INSIDE MISINFORMATION

HOW INDIVIDUALS ENGAGE WITH ONLINE MISINFORMATION

Sophie Morosoli

Faculty of Social Science

Department of Political Science

Inside Misinformation

How Individuals Engage With Online Misinformation

PhD thesis submitted for the degree of doctor in social science: political communication at the University of Antwerp

To be defended by

Sophie Morosoli

Promotor:

Peter Van Aelst

Antwerp, 2022

Members of the Doctoral Commission

Prof. dr. Karolien Poels, University of Antwerp

Dr. Edda Humprecht, University of Zurich

Dr. Julie Sevenans, University of Antwerp

Members of the Doctoral Jury

Dr. Desirée Schmuck, University of Leuven

Dr. Michael Hameleers, University of Amsterdam

TABLE OF CONTENTS

Funding

This work was supported by the Swiss National Science Foundation and the Research Foundation Flanders (Grant No. 100017L_182253). This was called “From Uninformed to Disinformed Citizens? Comparing Western Information Environments”. Edda Humprecht and Frank Esser (University of Zurich) are the principal investigator of the project in Switzerland and Peter Van Aelst (University of Antwerp) in Belgium.

CHAPTER ONE
INTRODUCTION

Introduction

> "Facts are beyond agreement and consent, and all talk about them – all exchanges of opinion based on correct information – will contribute nothing to their establishment."
> – Hannah Arendt in Truth & Politics (1967)

I could not have written this dissertation about the relationship between individuals and online misinformation without referring to Hannah Arendt. The thoughts and ideas of one of the most original thinkers of the past century on truth, lies, and politics seem nowadays as timely as ever. In her essay, *Truth & Politics* (1967), Arendt demonstrates the fragility of facts and the dangers of organized lying in politics. Organized lying will not only try to replace the truth with an untruth, but goes further by undermining the 'factual fabric of reality' (EUvs.Disinfo, 2021). Lies, propaganda, and misinformation can take on many shapes and forms and are as old as democracy itself, as Arendt (1967) further illustrates. Recently, we have witnessed war crimes committed by the Russian army being framed as a plot against Russian soldiers carried out by the Ukrainian government involving actors playing victims of war (Politifact, 2022). The Russian government is not only openly waging a war against Ukraine with its army on the ground, but also a (mis-)information war. The Russian propaganda machine targets international and national audiences to justify the invasion. To reach these audiences, the manipulated and misleading messages of Russia's disinformation campaign are strategically spread on social media. For instance, the Guardian expressed doubts about the war crimes committed by Russia in Ukraine and Der Spiegel warned about gas shortages (Scott, 2022). However, these articles were all fake. The posts shared on social media mimicked articles of legitimate Western news providers and these sites were especially promoted on Facebook where they reached countless individuals (Scott, 2022).

During the COVID-19 pandemic, the WHO coined the term 'infodemic' to describe the excess of true and false information, which was spread simultaneously with the coronavirus. A study of the Pew Research Center reveals that in 2020 around half of U.S. citizens indicated that they have been exposed to online misinformation regarding COVID-19 (Mitchell & Oliphant, 2020). One reason for this finding might be that a multitude of misleading messages went viral during the outbreak of the virus. One of those stories involves someone's uncle's advice on the coronavirus. It contained wrong advice against COVID-19 that should apparently prevent people from getting infected and gave a false sense of security by leading individuals to believe to not have the virus if they show certain symptoms (e.g. runny nose). The post went viral and was shared over 300'000 times before the author of the post

removed the false claims that the post contained (Full Fact, 2020). However, the original post was seen by hundreds of thousands of people.

By any means, false and misleading information, spread unconsciously or not, can have real world consequences. On December 4th 2016, the story of a pedophile ring run out of a pizza parlor by Hillary Clinton and other members of the Democratic Party made a man called Edgar Welch storm the pizzeria with his rifle drawn, in the hopes of breaking up this pedophile ring. Luckily, nobody got hurt. More serious, on January 6th 2021, an angry mob of Donald Trump's supporters stormed the Capitol Building in Washington D.C., disturbing a crucial aspect of democracy and killing multiple people. The reason for this attack on democracy and the raising doubt about democratic values was the false claim spread by Donald Trump that the 2020 elections were 'stolen' from him because of serious voter fraud. According to Oliphant and Kahn (2021) half of the Republicans still believe that the storm on the capitol was non-violent, or that it was an attempt of left-wing activists to show Trump in a bad light. Even more striking is the fact, that six out of ten Republicans believed in Donald Trump's misinforming claim that there was an incredible voter fraud going on, leading to Trump losing the 2020 presidential elections. Against the backdrop of these findings and with Donald Trump's enormous reach through social media, Fox News, and other media outlets, it makes it somewhat easier to comprehend why this large group of people came together to storm the Capitol shortly after the elections.

Events like this have shown us that the belief in misinformation can have extremely harmful real-life consequences for individuals and have raised the question about how to take on this worldwide societal phenomenon. A considerable amount of scientific literature has therefore focused on the impact that misinformation can have on people, especially in connection with the coronavirus. Because in the case of the COVID-19 pandemic, the effects of misinformation are not only about opinions, but a matter of public health. For instance, in a number of countries political leaders did little to mitigate the spread of COVID-19 related misinformation and sometimes even fostered it. For example, the Brazilian president Jair Bolsonaro promoted during the pandemic false information about the effects and treatment of the virus, giving the Brazilian population a false sense of security just to keep the economy running (Ricard & Medeiros, 2020). Research has shown that individuals who are more susceptible to misinformation about COVID-19 tend to have lower trust in scientists, politicians and journalists, are politically right leaning, and are rather young (Roozenbeek et al., 2020). People with a higher susceptibility are consequently less willing to get vaccinated and less willing to recommend the vaccine to their peers (Roozenbeek et al., 2020). General findings suggest, if individuals believe in COVID-related misinformation, they cannot form a correct assessment of the scope and the reality of the pandemic (Pantazi et al., 2021).

Overall, Lewandowsky and van der Linden (2021) describe the effect of misinformation very fitting – it 'sticks'. What the authors mean by that is, that false and misleading information stays in the minds of people and it is complicated to erase this kind of information out of one's memory (Lewandowsky & van der Linden, 2021). Even after being presented with corrective interventions, individuals rely on the initially seen misinforming content. This it called the 'continued influence effect of misinformation' (Chan et al., 2017; Lewandowsky et al., 2012; Lewandowsky & van der Linden, 2021). One reason for the persistence of misinformation is the familiarity with it. In the sense that individuals might think that they have heard of the false claim before, so there must be some truth to it (Ecker et al., 2020). Thus, misinformation is tough to erase from people's memory and can, on the contrary, even alter it (Pena et al., 2017).

