors have learnt to circumvent controls (for example, by directing their content towards private conversation channels such as WhatsApp) as well as making their messages more sophisticated. This is what happens with 'deepfake' where algorithms are used to generate false audio-visual content whose sophistication gives the appearance of reality.

CORONAVIRUS & INFODEMICS

The complexity of the threat is compounded by the current health crisis. Disinformation in the context of the coronavirus poses new risks to people's health and offers new opportunities for those aiming to undermine democracy. In recent months, misleading information has circulated that is not necessarily illegal, but poses a threat to the well-being of any individual. Some examples can be found in messages that contradict official recommendations to wear a mask or wash hands to prevent infection, or which claim that covid-19 only affects the elderly. Other messages support hate speech, relating the origin of the disease to certain ethnic or religious groups, such as the Chinese or Jews. We can also find fraudulent advertising for alleged cures.

Therefore, the pandemic has prompted the Commission to change its approach to disinformation: this is no longer seen as a debate between the legislator and the platforms, but as a challenge to society as a whole and in which civil society must also be involved. Moreover, member states must protect user rights.

Shadows in the fight against disinformation

However, this process of containing disinformation has not been without controversy. Combating disinformation on the internet is a major challenge because it involves fundamental rights such as freedom of expression or the right to information. Moreover, these rights are regulated as part of national legal systems – while digital platforms operate at a global level.

Furthermore, authorising platforms to decide what is true and what is false can be interpreted as imposing a privatised censorship that limits freedom of expression. The platforms would then control a quasi-monopolistic and self-regulated industry – becoming both judge and interested party.

Finally, the massive submission of political messages by leaders and parties to scrutiny by the platforms has led to an inevitable clash. We saw this during the recent American elections, when Facebook, Google, and Twitter warned their users that messages by Donald Trump may not be aligned with reality. The question raises doubts about where the frontier of legitimate political discourse should be drawn, and also about the platforms' ability to police such a frontier: **can the messages of all the world's politicians be properly monitored**?

There are many questions to which we have no clear answer: who should decide what is true or false? What content should be removed from a platform? What do we do with content that may violate organisations or coexistence – but is not illegal? Who checks the fact-checkers? How do we guarantee a balance between freedom of expression and user safety?

Given the complexity of these issues and the emergence of the pandemic, the European Commission has encouraged member states to become more involved in the regulatory fight against misinformation. However, there is plenty of controversy at national regulatory level. In most countries whose governments have passed legislation to mitigate the effects of disinformation, there have been protests from both society and the media. Some believe that fighting foreign-sponsored disinformation may be an excuse for governments to pass legislation that limits freedom of expression and political criticism.

National regulations

Below is a brief overview of some of the regulatory initiatives undertaken by European nations:

- **Germany.** The Bundestag passed the Network Enforcement Act (NetzDG) *(Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken),* also known as the Facebook Act (Facebook-Gesetz), which allows for millions of dollars in penalties for internet platforms and social networks with over two million users if they do not remove content reported as false or hate speech. In practice, the Facebook Act has revealed problems of implementation by platforms and the government has announced a review of the legislation, which has been widely criticised for imposing censorship, while violating freedom of expression and the separation of powers.

- **France.** Emmanuel Macron's government passed legislation in 2008 against the manipulation of information *(Loi contre la manipulation de l'information, or 'loi fake news')*, with the entire opposition voting against it. The act was also rejected by journalist unions and public opinion as a violation of press freedom and fundamental rights. The act establishes an independent authority, the Higher Audiovisual Council, to guard against interference by media controlled or influenced by a foreign state, especially if it carries disinformation that could alter the neutrality of an electoral process.

  The act authorises summary judicial interventions during election campaigns to interrupt the dissemination of obviously false information that has been massively and artificially distributed and could disturb the public peace or an electoral process. In addition, it requires platforms to be transparent regarding their algorithms, eliminate profiles that spread false information, and provide information on advertisers.

- **Britain.** These is no national regulation on disinformation, but Britain has established a working group on the issue, following foreign interference during the Brexit referendum. The National Security Communication Team (NSCT) is composed of senior officials and civil servants from the secret services to the Ministry of Culture and focuses on identifying false or misleading information relating to the British government and developing truthful content to counter such information. In addition, in 2019 the government announced the incorporation of a subject in the school curriculum that helps children identify disinformation.

- **Spain.** Although Spain does not have a specific law against disinformation, the Spanish government has just approved a procedure to monitor social networks and counteract disinformation campaigns. The lack of consensus with which the executive acted, failing to consult opposition parties or the media, has raised criticism and prompted accusations of censorship. The government claims that the protocol is designed to combat disinformation by foreign states. However, the document's lack of specificity and its intention to 'examine the freedom and pluralism of the media' have led to legal uncertainty and doubts about its real objectives.

- **Italy.** A committee of inquiry in the Italian parliament aims to ensure that the mass dissemination of false information is counteracted. In addition, it has been agreed that digital skills will be taught in schools to train pupils to identify false information and develop critical skills.

A liberal perspective

To argue in favour of a liberal approach to tackle disinformation, we first need to define what we consider as liberalism. Today, although the concept is still debated, it is widely accepted that John Rawls is the philosopher who has produced the concept and theory upon which most of today's liberalism is constructed and debated. This American philosopher wrote his 'Theory of Justice' in 1971, and it became one of the greatest works of political theory of the last century. It is often said that most of what has been written in this discipline after 1971 has either been in favour of Rawls or against, and his works have influenced the ideas of many philosophers and politicians ever since.

His theory is based on the assumption of a renovation of the idea of the social contract. He argues that if we were placed in an 'original position' and 'under a veil of ignorance' without knowing our future ideology, preferences, or social position, then there would be two principles of justice: first, 'a fully adequate scheme of equal basic liberties' which must be compatible with the same scheme for all other citizens. Second, an equality of opportunity in access to all positions with inequalities only being acceptable if they benefit the least advantaged members of society. The first of these principles takes priority over the second. These two principles have been the linchpin of liberal thought in recent decades.

Rawls later reformulated his theory (1995) to ground it in more 'realistic' bases, assuming that the idea of an original position and a 'detached' and rationalistic perspective on the two principles would not be accepted by everyone. Instead, he argued that from very different 'reasonable comprehensive doctrines' (meaning different ideologies which would be acceptable under a democratic outlook) we would end up reaching an 'overlapping consensus' where we would approve the two principles as the basis of our political system.

How can we design a regulation against disinformation with a liberal perspective? Taking a Rawlsian approach is useful. One of the main debates in the current regulation of fake news is where to establish the limits of intervention. Should we ban accounts that post disinformative material? What qualifies as such? Should we avoid its publication or simply the spread of misinformation? In the previous section we examined some of these dilemmas – but for a liberal there should be anever-present risk in any regulation: the violation of the first principle of Rawls.

For instance, let us suppose that we aim to ban illegal content and that which may be deemed as *hurtful* for some citizens. It is logical and compatible with the first principle of Rawls, to preclude illegal publicationsthat incite hate speech and could be dangerous for groups such as the LGTBI or racial minorities. By opposing such speech, we would be effectively protecting the scheme of equal basic liberties for all citizens.

However, when it comes to content regarded as hurtful, we could end up banning, for instance, material that is merely a parody of political or religious leaders. In certain contexts, this type of regulation could be used to pursue people posting in support of the same minorities that were protected by the banning of illegal content mentioned before. A liberal perspective must consider that defending the equal rights of everyone is the priority – and that the limits of freedom of expression include content that can be considered offensive by others. By engaging in an overly restrictive regulation we could damage basic rights.

Thus, a liberal perspective must be focused on what is *illegal* and not what is *offensive.* Similarly, it can work better by addressing the spread of disinformation and limiting the possibilities for sharing of messages that may include fake content. Of course, this entails a process of verification and fact-checking that ensures deeming certain content as fake is trustworthy. We will further discuss proposals in that direction in the final section of this report.

However, if we want to pose feasible goals, then we must also consider another key aspect: how our mind works. In the following section, we will address some of the findings of cognitive psychology regarding the foundations of our beliefs and ideologies, and how liberalism must take these into account to be effective in combating disinformation.

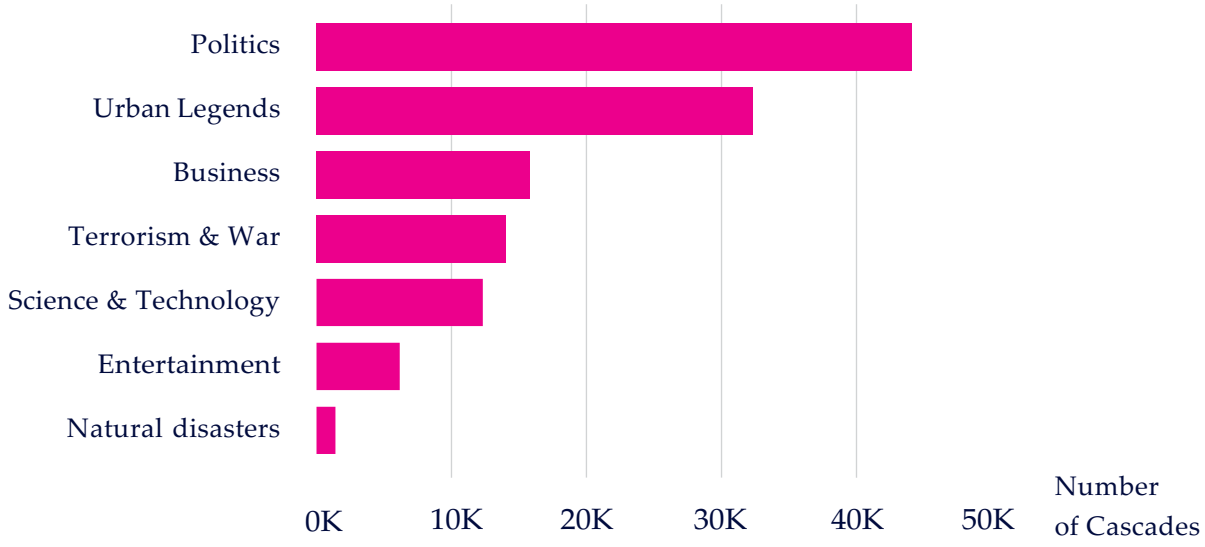
How our minds work: towards a 'liberalism for humans'

One of the most common assumptions in political behaviour models (such as Downs, 1957) is that the voter is rational. He or she has a defined set of preferences and will vote for the party that most closely resembles their 'ideological package'. This understanding of citizens as rational also underpins the basic tenets of liberalism and the Enlightenment as a whole. If this were not the case, and we were effectively more emotional or bound to **limited rationality,** some of the key proposals of liberals would collapse.

And that seems to be the case, according to recent studies. Bryan Caplan published *The Myth of the Rational Voter* (Caplan, 2011) in which he argued that people constantly elect bad politicians because those officials share their biases. This is especially prominent in corruption scandals. For instance, it has been shown that in Spain (Anduiza, Gallego and Muñoz, 2013) tolerance towards corruption is correlated with partisanship, with voters judging differently the same offense depending on the party to which the politician belongs. This phenomenon is not exclusive to Spain and is shared by developing and consolidated democracies (Vivyan, Wagner and Tarlov, 2012; Solaz, de Vries and de Geus, 2019).

However, this should not come as a surprise for social scientists or policymakers familiar with the theory of bounded rationality, recently popularised by the work of authors such as Daniel Kahneman and Amos Tversky (Kahneman, 2011; Lewis, 2016. See also Sunstein and Thaler, 2009). A core argument in these approaches is that our pre-conceptions and intuitions (or 'fast thinking') usually take primacy over our more rational impulses ('slow thinking'). Lewis presents the example of Darryl Morey, general manager in the Houston Rockets (an NBA basketball team) who, despite having carefully constructed a database and analysed vast amounts of data on players to select, did not select Marc Gasol for his team because he 'looked chubby'. Our fast thinking and biases usually betray our judgements, no matter how well informed.

This can be applied to political behaviour. We do not like cognitive dissonance and the popping of our own 'bubbles' (Festinger, 1957). Consequently, we tend to surround ourselves with people whose ideas are like ours, so we do not have to face the cost of either debating or changing our minds. Although this is a human feature, the apparition of social networks such as Twitter has allowed for even bigger echo chambers, where we can surround ourselves with people who share the exact same ideas, and so precluding any possibility of dissent within our circle (Barberá, 2015). Social media, as argued before, has also contributed to a faster spread of disinformation that reinforces our beliefs, especially in political matters. This was shown by Vosoughi, Roy, and Aral (2018) when they analysed 126,000 false stories shared by three million accounts and classified them in 'cascades' or 'spreads' with the following results.