The above discussed findings on the impact of misinformation on individuals and the named incidents were also the main reasons to start investigating the driving forces behind the spread of misinformation on the microlevel. How is it possible that these "unbelievable" events are unfolding, and why are people supporting or believing clearly false and misleading information? What happens when individuals are exposed to such harmful content, and who would further spread it, helping to expose even more people to it? The annexation of Crimea in 2014 or the 2016 presidential elections have revived the field of propaganda research and brought forward different strands of research on mis- and disinformation, which were again enriched by other affairs such as the Brexit referendum or the global COVID-19 pandemic. A considerable strand of research measures the exposure to misinformation and the consequences of it, for example, the belief in misinformation (see for example Halpern et al., 2019; Ognyanova et al., 2020). Next to exposure, scholars have increasingly studied the correction or debunking of false and misleading messages. This field of research has looked into the potential of ways to fight misinformation, for instance, through news literacy and fact checking, (e.g. Ecker et al., 2020; Hameleers & van der Meer, 2020; Lewandowsky & van der Linden, 2021; Vraga et al., 2020). However, little is known about why and how individuals go against misinformation on social media platforms, to warn others, for instance. We know by now that mis- and disinformation are mainly and effortlessly spread online and on social media platforms (e.g. Allcott & Gentzkow, 2017; Lazer et al., 2018). Against this backdrop, this dissertation takes an interest in the online dissemination of misinformation on the microlevel. It is crucial to investigate the microlevel because ordinary citizens are able to mitigate or accelerate the spread of false and misleading information. Psychologically, the question of why individuals believe false information to be true and therefore share it within their online network has gotten some scholarly attention. However, much about the exposure to online misinformation and its consequences has hardly been explored. For instance, the different motivations behind the spread of online misinformation, such as agreement with the message or entertainment

reasons, have only recently sparked the interest of the scientific community. This dissertation adds to filling this gap. In addition, it is important to provide more insight in the individuals who spread online misinformation. Even though research has singularly looked at sociodemographics or political orientation, this dissertation combines approaches from different disciplines to investigate a broader range of variables, including psychological factors such as personality traits. Hence, by answering who and why individuals would spread online misinformation, this dissertation can increase knowledge that can be used to combat the dissemination of misinformation.

Misinformation touches many different aspects of public life: political processes, public health, science, the environment and many more. It is therefore no surprise that the field of misinformation research is very diverse and multiple scientific disciplines are involved. For example, psychology scholars have, among other things, focused on the processes and consequences of individual misinformation consumption (e.g. Pennycook & Rand, 2021) and political scientists investigated the influence of misinformation on elections and other political processes (e.g. Grinberg et al., 2019). Field-specific theories do not exist. However, I identified some reoccurring theoretical frameworks used in previous research. Because misinformation is, among other things, a communication phenomenon, mass media and communication science provides useful theoretical backgrounds. Existing communication theories such as the third person effect, framing, the uses and gratifications approach, and the truth default theory have all been applied in misinformation research (Apuke & Omar, 2020; Koo et al., 2021; Rooke, 2021). Second, psychology offers some fruitful concepts, which misinformation studies have been relying on. Mainly motivated reasoning and the confirmation bias have been linked to studying misinformation (Hameleers & van der Meer, 2020; Miller et al., 2016). In addition, have personality traits been tested as a possible explanation for why individuals would engage with false and misleading content (Chen, 2016). Lastly, political science further contributes to the field. Concepts like partisanship and political polarization have been intensively studied in connection with misinformation (Hameleers & van der Meer, 2020; Osmundsen et al., 2020). These theoretical frameworks are by no means exclusive and other research on misinformation applies different theories, but they indicate the broader theoretical orientation of the field.

In this dissertation, I will combine different theoretical frameworks throughout to show the adaptability and dynamics of the field. By using different theoretical approaches and concepts, I can contribute to a better understanding of which frameworks work the best when studying individual behavior in the context of online misinformation. This dissertation empirically tackles four central aspects of the relationship between individuals and online misinformation. A considerable number of existing studies already focuses on individuals in the context of misinformation. However, existing research typically highlights only one side of the interplay between individuals and misinforming

content. For example, only focusing on individual debunking strategies and not on additionally considering the opposite – individuals who ignore or even consciously support misinformation. This study characterizes not only who and why people are willing to spread misinformation, but also who and why not. Furthermore, the thesis takes a closer look at the *how*. If confronted with real misinforming social media posts, how do individuals usually react? Do they support the content, or are they willing to warn others? Knowing more about how individuals react to online misinformation will, for instance, enable policymakers, journalists, and other stakeholder to produce appropriate tools to combat misinformation.

Zeroing in on these four overarching areas, will allow me to paint a broad picture of who and how individuals cope with online misinformation. This leads to the following four concrete research questions, which will be answered in four different chapters throughout this dissertation.

RQ1: Who engages with online misinformation?

RQ2: Why do individuals engage with online misinformation?

RQ3: How do individuals engage with online misinformation?

RQ4: Who counters online misinformation and why?

Answering these four research questions is not only relevant to capture the essence of the relationship between individuals and online misinformation, it is also vital in the fight against misinformation. Knowing who is susceptible to misinformation or what type of individuals are likely to disseminate or counter misinformation, can help create specific policy responses or other efforts to combat online misinformation tailored specifically to these individuals.

This dissertation defines misinformation as false or misleading information, which is created or spread without the deliberate intention to harm (Wardle & Derakhshan, 2018). Thus, misinformation is not intentionally misleading. I will use this term because frequently people share this type of information without knowing whether it is false and in that sense do not know that they are disseminating a falsehood. There are many different understandings of the term, and it is intrinsically related to other concepts such as disinformation that refers to the intentional and harmful use of false information. Therefore, I will elaborate here on the inconclusive efforts to clearly define the concept of misinformation.

In the remainder of the introduction, I will first discuss the terms and definitions used in the field of misinformation research. Next, I address the dissemination of misinformation online with a special focus on the individuals who engage with this type of content, and on the relationship between social

media platforms and misinformation. Also, possible counter efforts against online misinformation will be discussed. Finally, I describe the main goal and contributions of the dissertation while referring to the selected cases and applied methods.