Source: Vosoughi, Roy and Aral (2018)

It would be easy to argue that this is only the case for citizens who are uninformed or uninterested in politics. However, multiple studies confirm that well-informed and well-educated people are even more prone to believe in news that confirms their biases (for a great compilation and a compelling read, see Klein, 2020). Greater sophistication leads to a greater ability to cherry-pick the source of news and so reinforce the bubble effect. In fact, a recent work by Li and Wagner (2020) shows that **uninformed people are more likely to revise their beliefs when confronted with new evidence by, for instance, fact-checkers. However, people who are merely misinformed are less likely to revise their beliefs.**

In another strand of thought, Jonathan Haidt (2012) and his research group established his moral foundation theory. This perspective poses the existence of six foundations for moral behaviour: care/harm – fairness/cheating – loyalty/betrayal – authority/subversion – sanctity/degradation – and liberty/oppression. The problem, they argue, is that different political groups may tap into different foundations to build their ideology. This is also a clash between more communal and more universal values when it comes to shaping our political outlook (Enke, 2019). Although their studies are mainly focused on the US, it can be seen how this split on the moral sources of our beliefs represents a problem for bridging the distance between different ideologies. If the foundation of my ideas is entirely different from my neighbour, how can we ever find a point of agreement without betraying moral intuitions rather than rational political principles?

There has been some criticism of the moral foundation theory that proposes instead a 'morals-as-cooperation' mechanism (Scott Curry, 2019), where we 'have evolved to face problems as social groups, we have biological mechanisms to provide motivation for cooperative behaviour and criteria to evaluate the behaviour of others'. The morals-as-cooperation theory proposes seven foundations, including respect for prior possession and heroism, but in the end it refers to the same problem as traditional views of liberalism: we are not be perfectly rational, and we derive our political views not from a process of perfect reflection, as dreamt by Rawls (2012), but from moral intuitions within our brain.

These processes, in turn, can lead to greater affective polarisation. According to Lyengar et al. (2019) affective polarisation is the animosity towards members of other parties, including perceptions of the 'other' as hypocritical, selfish, and close-minded. Given our preference for news that reinforces our worldviews, our will to avoid cognitive dissonance and the possibly differing moral foundations of our ideologies, bridging the gaps between different groups can be increasingly difficult. Without a certain degree of civic compromise, the governance of a polity is impossible (Gutmann and Thompson, 2012)

Consequently, disinformation is problematic because it contributes to reinforcing these processes and widening the divide between citizens with competing 'reasonable comprehensive doctrines' as we have mentioned before. Not only that but given the way our brains work we are prone to believe them, regardless of our level of political sophistication or education. Consequently, **liberalism has to take into account what we know about our brains and emotions and become a 'liberalism for real people'.**

First, we should be wary of believing that other people are less informed than us, or that they are the only ones who believe in lies and fall prey to conspiracies. As Gray (2019) points out, liberals are also prone to believe in these kinds of theories. Additionally, the research of Philip Tetlock (2017) has shown that expert predictions are no better than those of average citizens in certain contexts (namely, experts adhering to a single theory that explains the world, which connects with the way our brains work).

Secondly, liberals should appeal to emotions and not just rationality. If we are bounded rationally by the limitations of our brain, then focusing exclusively on appealing to that rationality is not going to work. Crafting narratives that appeal to a wider range of people and trying to reach the vast majority of the public, instead of limiting our reach to niches (Klein, 2020), can contribute to reducing the credibility of fdisinformation material.

Finally, we must remember that demand and supply interact when determining political behaviour. A growing demand for partisan information and personalised content is not only the responsibility of the 'citizen-consumer', but also of political leaders and parties who are interested in keeping their base active in a time when party identification is fading and traditional cleavages are defrosting. Non-partisan campaigns against fdisinformation and the rejection of conspiracy theories that are beneficial to liberal interests will weaken demand for the reinforcement of bubbles.

These are merely some approximations to the problem that liberals face when combatting this phenomenon. In the next section, we discuss successful cases to understand the reasons behind liberal victories and what we could learn to deploy these techniques in other contexts.

Cases of succesful fight against disinformation

When it comes to analysing different cases of fights against disinformation, we must distinguish between campaigns organised by institutions, such as national governments, and those created by civil society organisations. We should also distinguish between different territorial levels, from local actions to campaigns that transcend national boundaries. In this section, we will first focus on a campaign conducted by civil society at the country-level (Operation Libero in Switzerland), an effort by the European External Action Service (EuVsDisinfo), and an action that combined efforts by the Taiwanese government and civil society to combat disinformation.

Let's start with Operation Libero. This is one of the best-known cases of fighting against disinformation and, at the same time, defeating a populist party by crafting new narratives (Henley, 2019). Founded in Switzerland in 2014 by a group of young students led by Flavia Kleiner, its goal was to be 'the last line of defence but, at the same time, an operation to launch a counter-attack'. Thanks to its efforts, Libero managed to force the SVP (a Swiss populist party) to lose the referendum it had called on the expulsion of 'foreign criminals'. Libero also ran campaigns to protect same-sex marriage and continues to uphold liberal values.

Why was Operation Libero successful? Kleiner argued (De Gruyter, 2017) that it was due to a combination of offline and online actions. Libero features a section on its website where it debunks the disinformation spread by the SVP and other rivals. For instance, throughout the campaign against the expulsion of foreign criminals, the group uploaded a 'bullshit bingo' that mocked the lies of its rivals. Operation Libero is also willing to combat those internet trolls in social media who could be potential amplifiers of disinformation by deploying trolls of its own to counter their arguments. Kleinerargues that Operation Libero members mock SVP arguments instead of insulting populist voters; take an active rather than merely reactive position; and speak to voters offline to make their arguments more widely appealing. We should not forget that despite the importance of social media, television news is still the main source of news and political information (Alcott and Gentzkow, 2017; Beam et al. 2017), and that Twitter, for instance, is still mostly used by young people. A message that encompasses all society must go beyond the digital gap and be present off-line. This, at the same time, underpins the importance of having committed activists and volunteers willing to work in both realms.

Our second case is EuVsDisinfo, a project by the EastStratCom Task Force, part of the European External Action Service (EuVsDisinfo, 2015) was launched in 2015 to combat the disinformation spread by the Kremlin. This project, in its own words, aims to 'increase public awareness and understanding of the Kremlin's disinformation operations, and to help citizens in Europe and beyond to develop resistance to digital information and media manipulation'. In an effort to turn the tide against disinformation spread by Russia, the project includes a database with samples of the Kremlin's disinformation and 'data analysis and media monitoring services in 15 languages.'

EuVsDisinfo is led by political elites, and thus it has a more 'defensive' function than Operation Libero. It is also more analytical and provides tools that can be used by governments (a database and advisory service) to construct their own defences, and has a strong social media presence with some 52,000 followers on Twitter. However, despite the importance of this cooperation with governments, we must also consider several problems: first and foremost, the reliance on fact-checkers who may not be accurate and instead serve as a source of counterpropaganda (Giorio, 2018). Another weakness is the lack of a presence outside of social media for the general public. While Libero is well-known by the Swiss population, the EuVsDisinfo project is only known by the political elites and those well-informed citizens who are especially interested in this topic. Finally, this 'exclusivity' contributes to reinforcing the beliefs of those who already knew about the project and so reduces its efficacy.