Terms and Definitions[1]

The debate around the term 'misinformation' has been ongoing and has been affected by some conceptual confusion. This is in that sense problematic that different understandings of the supposed same concept can lead scholars to measure different things while building upon research which measured something completely different. Based on the results of a systematic literature review (see Appendix 1), we will disentangle the different terms and definitions used in connection with misinformation research. One central observation is that in recent studies scholars have used the same terms to describe different concepts (Egelhofer & Lecheler, 2019) and there seems to be no shared understanding regarding the terminology yet (Kapantai et al., 2020; Mosinzova et al., 2019). For example, the term 'fake news' has been used to describe the concept of disinformation (Allcott & Gentzkow, 2017), news satire such as The Daily Show (Reilly, 2012), or dubious and unprofessional 'news' sources online (Grinberg et al., 2019). Not only are the same terms used to describe different concepts, different terms are also used to define the same concept. Disinformation, misinformation and fake news have in the past all been defined as fabricated misleading information (Allcott & Gentzkow, 2017; Humprecht, 2019). With regard to this definitional confusion, Kapantai et al. (2020) even go as far as stating that some of the terms have been mis- and abused in the recent past by different actors such as scholars, journalists and politicians. Additionally, we find that only half of the consulted studies even define the respective concepts.

To structure the amount of different understandings within the empirical research, we mainly rely on the conceptualization of Wardle and Derakhshan (2018) to define mis- and disinformation, since coherently defining the terms has been challenging for the field (Vraga & Bode, 2020). We understand misinformation as incorrect information that is not intentionally shared to harm (Allcott & Gentzkow, 2017; Nielsen & Graves, 2017; Wardle & Derakhshan, 2018), in contrast to disinformation, which we define as willfully produced incorrect or misleading information to serve a strategical purpose, deceive and to harm (Humprecht, 2019; Wardle & Derakhshan, 2018). To characterize the term 'fake news', we reference a recently introduced conceptualization, which suggests that the concept is two-dimensional (Egelhofer & Lecheler, 2019). The authors describe fake news as a genre, "describing the deliberate creation of pseudojournalistic disinformation, and there is the (2) fake news label,

[1] Some insights of this chapter are based on a systematic literature review I have written together with Anna Staender in 2019. I am very grateful for her support and brilliant insights. The figures and detailed description of the results can be found in the appendix of this chapter.

describing the political instrumentalization of the term to delegitimize news media" (Egelhofer & Lecheler, 2019, p. 97). Another concept, which is also used in the context of misinformation research, are conspiracy theories. We adapt the definition of Keeley, which defines conspiracy theories as "a proposed explanation of some historical event (or events) in terms of the significant causal agency of a relatively small group of persons the conspirators acting in secret" (1999, p. 116). According to this understanding, a group of people is trying to benefit politically or economically by acting in secret, harming the common people and violating certain rights (Uscinski et al., 2016). For instance, the widely circulated conspiracy theory that people in powerful positions intentionally planned and executed the global COVID-19 pandemic (Schaeffer, 2020). It becomes apparent and is easy to imagine that dis- and misinformation can be linked to conspiracy theories. False and misleading information plays effortlessly into the narratives of conspiracy theories, and individuals who believe in conspiratory information are consequently misinformed.

More recently, the field has undertaken some efforts to synthesize the different understandings of the terms and a consensus seems to develop (Kapantai et al., 2020; Tandoc et al., 2018; Vraga & Bode, 2020). These literature reviews offer valuable suggestions for how to best define and study the different concepts and hereby unify the field. Vraga and Bode (2020) advocate to use two criteria to contextualize misinformation: evidence and expertise. Information either aligns with experts or with evidence, which is then considered as correct information, or it does not align with these two criteria and is consequently understood as misinformation (Vraga & Bode, 2020). However, this suggests that misinformation is linked to an universal understanding of what is true and what is false, but misinformation is context dependent. What is true for one person, considering different expert beliefs or a different information environment, might be inaccurate for another person. Revisiting this thought is essential when studying misinformation.

In this dissertation I focus on the spread of false and misleading information[2] and not on the production side of this type of content. Although we know that many false and misleading stories are created for commercial or political reasons, the intentionality of the producers of such information is often not apparent to scholars or impossible to measure. This is even more the case for the dissemination of this content: we often simply do not know whether individuals spread false information with the intention to harm. This harm can take on many shapes and forms. For instance, individuals might share information, which they know is false, for personal gain or to spread hate against out-groups. Based

[2] The discussion about what is precisely true or false is complex. In the context of this dissertation, I often refer to 'false and misleading' information. I consider false and misleading information as information that has been created and is factually incorrect or accurate information that has been taken out of context to mislead people. This type of information can be proven inaccurate by trustworthy sources such as, journalists, fact checkers or scientists.

on this, I decided to use the term misinformation throughout this dissertation and measure the concept based on the understanding of misinformation as being false and misleading information that is spread or created without the deliberate intention to harm (Wardle & Derakhshan, 2018). Misinformation looks accurate and authentic, and can therefore lead to individuals believing in false information (Hameleers & van der Meer, 2020).

Social Media and the Spread of Misinformation

The reason why misinformation is so successful lies in its ability to spread like wildfire. Especially on social media, it is easy to contribute to the dissemination of misinformation. Through likes, shares, and comments, false and misleading content can travel fast and far. From personal experience and research, we know that the majority of the people use social media platforms daily, and they make a big part of their daily lives. In the past years, social media have become more and more important for people's news consumption (Newman et al., 2022). As of July 2022, Twitter announced its platform had 238 million monetizable daily active users around the world (Twitter, 2022) and Facebook reported to have 1.97 billion daily active users in June 2022 (Meta, 2022). This shows the incredible scope of two of the biggest social media platforms. Meanwhile, Instagram and TikTok are also on the rise, attracting more and more users. In 2021, Instagram had 1.21 billion monthly active users and TikTok 655.9 million users worldwide with an increasing trend for 2022 (Statista, 2022a, 2022b).

Generally speaking, social media platforms can be seen as an amplifier of information – true and false (Allcott & Gentzkow, 2017; Islam et al., 2020). According to Rossini et al. (2020), there are two main reasons why social media platforms alter the way we consume information compared to traditional media. First, a big part of the social networks on these platforms are weak ties (e.g. coworkers), collected over time, which leads to the possible exposure to diverse opinions and beliefs. Thus, we are confronted with a multitude of viewpoints and do not only see information that is consistent with our preexisting attitudes. Second, social media users who are not looking for news might either way be confronted to it because of how the people in their network select and share news. This means that people, who share a lot of selected (political) news, can somewhat influence what others see (Rossini et al., 2020). It becomes evident that social media platforms are not just a fertile ground to spread political information, but also political *mis*information. One of the dangers of misinformation, which is circulating on social media, is that individuals are more prone to believe false information when it is shared by their friends or other trusted contacts in their networks (Metzger et al., 2021). And once individuals believe in misinforming messages, they are also more likely to further promote them on social media, exposing their friends to it. In the 2021 Digital News Report of the Reuters Institute (2021), over 40% of the participants indicated that they have been exposed to misinformation and

almost 60% are concerned about false and misleading content. In the UK, a staggering 42% of individuals admitted having shared misinformation at one point (Chadwick & Vaccari, 2019). It is no surprise that these observations contributed to the fact that citizens are becoming more and more concerned about online misinformation – all over the world (Newman et al., 2020, 2021). This worry led U.S. citizens to increasingly wish for governmental help in restricting online misinformation (Pew Research Center, 2021). Even if they would have to give up certain levels of access or freedom to produce content, almost half of the participants wanted the government to take action against online misinformation (Pew Research Center, 2021).

As a response to the effortless spread of misinformation on social media, fact checking organizations are trying to intervene by posting corrections on the platforms. Research has shown that offering compelling corrections which are coherent and easy narratives are an effective way to counter online misinformation (Lewandowsky et al., 2012). Furthermore, social media platforms such as Facebook have begun to flag articles to warn users of false and misleading information but eventually stopped with it and added the term 'related information' to questionable posts (Garrett & Poulsen, 2019; Kirchner & Reuter, 2020). However, several scholars observed that debunking might even increase the belief in misinformation (e.g. Berinsky, 2017; Nyhan & Reifler, 2010). Garrett and Poulsen (2019) found that warnings from fact checking organizations were not effective. More in general, individuals might be resistant to these warnings, when they counter their political ideology (Flynn et al., 2017; Garrett & Weeks, 2013). Thus, research is still somewhat inconclusive regarding debunking or corrective measures and social media platforms seem therefore to be the 'perfect' tool to disseminate misinformation. Despite this ambiguity about debunking measures, it is essential to know which characteristics individuals display who engage with online misinformation and which motivations are behind this behavior to effectively implement different debunking strategies.

Individuals who Spread Misinformation

The first main goal of this dissertation is to better understand the 'profile' of individuals that engage with online misinformation (chapter 1 & 2). Focusing on the microlevel, a large part of the scholarly community is interested in the individuals who like or share false messages on social media and therefore (willingly or not) contribute to the spread of online misinformation. What are the reasons behind this online behavior, and what kind of individual characteristics do these people display?

Who?

Previous research has explored different characteristics that individuals, who disseminate online misinformation, might exhibit. An evident starting point are sociodemographics, which are not only applied as control variables. Numerous studies have found sociodemographic factors to be valuable in

researching the spread of false and misleading content (Chadwick & Vaccari, 2019; Grinberg et al., 2019; Guess et al., 2019). For instance, results indicate that older social media users are more likely to engage with misinformation than younger individuals (Grinberg et al., 2019; Guess et al., 2019; Osmundsen et al., 2020). Closely related to that are political variables, which are crucial factors to get a better idea about who engages with misinformation. Especially political orientation and political interest have been highlighted in previous research (Bennett & Livingston, 2018; Chadwick & Vaccari, 2019; Van Bavel et al., 2021). Numerous studies find, for example, that conservative individuals were more prone to share inaccurate information and were less willing to engage with fact checked claims (Guess et al., 2019; Osmundsen et al., 2020; Shin & Thorson, 2017). Because misinformation spreads easily through social media, users' general social media behavior was previously taken into account by scholars when investigating people's propensity to disseminate false or misleading content. More recently, research focused on psychological factors and personality traits. Concepts like the fear of missing out (FoMO), social media fatigue, and dark personality traits have been proven to be valuable predictors for the sharing of misinformation (Balakrishnan et al., 2021; Buchanan, 2020b; Buchanan & Benson, 2019; Talwar et al., 2019).

The findings from these previous studies suggests that people who disseminate misinformation on social media, frequently share similar attributes and are not a unique appearance (Guess et al., 2019; Pennycook & Rand, 2019). Basic individual attributes such as gender, education, and age have been connected to the spread of online misinformation. Previous research shows that older social media users tend to be more prone to engage with possibly misleading content than younger individuals (Grinberg et al., 2019; Guess et al., 2019; Osmundsen et al., 2020). On the other hand, Chadwick and Vaccari (2019) observed that younger British citizens would rather disseminate misinformation. Hence, findings on the importance of age are still inconclusive. Furthermore, studies have proven that men are more likely to engage with misinformation than women (Chadwick & Vaccari, 2019; Grinberg et al., 2019). Other findings suggest that next to gender and age, education plays a crucial role in the spread of online misinformation. Higher educated people are less likely to spread false and misleading information (Pop, 2019). Another defining factor, which is shared by individuals who spread misinformation, is heavy social media use. Research has found a significant correlation between the scope of social media use (e.g. in terms of sharing and liking news) and the prospect of disseminating misinformation (Buchanan, 2020b). Furthermore, high trust in news on social media can overall lead to more engagement with social media content (Sterrett et al., 2019). Aside from leading people to be more engaging with content, findings also suggest that online trust makes people less cautious in their sharing behavior. For example, Talwar et al. (2019) found that the people who trusted news on social

media more were also more prone to share inaccurate news stories on WhatsApp and less likely to authenticate the news story.

In the past years, scholars begun to study the impact of personality traits on disseminating online misinformation (Buchanan, 2020b; Buchanan & Benson, 2019). To answer the question of *who* shares misinformation, it is important to understand what role dark or callous personality traits play. Notably, the Dark Triad of personality traits has been identified as a characteristic that explains why people express malevolent behavior on social media (Buckels et al., 2014; Craker & March, 2016; Lopes & Yu, 2017). The Dark Triad consists of three elements: psychopathy, narcissism, and Machiavellianism. Psychopathy describes individuals with a lack of remorse, empathy, and anxiety, as well as thrill-seeking behavior. The main components of psychopathy are deficits in self-control and affect. Narcissism is described by feelings of grandiosity, superiority, and dominance. Machiavellianism entails the tendency to manipulate other people in a strategic, calculating way. The impulsivity linked to psychopathy is central in differentiating it from Machiavellianism. Machiavellians are concerned about their reputation and tend to plan ahead, unlike psychopaths (Jones & Paulhus, 2014; Paulhus & Williams, 2002).

Another strand of misinformation research has highlighted the importance of party affiliation and political orientation in connection with spreading online misinformation (e.g. Bennett & Livingston, 2018; Chadwick & Vaccari, 2019; Grinberg et al., 2019; Guess et al., 2019; Van Bavel et al., 2021). For instance, individuals are more likely to disseminate false information on social media if it is consistent with the positions of their party (Osmundsen et al., 2020). This process is also a likely explanation for the detected party divides: Literature focusing on the United States, postulates that a significant asymmetry exists between Democrats and Republicans when it comes to the dissemination of misinformation or fact checks. Overall, conservative individuals were more inclined to spread inaccurate information and were less likely to engage with fact checking claims (Guess et al., 2019; Osmundsen et al., 2020; Shin & Thorson, 2017). This is in line with research in the UK, where conservatives were also found to be more prone to spread false and misleading messages, whereas left-leaning citizens rather corrected other users (Chadwick & Vaccari, 2019).