Another successful case is that of the government of Taiwan – as recently praised in Foreign Policy magazine (Kerr and Phillips, 2020). The island has been under constant cyber-attack by China and this includes the spread of misinformation. Taiwan's strategy, according to Aaron Huang (Wang and Huang, 2020) has been based on three pillars: first, constant control of media platforms and the use of memes (short, funny, and understandable) to counter fake narratives, while cooperating with social media to spread clarifications in response to disinformation. Secondly, by treating misinformation as a virus and making sure that the whole community is involved (including 'literacy trucks' that visit rural areas). The third pillar is legislation to avoid foreign interference in elections. Although we can learn from this, we also must be careful, as the constant monitoring of media could become a tool for illiberal purposes. We need to develop strategies that learn from these good practices – but do not copy them without considering context.

The success of these strategies is founded upon a 'whole-society' approach (Steger, 2018). Community leaders, activists, priests, students, and political elites are aligned with the goal of combating the spread of disinformation. Cooperation between civil society organisations and the government is key, including the possibility of citizens sending news to be reviewed by an independent fact-checker. This approach combines the resource advantages of a top-down project with the vitality and local and offline implementation of bottom-up strategies.

Although there are many European initiatives fighting disinformation (for a compilation of other cases, see Lessenski and Kavrakova, 2019), we have presented three that serve as a model for civic society initiatives, government initiatives, and a combination of top-down and bottom-up approaches. Drawing from these examples we can agree on some independent variables that are key to success:

- **Grassroots:** the availability of volunteers or workers who are willing to help in the fight is perhaps the most important variable in all the successful cases. A convincing message must reach the whole community (either at local or European level), and an activated and energised group of activists is needed for this task. Without implementation in key demographic areas, or alliances with civic leaders, it is much more difficult to spread the anti-disinformation message.

- **Online presence:** as we pointed out; television news is still the main source of information for citizens. However, social media has become a key tool for governments and citizens to send messages. It is also an environment where bubble-effects and disinformation can thrive given the opportunity to exclusively follow people from your own in-group. Fighting in the online realm is crucial for any successful project.

- **Off-line presence:** as said before, there is still a digital breach in terms of use of technology and sources of information. For campaigns that address the whole community, it is necessary to, as Flavia Kleiner said, 'shake hands and get dirty' by engaging in debates on TV, radio, and even civic centres. Online disinformation is easier to track, but those who are outside of the internet can only be tracked by a community effort.

- **Non-partisan approach:** it can be problematic in terms of liberal rights and efficacy in combating disinformationif the government is perceived as silencing certain political options. An approach that relies on independent fact-checkers, checks and balances in the action of government, transparency, and civil society seems to be most successful as it dissipates doubts about a partisanship bias in the narrative provided by the government.

More research is needed. Research that reports on successful cases at different levels (especially locally) and makes comparisons with larger samples would be useful in testing these four hypotheses (in which independent variables are key for the success of such initiatives). We hope that this section can be used as a foundation for these future studies and contribute to creating a framework for research on organisations that oppose disinformation.

Best practices against disinformation

As we have seen in this report, disinformation has always existed, but the technological revolution and globalisation have greatly extended its reach and the consequences are potentially disastrous for individual safety and the stability of democratic institutions. In recent years there has been a growing awareness in the West about the need to counter disinformation, especially from illiberal states. However, disinformation is difficult to address in democratic societies as it involves fundamental rights such as freedom of expression and the right to information. The liberal starting point for tackling disinformation should not be over-regulation. Rather than invent new rights, we need to protect existing rights in the digital environment (Adsuara, 2020). The goal should be preserving the benefits of digitalisation while minimising its vulnerabilities and building a public-friendly digital sphere (Gardels, Giddens and Thorning-Schmidt, 2020).

Self-regulation by digital platforms to combat disinformation has generated mistrust and the suspicion that we may be moving towards a future in which the industry's monopolists together act as the judges of the digital environment by establishing and privatising censorship. The realisation that disinformation cannot be left solely in the hands of private platforms has led many countries to legislate to deal with the phenomenon. Here again, the democratic legitimacy of governments has not prevented criticism from the opposition, public opinion, and the media of what can be interpreted as a policy of social control that curtails fundamental freedoms and limits political criticism.

Social reaction to regulatory policies by private platforms and governments should be accepted as normal, necessary, and a symptom of democratic vigour. After all, a suspicion of power and a reluctance to expand government intrusion into the individual sphere are features of liberalism (Fawcett, 2014). However, such healthy mistrust must not lead to paralysis when faced by disinformation, but rather to a heightened standard of rigour, respect for fundamental rights, and the consensus needed to address such a sensitive issue.

By analysing several successes, we have seen that it is possible to take measures to counteract misinformation. Platforms have taken on greater responsibility and have recently started working to examine the content circulating on their networks, collaborating in this task with independent fact-checkers and experts. For their part, the traditional media have also seen the need to improve the quality of their information and gain the confidence and trust of users. Finally, the European Union has identified disinformation as a strategic issue and member states have begun to protect themselves.

Despite the efforts made, disinformation is far from being a disarticulated threat. When confronted with disinformation, there are more doubts than certainties, and with every measure taken to counter disinformation, the spreaders respond by sophisticating their messages to avoid controls. Respect for individual freedoms obliges us to reject forceful actions that could clash with fundamental rights, and the lack of solid empirical evidence limits our ability to act. In addition, the efforts required to counter disinformation often collide with the economic interests of platforms. Nevertheless, there are several things we can do in the fight against disinformation, but these actions require the commitment and participation of platforms, governments, and societies as a whole:

01. Checking fact-checkers

Taiwan has shown the important role that fact-checkers can play (CoFacts, for example, in partnership with Line). However, offloading all responsibility onto a single fact-checker can have negative consequences. Imagine, for example, that the government, or a media conglomerate, captures the fact-checker. In the absence of another fact-checker, disinformation could then be redirected and amplified to the benefit of this new actor. For this reason, we must include control mechanisms that maximise trust in fact-checkers.

This can be achieved with **two types of policies: firstly, by increasing the number of fact-checkers working with media and digital platforms** (competition is always a good strategy to improve quality standards, while capturing several fact-checkers is more difficult than capturing just one); and secondly, by independently auditing fact-checkers on a regular basis to assess if they are meeting the quality, transparency, and independence requirements.

Taiwan, as mentioned above, is the best example of this policy. The Taiwanese messaging app, Line, launched its own fact-checking bot. However, it was not the only watchdog. Citizens could send news to this bot, which passed the content to five independent fact-checking organisations. None of these fact-checkers had the absolute power to determine what was, and what was not, disinformation (Lange and Lee, 2020). This dispersion of control enables greater oversight and makes it difficult for fact-checkers to be captured by the government or others.