Following the findings that social media users are more willing to share false and misleading content if it is consistent with their political orientation, another central predicting variable for spreading misinformation is attitudinal congruence (e.g. Buchanan, 2020; Hameleers et al., 2020). The concept of attitudinal congruence dates back to the beginning of cognitive consistency theories and is related to the confirmation bias (Festinger, 1957). Individuals consciously or unconsciously tend to consume information that is consistent with their existing attitudes and avoid uncongenial information to

decrease cognitive dissonance (Hameleers, 2019; Iyengar & Hahn, 2009). Disseminating attitudinally consistent content can mitigate "states of negative dissonance incurred by the existence of contradictory information, particularly as it pertains to one's in-group" (Hopp et al., 2020, p. 6). Essentially, this belief-based information selection on social media platforms leads to people primarily interacting with attitude-congruent social media content. Even if individuals do not know whether the information is true, they might uncritically decide to spread it to endorse the views of their in-group (Clemm von Hohenberg, 2019; Hameleers, 2019). In contrast, the willingness to support (misinforming) social media content declines when individuals expect incongruity within their network (Nekmat & Ismail, 2019). This suggests that social media users decide what to spread and support based on their understanding of the issue at hand and on preexisting beliefs.

The interplay between these different characteristics continues to be somewhat unclear. So the question remains which individual factors are the most appropriate to study the individuals who spread online misinformation. In addition, we do not know which characteristics turn out to be the most prevalent. In chapter 1, we delve deeper into how certain characteristics (i.e. personality traits, attitudinal congruence, issue salience) can explain who is willing to disseminate misinformation on social media. In chapter 4, we test the importance of individual factors in the context of countering online misinformation, since they are equally important when examining individuals who are not willing to further spread false and misleading content.

Why?

The answers to the question of who is the most likely to disseminate online misinformation, naturally lead to another essential part of this dissertation: *The Why.* What are the reasons why individuals spread online misinformation?

As low trust in news media has manifested itself over the past years in many different countries (Newman et al., 2021), ample research has linked (dis-) trust in institutions to the willingness to share online misinformation (Ognyanova et al., 2020; Stubenvoll et al., 2021). In contrast to the declining trust in traditional media, research suggests that individuals display high levels of trust in social media as an information source. Users are more inclined to trust and share content posted by other users and perceive the information to be more accurate (Anspach & Carlson, 2020). Not only does a trusted source matter, but also the indication of how many other users have already shared the content. When individuals notice that a large number of users have already engaged with the misinforming post, they are too more willing to share it further (Epstein et al., 2022). Furthermore, once social media users trust manipulated information they will not try to verify its claims and in consequence spread it further with large audiences, without the intention of actually doing any harm (Kalsnes, 2018). A frequently

considered reason why individuals would share misinforming content in the field of psychology is the implied truth effect (Pennycook et al., 2020; Pennycook & Rand, 2021). Social media platforms and media outlets are nowadays adding labels or fact checks to disputed content to warn users about misinformation. The implied truth effect suggests that people disseminate false and misleading content because they believe that if such warnings are missing, the information must be accurate (Pennycook et al., 2020).

As illustrated, different scholars and different disciplines use different approaches to study why individuals share misinformation. We argue that the classic Uses and Gratification Theory (Katz et al., 1974) is a fruitful approach to explain why people *willingly* spread online misinformation. This fundamental work greatly contributed to a transformation in thinking about media effects from "what the media do to people" to "what the people do with media". This approach proved to be a beneficial method to investigate the motivations behind the sharing of false and misleading information (e.g. Chen et al., 2015). We argue that at least three overarching reasons or motivations connected to this approach capture the rationales why people would disseminate online misinformation. First and foremost, there is *conviction*. The need or gratification of being able to signal agreement with what you consume online and consequently sharing your beliefs and views with others (Leiner et al., 2018). Relying on the literature on attitudinal congruent misinformation dissemination, individuals actively or unconsciously tend to choose content that is in line with their existing beliefs and like that avoid incongruent information to limit cognitive dissonance (Hameleers, 2019; Hopp et al., 2020). Hence, correctness seems to be much less important and this can cause the deliberate dissemination of misinformation because for people it matters much more that the content is attitudinally consistent with them, than it being correct (Pennycook et al., 2021). Thus, personal relevance and message support are considered a central motive for why individuals share online misinformation.

The second gratification social media users can get out of misinformation sharing is *entertainment*. Engaging with misinforming content might be seen as a 'fun' thing to do online. Social media platforms have been characterized as a "hedonic information system" (Islam et al., 2020, p. 4), where pleasure is one of the main reasons for people to use social media. Based on this, we suspect that perceived entertainment is an important predictor for why individuals would spread false or misleading information. Individuals do not necessarily perceive misinformation as harmful, but they can also experience this type of content as entertaining or funny (Daniel & Harper, 2020; van Prooijen et al., 2021). Findings further suggests that the experienced entertainment increases the likelihood to share online misinformation (Islam et al., 2020; Metzger et al., 2021). Research on conspiracy theories, which are linked to the concept of misinformation, finds that the provoked entertainment value through conspiracy theories promotes the belief in them (van Prooijen et al., 2021). People experience false

and misleading content as fascinating and exciting. This can trigger positive and intense emotions, which can explain the appeal of and the belief in misinformation (van Prooijen et al., 2021). Hence, people who look for fun and entertainment on social media might not be concerned about the authenticity and accuracy of a post and therefore spread unverified or simply false information (Islam et al., 2020). Against this backdrop, we assume that social media users possibly share online misinformation for fun, even if individuals do not necessarily support the content.

As a final motivation, we identified *reaction provocation* on social media as a third reason for the willful spread of online misinformation. As already touched upon, social media use is considered to be highly gratifying (van Koningsbruggen et al., 2017). For instance, people publish content on social media to trigger conversations and to find out about the different viewpoints of others (Zivnuska et al., 2019). To get reactions and attention from other users, liking, sharing, and commenting on social media content can be seen as one way to entertain those needs. On a content level, findings indicate that information that amplified COVID-19 related misinformation has been proven to get more engagement on social media and to be more viral (Papakyriakopoulos et al., 2020). Similar findings suggest that false information such as conspiracy theories becomes more viral than scientific information and lasts longer on platforms (Zhang et al., 2021). All in all, research suggests that spreading this type of harmful content is appealing for social media users who want to engage with others. We assume that a third motivation for why misinformation gains such high numbers of interactions on social media is that people want to provoke reactions and obtain attention within their social network.