02. Digital literacy policy and awareness campaign.

The European Commission has recommended that member states adopt policies to prevent the harmful social effects of misinformation by launching digital and media literacy campaigns to educate the general public and schoolchildren. The Digital Education Action Plan Update (2021-2027) points in the same direction. The aim is to provide the public with the tools to identify disinformation and misinformation, and to recognise reliable media and information. This is not a miracle measure represents a costly policy whose benefits can only be assessed in the long term; however, it is supported by academics and experts (Livingstone, 2018).

One of the main difficulties when tackling disinformation has been the lack of awareness of its very existence. One of the objectives of the EuVsDisinfo strategy is to increase the European public awareness of the Kremlin's tactics in spreading disinformation. In the same vein, the digital divide (Barber, 2006) is problematic because it leaves many people vulnerable to disinformation and misinformation techniques. For this reason, literacy and awareness-raising are crucial.

→

Such measures have already been adopted by several countries. The Dutch government launched a public awareness campaign a few months before the last European elections to inform the public about how misinformation goes viral on the internet. The campaign was focused on social media, because it is the main vehicle used by misinformation promoters.

In terms of education, in July 2019, Britain announced the addition to the school curriculum of content that teaches pupils about disinformation and misinformation. The aim is for children to acquire the skills to identify reliable information and use the internet safely.

Outside Europe, several US states have passed similar legislation. This is the case of California, which introduced media literacy programmes in state schools. The measure was based on the work of a Stanford University researcher who showed that most students are unable to differentiate between information and sponsored content.

Australian authorities launched a campaign in the run-up to the federal election called 'Stop and Consider'. Its aim was to encourage critical thinking among citizens and bring attention to the sources of news.

03. Refuting and informing is better than prohibiting.

A liberal policy that is averse to interventionism and the temptation to establish an 'official truth' will rely on well-targeted refutation and the provision of reliable, pluralistic, and quality information – rather than prohibition. A fundamental task for media and internet platforms should be the elimination of content that breaks the law (for example, content that defrauds consumers, constitutes hate speech, or endangers the physical well-being of people). When faced with content that does not break the law, but could be considered misinformation, media and internet platforms should provide users with reliable complementary information and the context needed to form an informed opinion.

This policy can include interventions by platforms to encourage content from media outlets that have proven to be trustworthy by external evaluators because of their rigour, political independence, and funding transparency. At the same time, strong action is needed against botnets that viralise disinformation. Actions in this area are delicate because they can violate fundamental rights, but in the case of bots there must be no doubt: bots have no rights (Adsuara, 2020).

→

Action to prevent the excessive dissemination of certain messages have been successfully carried out – Operation Libero in Switzerland is a good example. By means of an offline and online presence, and always stressing the need to counteract hate speech and fake news with discussion that does not attack the sender, Operation Libero managed to limit the media power of the Swiss extreme right.

Based on the psychological foundations outlined above, the example of Operation Libero can help us design campaigns and policies that instead of attacking those who believe and share misinformation, succeed in counteracting their arguments.

04. Defending democracy: a public handbook.

Disinformation practices are a threat to liberal democracies, whether they come from other states, from foreign non-governmental actors, or from within the system. Liberal democracies must become aware of this danger and defend themselves. Some countries have already started to act.

This is the case of Sweden, which announced in 2018 the creation of a new authority dedicated to psychological defence, and whose work focuses on detecting and countering disinformation campaigns launched from abroad to interfere in electoral processes. The Swedish authority has an objective that differs from proposals in other countries because it is based on the aforementioned liberal ambition: it does not fight false information directly, but encourages verified quality content.

The Danish government also decided to protect electoral processes from foreign interference through the Defence Intelligence Service and the Danish police (working in cooperation with the media and political parties).

Outside Europe, the Canadian government created a Critical Election Incident Public Protocol that monitors foreign interference in digital space. This protocol involves senior government officials and is part of a rapid response mechanism organised by the Department of Foreign Affairs. In addition, the country launched its digital charter to promote trust in the digital environment. It states that: ‘the Government of Canada will defend freedom of expression and protect against online threats and disinformation designed to undermine the integrity of elections and democratic institutions’ (Funke and Flamini, 2020).

→

The value of civil society: while government action in defence of democracy is essential, countries also need a conscientious public that participates in civil associations. The benefit is twofold: firstly, **a vigorous and well-organised civil society is an indispensable element of liberal democracies for preventing the state** from using the fight against disinformation as an excuse to intrude on liberties and the individual sphere. Secondly, a vigilant society has a critical outlook and is therefore better prepared to withstand the onslaught of disinformation.

Even a democratic government may spread fake news to avoid accountability or damage the opposition. It is important to encourage civil society, which is the best safety net that liberal democracies have against abuses of power and those who wish to destabilise the nation. Operation Libero is a striking example of what we want to illustrate.

05. Liberalism for real people.

While education campaigns are much needed, we must be cautious about the prospects for success in countering misinformation and its political by-product – polarisation. There is not enough evidence to conclude that education and faith in reason will defeat the tribal instinct that fuels populism. On the contrary, education often serves to sophisticate the arguments and reinforce prejudices. Despite Macron's victory in 2017, or triumphs like that of Caputova in Slovakia, liberalism (in its widest definition as defenders of liberal democracy) has recently suffered many defeats.

However, we note this not in a spirit of defeat, but to indicate the aspects that can determine the success or failure of a liberal strategy. The excessive rationalism of liberalism can be a disadvantage against the emotional campaigns of populist leaders who mobilise the masses with sentimental appeals. A successful liberalism must also be a popular liberalism that remembers the value of historical narratives and emotions.

Speaking the emotional language that makes populists strong is important – as Emmanuel Macron's campaign demonstrated. In his campaign, the French president appealed to rationality, but also managed to awaken Europeanist pride and combine it with French pride. This type of discourse, which is emotional and proactive, as well as rational without renouncing identity, helps build a community and provide narratives that avoid the more perverse effect of disinformation.

We know that disinformation is especially successful in environments with high levels of affective polarisation. Liberalism should therefore try to steer the political agenda towards issues that are not especially confrontational and where the chances of generating political and social consensus are high. Reducing affective polarisation can help create more stable and civic political climates, as well as creating environments where misinformation finds it difficult to flourish.

06. A duty of care for platforms.

Authors such as William Perrin (2020) have pointed to the need to impose a 'duty of care' on social media multinationals. Such a policy implies **greater transparency about how the algorithms used by the platforms work – and this would help us understand how these algorithms act on the prioritisation of content** and which biases they introduce. It has also been proposed limiting the hyper-targeting of vulnerable groups (such as people with addictions or prone to depression), and a transnational regulation has even been suggested that would establish an independent authority with the power to sanction platforms that fail to apply duty of care policies (Gardels, Giddens, Thorning-Schmidt, 2020).