The here listed reasons and motivations are by no means exhaustive. There is a myriad of other possibilities why individuals would share misinformation on social media, for example if individuals feel politically marginalized (Petersen et al., 2018). However, the motivations specified in this dissertation display a wide variety of microlevel and theory-based reasons which have been applied throughout the chapters and seem to be a valid selection. Chapter 2 of this dissertation examines therefore three concrete motivations (conviction, entertainment, reaction provocation) to provide deeper insights of why individuals would disseminate online misinformation.

Efforts to Counter Misinformation

The final contributions of this dissertation lie in the empirical study of individuals who actively counter online misinformation and in the investigation of the manner how social media users reply to real political misinformation. After listing the far-reaching effects misinformation can have on people, it appears that online (political) misinformation is tricky to get rid of in the minds of people. It is ‘sticky’ and persistent (Lewandowsky & van der Linden, 2021). However, we can be optimistic, there are multiple efforts to counter and combat online misinformation. For instance, on a policy level, holding

social media platforms accountable for hosting harmful content by implementing specific laws and fining the platforms if they violate them (Alemanno, 2018). Furthermore, governments and supranational institutions such as the WHO or the EU attempt to control the problem with, for example, international task forces. The EU, for instance, has a project since 2015 called EUvsDisinfo where they identify and expose disinformation in multiple languages. It started out focusing on the spread of pro-Kremlin disinformation, but since 2020 the project expanded its scope further to disinformation around the coronavirus. On a platform level – one of the most prominent interventions – flagging or attaching warnings to possibly false and misleading information was introduced as a remedy against misinformation (Pennycook et al., 2020). One strand of research proposes that if people are warned about the inaccuracy of messages, they would be less likely to believe it and misinformation is consequently being corrected (e.g. Chan et al., 2017; Lewandowsky et al., 2012). Similarly, three debunking strategies have been identified by previous research: (1) alarming users about misinforming content at the moment of the first exposure; (2) repeating the warning, and (3) offer compelling corrections which are coherent and accessible stories (Lewandowsky et al., 2012). Nonetheless, several scholars postulated that debunking might even enhance the belief in misinformation (e.g. Berinsky, 2017; Nyhan & Reifler, 2010). Individuals might grow to be resistant to these interventions, when they go against their (political) ideology (Flynn et al., 2017; Garrett & Weeks, 2013). Thus, findings are still inconclusive regarding debunking measures connected to online misinformation.

Since this dissertation focuses on the individual level, it is of high importance to explore what actions can be taken by the everyday social media user to combat online misinformation. On the microlevel, correction on social media platforms can happen through users commenting on false and misleading messages to call attention to the possibility that this content is inaccurate (Bode & Vraga, 2018; Tandoc et al., 2020). Vraga and Bode (2017) observed that social media users would correct or adapt their attitudes after seeing other users being corrected. This is especially successful if other users present believable corrective sources (Vraga & Bode, 2018). Moreover, have three reasons, which affect the motives why individuals would debunk misinformation, been identified by Tandoc and colleagues (2020): First, if people have a personal connection or relationship with the user who posted the false content. Second, if the topic of the misinforming message is personally relevant, and third, the user's personal efficacy, which refers to feeling successful in correcting misinformation.

Drawing on rumor research, which is also related to misinformation research, through interviews, Arif et al. (2017, p. 160) identified multiple rumor correcting methods: "correcting oneself, correcting the information space, and correcting another person (or organization)". Correcting oneself can be done through posting a correction of a false or misleading post one has published or shared beforehand. To

correct the information space, users indicated they wanted to keep the information on the platform as correct as possible. Thus, an individual mistake such as posting misinformation was not the main concern of the participants, but rather the greater good of the information quality on the platform itself. Correcting another person by directly tagging them was the final strategy to fight online rumors (Arif et al., 2017). Confirming these findings through interviews with young adults, Borah et al. (2021) found that directly confronting users who spread misleading content was an effective strategy by the respondents to fight online misinformation.

Even though confronting someone seems to be an often-quoted counteraction by social media users, a number of findings indicate that only a limited number of individuals corrects others. In the UK, only 8.5% of individuals confronted others for sharing false information (Chadwick & Vaccari, 2019). It appears that people can also take on a non-existent or passive role in combatting online misinformation by, for example, simply ignoring corrections or news in general (Borah et al., 2021; Tandoc et al., 2020; Young, 2021). Interviews illustrated that young adults ignore misinformation out of the feeling of getting frustrated and even angry (Borah et al., 2021). Tandoc et al. (2020) found in their study that almost three fourths (73%) of their participants would ignore inaccurate messages and not counter it by commenting or directly addressing the person who published misinforming content. One reason for ignoring possibly misinforming messages is the perceived accuracy of the content. It is difficult to assess the correctness of misinformation (Pennycook et al., 2020). Hence, people might be hesitant to interact with certain suspicious social media content and just avoid it. Furthermore, we should not underestimate the news literacy and social media literacy of people. The majority of social media users sees right through misinformation and does not further disseminate it (Guess et al., 2019). Chapter 3 and chapter 4 investigate closely how individuals counter real online misinformation and experimentally examine who and why someone goes against misinformation.

Case & Methodological Approach

To answer the central questions posed in this dissertation, namely who, why and how people engage with or counter online misinformation, I draw on microlevel data acquired among social media users. The methodological approach is innovative in three ways: First, it measures the behavior of individuals in six Western democracies, which enables me to observe effects across different national contexts and broadens the scope of the majority of the research, which focuses heavily on the U.S. Second, in our survey, we exposed the social media users to three highly polarized, highly realistic misleading social media posts. This helps to better understand the diversity and versatility of online misinformation and allows measuring the impact of issue importance in connection with misinformation sharing. Third, I focus on individual user comments on, by fact checking organizations

as misinformation identified, social media content. I will concisely discuss the six selected cases and the two applied methods. More detailed information on the specific methods and data used can be found in the corresponding chapters.

Case: Six Western Democracies [3]

This dissertation takes six countries into account: Belgium, France, Germany, Switzerland, the UK, and the U.S. The countries were selected based on certain characteristics of their information environments. They had to systematically vary in the strength of public broadcasting, level of media trust, and the level of media fragmentation. Germany, Switzerland and the UK have a strong public broadcasting system, measured by the public revenue and market share of the public broadcaster. The public broadcaster in France and Belgium is slightly less strong. The U.S. has by far the weakest public broadcasting system (Brüggemann et al., 2014). The trust in media also heavily varies across countries. In Germany, Belgium and Switzerland the trust in news is over 45%, whereas for the UK, the U.S. and France it is far below that level. The U.S. has by a great deal the lowest score for media trust (26%), as it stands in 2022 (Newman et al., 2022). Turning to national media fragmentation, we observe that in the UK most of the people consume news online, followed by TV and social media as a news source (Newman et al., 2022). We see the same pattern in Belgium, Germany, Switzerland, France, and the U.S. However, in contrast to Belgium and the UK, in France and Germany TV as a news source is almost as much used as online sources. In every country, print media score the lowest as a news source, especially in the U.S. In Switzerland, however, print is used roughly as frequently as social media (Newman et al., 2022). Research suggests that there is a strong link between the strength of the domestic public broadcaster, high media trust and the resilience to false and misleading information (Humprecht et al., 2020).

To measure whether certain opportunity structures facilitate the spread of misinformation, differences in political and party systems were also of high relevance. Belgium, Germany and Switzerland are consensus democracies in contrast to the UK and the U.S., which are majoritarian democracies (Lijphart, 2012). France can be seen as a mixed democracy. Belgium and Switzerland have the most parties in government (n = 4) (Europa Publications, 2018). France and Germany have each two parties in the government, whereas the UK and the U.S. only have one party. The different forms of democracies raised the question if different models lead to more or less resilience towards

[3] The case selection is based on the research plan of the Disinformation Project, led by Edda Humprecht, Peter Van Aelst, and Frank Esser. I hold one of two PhD positions of this project, and this dissertation is therefore closely linked to this collaboration. In this dissertation, I do not focus on potentially relevant country differences (but see Humprecht et al. 2021). The different countries under study rather function as strong robustness tests and allow for a broader generalization.

misinformation. For instance, the U.S. with its two party system and majoritarian democracy has been proven quite vulnerable during the 2016 presidential election (Allcott & Gentzkow, 2017). In the UK, misinformation was widely spread during the Brexit referendum with the aim to leave the European Union (Nielsen & Graves, 2017). Contrary to that, leading up to the general elections in Germany in 2017, misinformation did not have a big impact on the campaigns (Sängerlaub, 2017). The suspected reason behind this pattern, is that in consensus democracies the level of conflict in political debates is lower, and that partisan networks are less isolated, leading to hamper the spread of misinformation. Moreover, misinformation is spread more easily and willingly in polarized societies, such as the U.S. (Humprecht et al., 2020; Tucker et al., 2018). Besides, in smaller countries such as Belgium and Switzerland, the scope of websites providing false and misleading messages is minimal. Possibly as a result of missing economic incentives in these small audience markets (Humprecht et al., 2020). For example, the U.S. is a much more attractive country for the producers of misinformation – not only because of its level of polarization but also because of its enormous advertising market (Humprecht et al., 2020).

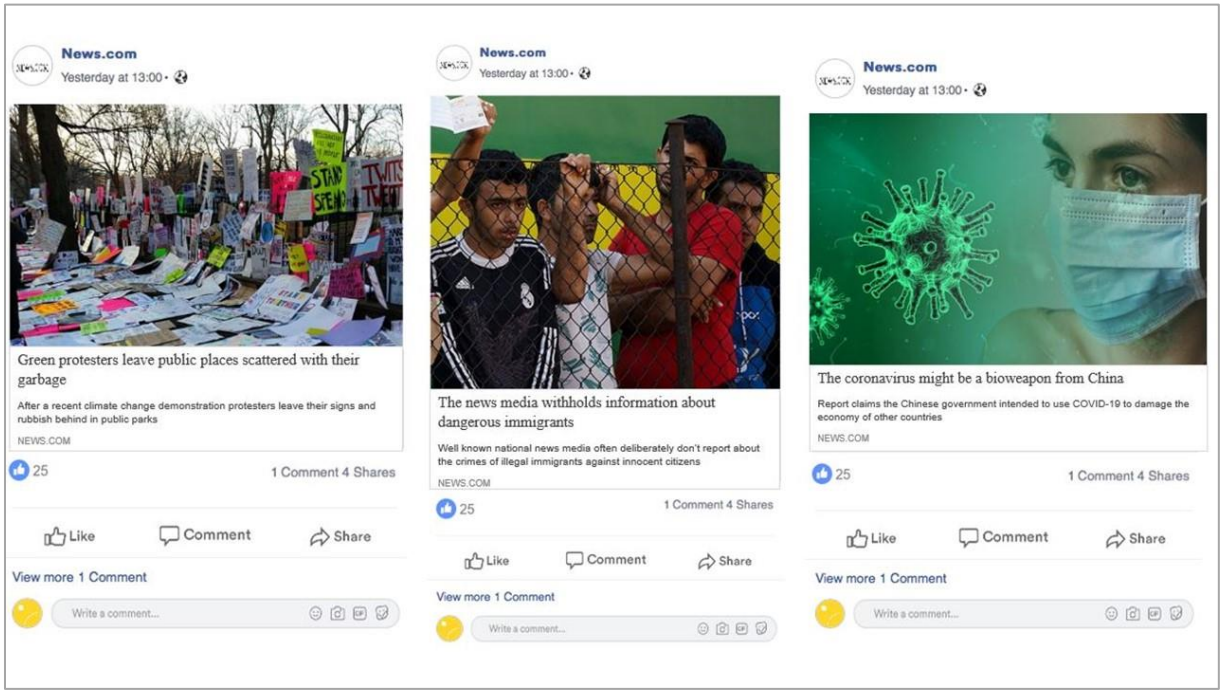
Methods

Examining the microlevel connected to the dissemination of online misinformation, I rely on two methods: a large-scale survey with social media users and a manual content analysis of user comments. To answer three out of the four overarching research questions, I draw back on the survey data. The survey was conducted during four weeks in April and May 2020 in all six cases under study, and the polling company Respondi recruited representative samples of social media users based on country-specific census data (see Appendix 1.1).