The authors of this proposal start from the position that social media companies, due to the nature of their business model, exploit attention and advertising – and this generates negative externalities in society. As an example of what they propose, they cite the Britain: where, it was decided that parliament will introduce a legal obligation for social media companies to prevent foreseeable harm to individuals. In addition, an independent regulator should oversee the whole process (UK Government, 2019).

07. A tax on platforms to finance quality local journalism.

This proposal is closely related to the duty of care. The proposal starts from the same assumption that social media companies pursue economic interests that generate negative externalities in society. One of these externalities is the erosion of the traditional media ecosystem (especially local media) and the constant drift of information consumption towards digital platforms (which are also behind the rise of disinformation).

The fall in advertising revenue for traditional media due to the emergence of platforms has put an end to a certain type of investigative and quality journalism. Money now flows to multinationals that do not generate local knowledge or accept any local responsibility (Steinberg, 2020). Consumers of news in traditional newspapers spend more time on average reading information than online readers (Thurman, 2017). In contrast, the viralisation of content takes place on platforms where people often indiscriminately share information after just reading a sensationalist clickbait headline. In addition, the platforms offer mixed bags of information and social or entertainment content that confuses users.

Some experts have called for a platform tax, the proceeds of which would be invested in the recovery of a local media ecosystem with quality content and research (Steinberg, 2020). In addition, in October 2020, the International Federation of Journalists launched a campaign to levy platform taxes and start an ad hoc fund to support and promote both journalists and the media. The measure has been adopted in Brazil.

08. Advertising transparency.

Closely related to previous measures is another aimed at advertising. Most of the advertising traffic that traditionally financed local journalism has shifted to platforms, which now control the lion's share of the adverting revenues. The consequences go beyond the closure of local newspapers. Unlike social media, traditional journalism had a social responsibility and a commitment to quality.

In recent years, platforms have become aware that a real commitment to the fight against disinformation needs greater control over advertising. In this regard, it is worth noting that Facebook has implemented a strategy with three levels of control for political or socially sensitive advertising:

a) Verification: complete documentation identifying the advertiser.

b) Labels added to advertisements specifying who paid for it.

c) Finally, this information is recorded in the Ad Library, where it is made available to researchers, governments, institutions, journalists, and anyone who wishes to consult it

This commitment has been prompted by governmental pressure. The United States Congress passed a clause in 2017 requiring platforms to retain and publish advertising contracts. Other countries, such as Israel, have banned the publication of anonymous advertising on the internet. Australia has imposed transparency measures on the funding of platform advertising (Funke and Flamini, 2020).

09. Working groups.

One proposal that several countries have already implemented is the creation of an intermediate body between government and civil society: working groups made up of experts to monitor the media, study the proliferation of disinformation, prevent and detect threats, and propose actions aimed at counteracting both disinformation and misinformation campaigns. Working groups can help monitor digital space and provide solutions – but appointments to these groups (where they come under government control) need to be transparent and public, especially if the findings of expert groups may have regulatory consequences.

One of the first countries to adopt such working groups was Britain, which decided to act after the 2016 Brexit campaign suffered interference from Russia. Belgium has also set up groups of journalists and academics to work on this issue. Outside Europe, Australia set up the Electoral Integrity Assurance Task Force in 2018, led by the Home Affairs Department, to identify potential cyber-attacks of foreign origin, especially in electoral processes (Funke and Flamini, 2020).

10. Online reporting portal.

Finally, in addition to the measures directed from governmental organisations, as well as the media, platforms, and academia, it is essential to involve the public in the fight against disinformation. Civil association are a key pillar of a strong civil society, but we also need tools that can enable individual collaboration. Online portals are an easily accessible tool for people to denounce harmful or false internet content, and obtain verified information on the viralisation of disinformation.

Italy has such a portal, where the public can report disinformation to the police. Belgium has also launched a disinformation reporting website where citizens can even vote on measures proposed by the government to fight disinformation and misinformation (Funke and Flamini, 2020).

There are no magic recipes for combating disinformation. The success of any regulation largely depends on whether **we can generate a shared vision of what disinformation represents** – while safeguarding fundamental rights and freedoms. The challenge is great: the line between a media outlet's legitimate editorial line and information manipulation can be very thin.

Legislation that creates new offences related to disinformation in overly broad terms and with disproportionate penalties may stop sources from talking to journalists, infringe freedom of expression, and even lead to the worst kind of censorship, namely, self-censorship. Encouraging quality reporting can be a difficult task: how do we set the criteria for truth and lies, how do we set standards of quality and public interest that generate consensus? Finally, how do we limit our tribal instincts and accept information that challenges our prejudices?

We do not have answers to everything, but we have accumulated enough knowledge in recent years to begin to understand the new media and platform ecosystem. We hope that this report can serve as a starting point for more work and research, and that it will contribute to promoting a pluralistic public discussion that can overcome polarisation and strengthen democratic structures. Finally, let the words of John Stuart Mill serve as an example of the liberal spirit that should guide us:

> *We can never be sure that the opinion we are endeavouring to stifle is a false opinion; and if we were sure, stifling it would be an evil still. First: the opinion which it is attempted to suppress by authority may possibly be true. Those who desire to suppress it, of course deny its truth; but they are not infallible. They have no authority to decide the question for all mankind, and exclude every other person from the means of judging. To refuse a hearing to an opinion, because they are sure that it is false, is to assume that their certainty is the same thing as absolute certainty. All silencing of discussion is an assumption of infallibility."*

Bibliography

Adsuara, B. (2020a): "La Guardia Civil no está para perseguir 'bulos', sino para perseguir delitos". *La Información,* 22 de abril. Available online at:
https://www.lainformacion.com/opinion/borja-adsuara/guardia-civil-perseguir-bulos-perseguir-delitos/6561145/

Adsuara, B. (2020b): "Las 4 'trincheras' para la defensa de la libertad de expresión". *La Información,* 4 de noviembre. Available online at:
https://www.lainformacion.com/opinion/borja-adsuara/las-4-trincheras-para-la-defensa-de-la-libertad-de-expresion-en-internet/2819958/

Adsuara, B. (2020c): "Lo contrario de la 'información veraz' no son las 'noticias falsas'". *La Información,* 8 de noviembre. Disponible *online* en: https://www.lainformacion.com/opinion/borja-adsuara/lo-contrario-de-la-informacion-veraz-no-son-las-noticias-falsas/6557781/

Allcott, H.; Gentzkow, M. (2017): "Social media and fake news in the 2016 election". *Journal of Economic Perspectives,* 31(2): 211-236.

Arteaga, F. (2020): "La lucha contra la desinformación: cambio de modelo". Real Instituto Elcano. *Disponible online* en: http://www.realinstitutoelcano.org/wps/portal/rielcano_es/contenido?WCM_GLOBAL_CONTEXT=/elcano/elcano_es/zonas_es/comentario-arteaga-la-lucha-contra-la-desinformacion-cambio-de-modelo

Barber, B. (2006): "How democratic are the new telecommunication technologies". IDP: *Revista de los Estudios de Derecho y Ciencia Política de la UOC,* 3, p. 6.

Barberá, P. (2015): "Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data". *Political Analysis,* 23(1): 76-91.

Beam, M. A.; Haridakis, P. M.; Hutchens, M. J.; Hmielowski, J. D. (2017): "Social media, news platforms, and partisan exposure: Voters' media preferences during the 2016 presidential campaign season". En: *The Presidency and Social Media,* pp. 37-55. Routledge.

Berggruen Institute (2020): Renewing democracy in the digital age. Marzo. Available online at: https://www.berggruen.org/docs/BER_2020_RenewingDemocracy.pdf

Caplan, B. (2011): The Myth of the Rational Voter: Why *Democracies Choose Bad Policies-New Edition.* Princeton University Press.

Cembrero, I. (2020): "Desinformación 'made in Russia': un desafío para las democracias europeas". *El Confidencial,* 11 de noviembre. Available online at: https://blogs.elconfidencial.com/mundo/la-historia-no-acaba-aqui/2020-11-11/desinformacion-rusia-desafio-democracias-europeas_2827024/

Clegg, N. (2020): "Combating COVID-19 Misinformation Across Our Apps". *Facebook,* 25 de marzo. Available online at: https://about.fb.com/news/2020/03/combating-covid-19-misinformation/

Comisión Europea (2018a): *Tackling online disinformation: a European Approach.* 26 de abril. Available online at:
https://ec.europa.eu/digital-single-market/en/news/communication-tackling-online-disinformation-european-approach

Comisión Europea (2018b): *The Code of Practice on Disinformation.* 26 de septiembre. Available online at:
https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation

Comisión Europea (2019): *The Action Plan Against Disinformation.* Junio. Available online at:
https://ec.europa.eu/commission/sites/beta-political/files/factsheet_disinfo_elex_140619_final.pdf

De Gruyter, C. (2017): *The Rise of Europe's Antipopulists. Carnegie Europe.* Available online at:
https://carnegieeurope.eu/2017/04/25/rise-of-europe-s-antipopulists-pub-68764

Downs, A. (1957): *An Economic Theory of Democracy.* Harper Collins Publishers.

EUvsDisinfo (2015): Available online at: https://euvsdisinfo.eu/

Festinger, L. (1957): *A Theory of Cognitive Dissonance.* Vol. 2. Stanford University Press.

Fletcher, R.; Cornia, A.; Graves, L.; Nielsen, R. K. (2018): "Measuring the reach of 'fake news' and online disinformation in Europe". *Reuters Institute Factsheet.* Febrero. Available online at: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-02/Measuring%20the%20reach%20of%20fake%20news%20and%20online%20distribution%20in%20Europe%20CORRECT%20FLAG.pdf

Freedom House (2020): *Freedom on the Net 2020: The Pandemic's Digital Shadow.* Available online at: https://freedomhouse.org/sites/default/files/2020-10/10122020_FOTN2020_Complete_Report_FINAL.pdf

Funke, D.; Flamini, D. (2020): "A guide to anti-misinformation actions around the world". *Poynter,* diciembre. Available online at: https://www.poynter.org/ifcn/anti-misinformation-actions/

Giddens, A. (2018): "A Magna Carta For The Digital Age". *The Washington Post,* 2 de mayo. Available online at: https://www.washingtonpost.com/news/theworldpost/wp/2018/05/02/artificial-intelligence/

Giorio, L. (2018): *War on Propaganda or Propaganda War?: A case study of fact-checking and (counter) propaganda in the EEAS project EUvsDisinfo.*

Gray, J. (2019): "Why liberals now believe in conspiracies". New Statesman, 14 de agosto. Available online at: https://www.newstatesman.com/politics/uk/2019/08/why-liberals-now-believe-conspiracies

Henley, J. (2019): "Change the narrative: how a Swiss group is beating rightwing populists". *The Guardian.* 7 de abril. Available online at: https://www.theguardian.com/world/2019/apr/07/we-had-to-fight-operation-libero-the-swiss-youth-group-taking-on-populism

Hobbs, R. (2010): Digital and Media Literacy: *A Plan of Action. A White Paper on the Digital and Media Literacy Recommendations of the Knight Commission on the Information Needs of Communities in a Democracy.* Aspen Institute. Available online at: ver link

Howard, P. N. (2020): *Lie Machines: How to Save Democracy from Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives.* Yale University Press.

Iyengar, S.; Lelkes, Y.; Levendusky, M.; Malhotra, N.; Westwood, S. J. (2019): "The origins and consequences of affective polarization in the United States". *Annual Review of Political Science*, 22, 129-146.

Jeong, S. H.; Cho, H.; Hwang, Y. (2012): "Media Literacy Interventions: A meta-analytic review". *Journal of Communication,* 62(3): 454-472.

Jourová, V. (2020): Discurso de apertura de la vicepresidenta Věra Jourová en la conferencia *"Disinfo Horizon: Responding to Future Threats".* 30 de enero. Available online at: https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_20_160

Kahneman, D. (2011): *Thinking, Fast and Slow*. Macmillan.

Kerr, W.; Phillips, M. (2020): "Taiwan is beating political disinformation. The West can too". *Foreign Policy,* 11 de noviembre. Available online at: https://foreignpolicy.com/2020/11/11/political-disinformation-taiwan-success/

Lange, E.; Lee, D. (2020): "How One Social Media App Is Beating Disinformation". *Foreign Policy,* 23 de noviembre. Available online at: https://foreignpolicy.com/2020/11/23/line-taiwan-disinformation-social-media-public-private-united-states/

Lessenski, M.; Kavrakova, A. (2019): *Societies outside Metropolises: the role of civil society organisations in facing populism.* Estudio del Comité Económico y Social Europeo. Available online at: https://www.eesc.europa.eu/sites/default/files/files/qe-04-19-236-en-n.pdf

Lewis, M. (2016): *The Undoing Project: A friendship that changed the world.* Penguin UK.

Li, J.; Wagner, M. W. (2020): "The Value of Not Knowing: Partisan Cue-Taking and Belief Updating of the Uninformed, the Ambiguous, and the Misinformed". *Journal of Communication,* 70(5): 646-669.

Livingstone, S. (2018): "Media literacy – Everyone's favourite solution to the problems of regulation". *LSE blog,* 27 de julio. Available online at: https://blogs.lse.ac.uk/parenting4digitalfuture/2018/07/27/media-literacy-problems-of-regulation/

LSE Commission on Truth, Trust and Technology (2018): *Tackling the Information Crisis: A Policy Framework for Media System Resilience.* LSE. Available online at: https://www.lse.ac.uk/media-and-communications/assets/documents/research/T3-Report-Tackling-the-Information-Crisis.pdf

Martens, H.; Hobbs, R. (2015): "How media literacy supports civic engagement in a digital age". *Atlantic Journal of Communication,* 23(2): 120-137.

Perrin, W. (2018): *Harm Reduction in Social Media – A Proposal.* Carnegie UK. Available online at: http://www.carnegieuktrust.org.uk/blog/social-media-regulation/

Rawls, J. (1995): *Liberalismo político.* UNAM.

Rawls, J. (2012): *Teoría de la justicia.* Fondo de Cultura Económica.

Revealing Reality (2018): Scrolling News: *The changing face of online news consumption. A report for Ofcom.* Available online at: https://www.ofcom.org.uk/__data/assets/pdf_file/0022/115915/Scrolling-News.pdf

Rosen, G.; Lyons, T. (2019): "Remove, Reduce, Inform: New Steps to Manage Problematic Content". *Facebook,* 10 de abril. Available online at: https://about.fb.com/news/2019/04/remove-reduce-inform-new-steps/

Shearer, E. (2018): "Social media outpaces print newspapers in the US as a news source". *Fact Tank,* 10. Pew Research Center. Available online at: https://www.pewresearch.org/fact-tank/2018/12/10/social-media-outpaces-print-newspapers-in-the-u-s-as-

Solaz, H.; De Vries, C. E.; de Geus, R. A. (2019): "In-group loyalty and the punishment of corruption". *Comparative Political Studies,* 52(6): 896-926.

Soll, J. (2016): "The Long and Brutal History of Fake News". *Politico,* 18 de diciembre. Available online at: https://www.politico.com/magazine/story/2016/12/fake-news-history-long-violent-214535

Steger, I. (2018): "How Taiwan battled fake anti-LGBT news before its vote on same-sex marriage". *Quartz.* 22 de noviembre. Available online at: https://qz.com/1471411/chat-apps-like-line-spread-anti-lgbt-fake-news-before-taiwan-same-sex-marriage-vote/

Tetlock, P. E. (2017): Expert Political Judgment: *How Good Is It? How Can We Know?* Princeton University Press.

Thaler, R. H.; Sunstein, C. R. (2009): *Nudge: Improving Decisions about Health, Wealth, and Happiness.* Penguin.

Thurman, N. (2018): *Time Spent Reading News Brands: Print–Digital Performance, Online-Only Effects, and Long-Term Trends.* World Association of Newspapers and News Publishers (WAN-IFRA). Available online at: https://neilthurman.com/files/downloads/journalists-in-the-uk.pdf

Thurman, N.; Aquilina, J. (2020): *BBC Three's reinvention online: Effects on viewing and content.* Londres: Enders Analysis. Available online at: https://neilthurman.com/files/downloads/time-spent-reading-newsbrands.pdf

Thurman, N.; Lewis, S.; Kunert, J. (eds.) (2019): *Algorithms, Automation, and News: New Directions in the Study of Computation and Journalism.* Londres: Routledge

UK Government (2019): *Online Harms White Paper.* 8 de abril. Available online at: https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper

Vosoughi, S.; Roy, D.; Aral, S. (2018): "The spread of true and false news online". *Science,* 359(6380): 1146-1151.

Volkova, S.; Jang, J. Y. (2018): "Misleading or falsification: Inferring deceptive strategies and types in online news and social media". En: *Companion Proceedings of The Web Conference 2018,* pp. 575-583).

Wang, C.; Huang, H. (2020): "When 'Fake News' Becomes Real: The Consequences of False Government Denials in an Authoritarian Country". *Comparative Political Studies,* septiembre.

Wendling, M. (2018): "The (Almost) Complete History of 'Fake News'". *BBC News,* 22 de enero. Available online at:https://www.bbc.com/news/blogs-trending-42724320

Woodhouse, J. (2019): *Social media, how much regulation is needed?* UK Parliament. Available online at: https://commonslibrary.parliament.uk/social-media-how-much-regulation-is-needed-2/

CONTACT

Esade Business School
Avda. Pedralbes, 60-62
E-08034 Barcelona · Spain
Phone: +34. 93 280 6162
dobetter@esade.edu
www.esade.edu

EsadeBS
esade
esade
esade
_esade

Friedrich Naumann Foundation for Freedom
Calle de Fortuny 3, 1º izquierda
28010 Madrid · Spain
Phone: +34 913 08 94 80
madrid@freiheit.org
www.freiheit.org

fnfmad
fnfmad
Friedrich Naumann Foundation Madrid
MD Go Friedrich Naumann Foundation Madrid
fnfmad

esade

EsadeEcPol - Center for Economic Policy

FRIEDRICH NAUMANN FOUNDATION For Freedom.

Spain, Italy, Portugal

About Esade

Founded in 1958, Esade is a global academic institution, with campuses in Barcelona and Madrid, and present throughout the world through collaboration agreements with 185 universities and business schools. Each year, more than 11,000 students participate in its courses, in the three training areas: Business School, Law School and Executive Education. Esade Alumni, Esade's alumni association, has more than 60,000 alumni and has an international network of 72 chapters, with alumni of up to 126 nationalities, present in more than 100 countries. Esade also participates in the Esade Creapolis business innovation park, a pioneering ecosystem that aims to inspire, facilitate and accelerate the innovation processes of the companies that participate in it. With an international vocation, Esade occupies prominent positions in the main world rankings of business schools such as Financial Times, QS, Bloomberg Businessweek or América Economía. Esade is a member of the Ramon Llull University.

About the Friedrich Naumann Foundation for Freedom (FNF)

FNF is a German foundation dedicated to the promotion of political education and liberal principles, such as human rights, free market economy, the rule of law and democracy. The foundation is working to promote the principles of freedom and dignity for all people in all areas of society in more than 60 countries around the world (Europe, Africa, Asia, the Americas). The Madrid office seeks to strengthen cooperation and political dialogue between political parties, scientific institutions and civil society organisations from Spain, Italy and Portugal and the European sphere. Committed to contribute to solutions to specific regional challenges in Southern Europe, FNF Madrid strives to highlight best practices from the region for the European integration process.