To investigate *who and why* engages with online misinformation and *who and why* would someone counter misinforming social media content, we exposed our respondents to three fabricated false and misleading posts. The created vignettes resembled news articles posted on Facebook by a fictitious news source (news.com). Each post contained a different issue: immigration, climate change protests, and COVID-19 (see Figure 1). We argue that these three topics are appropriate for the study of misinformation, because they had been linked to inaccurate and misleading information circulating widely on social media platforms in the months and weeks prior to the survey (Mimikama, 2019; Petersen et al., 2018) and are highly polarized. Furthermore, this also enhances the dissertation's external validity. The statement in the post regarding immigration read as follows: "The news media withholds information about dangerous immigrants – Well known national news media often deliberately don't report about crimes of illegal immigrants against innocent citizens". The COVID-19 post included the following claim: "The coronavirus might be a bioweapon from China – Report claims

the Chinese government intended to use COVID-19 to damage the economy of other countries". The post regarding climate change had the following message: "Green protesters leave public places scattered with their garbage – After a recent climate change demonstration, protesters leave their signs and rubbish behind in public parks". Not only can these three manipulated posts be linked to misinformation, but also to specific conspiracy theories. The post on immigration is in line with the Great Replacement conspiracy theory where white the European identity and civilization is in danger because (Muslim) immigrants are plotting against the West, to eventually taking it over in a hostile manner (Bergmann, 2021; Cosentino, 2020; Obaidi et al., 2021). The claim of the COVID-19 post suggests that the Chinese government is to be blamed for the planned outbreak of the virus. Blame attribution and secretive plotting are common features of conspiracy theories that can be found on social media (Uscinski et al., 2016). The central questions, that we asked the respondents, inquire about how individuals would normally react to the posts. We asked, for instance, whether they agree with the message of the posts and what the possible reasons are for why they would react to it or why they would not interact with a post.

Figure 1. Online misinformation vignettes



Relying on self-reported survey data consistently bears two main limitations, which should be addressed: Social desirability and post hoc rationalization. The issue of social desirability is especially important, considering that I am dealing with a very delicate topic. There is never certainty that individuals would answer according to their true beliefs when investigating their engagement with online misinformation. However, we attempted to contain the problem of individuals giving desirable answers by assuring that their answers will be anonymized and that their data will not be shared with

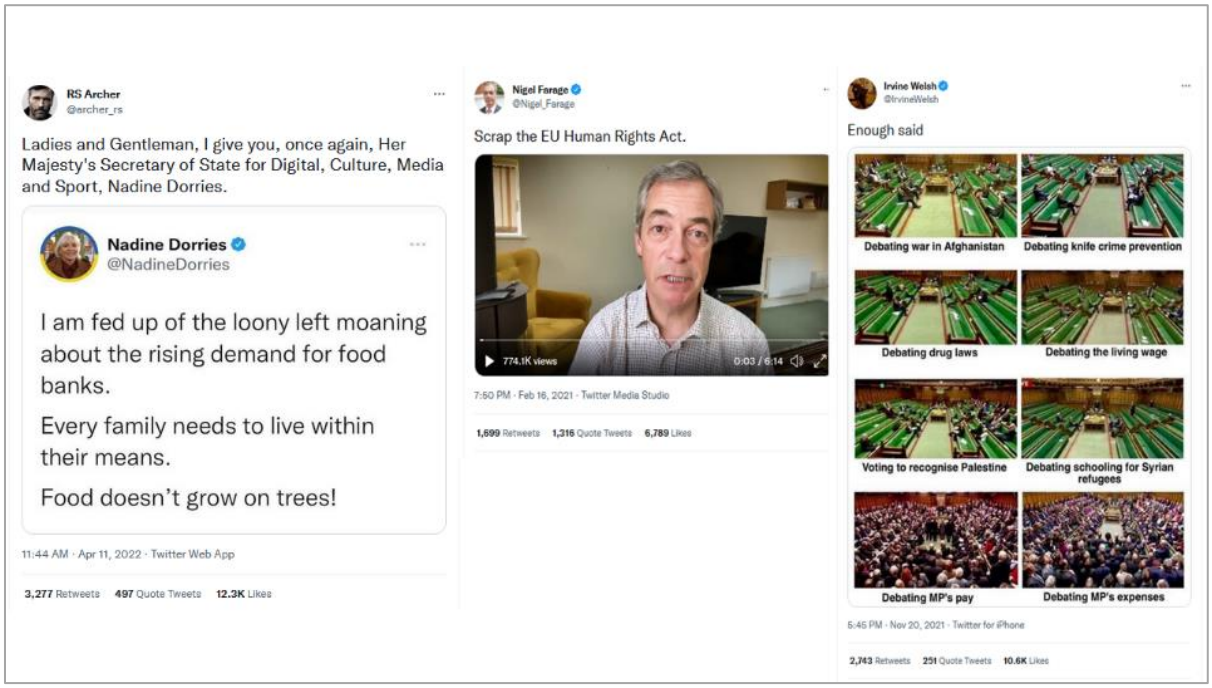
third-parties. Additionally, did the respondents fill in the survey online and no researcher was present to evoke a possible bias. Knowing how many individuals answered that they would engage with at least one of the three presented stimuli, I am positive that social desirability played only a marginal role.

Furthermore, post hoc rationalization can be seen as a second major problem with self-reported data. Giving desirable answers is a rather conscious act, while rationalization can happen unconsciously. Correctly remembering one's decisions is tough, and we often trust our intuition when making the decisions. Being challenged to elaborate on these choices afterward, leads us to rationalize why we made a certain choice, and that reason might be completely different from the processes that actually led us to this decision. In a nutshell, we first react with our gut instinct and think about the reasons for this reaction in a later step. We do so because decision processes are too complex for us to understand, and we are also unaware why we sometimes behave in a certain way (Summers, 2017). In this dissertation, this also applies to the question of why individuals would share misinformation. It is easy to imagine that individuals give rather a socially accepted answer why they consciously shared misleading information, than admitting they want to mess with other users. However, we argue that post hoc rationalization was rather unlikely to singularly drive the findings of the survey. The fact that we uncover social media users who are willing to promote themselves as 'trolls', by, for instance, indicating they would spread the misleading posts just for fun, or to gain attention from their network, indicates that individuals are not only reasonable beings. Nevertheless, we acknowledge that the spreading of inaccurate information is not necessarily a conscious process, but can also be considered as a rather spontaneous or automatic one (Buchanan, 2020b) that does not require much cognitive effort (Pennycook & Rand, 2019).

The findings from the survey allowed me to understand who and why individuals disseminate misinformation, but not how they react to it in a real-life setting. To answer the question of *how* precisely individuals react to political online misinformation, I conducted a more in-depth analysis in one country, the UK. More concretely, a manual content analysis of the comments belonging to three misinforming posts involving UK politicians (see chapter 5). The selection was based on certain criteria. First, the tweets needed to clearly contain misinformation according to well-established fact checking organizations. Second, the posts should have gained a lot of traction when published (more than 5000 likes) and finally, they had to be diverse in the type of misinformation and political actor involved. The tweets selected varied therefore in style, affected politician and user engagement. The first tweet involves a bluntly fake tweet of a former Secretary of State posted as a screenshot by a fake account. The fact checking organization Full Fact found out that the tweet is fake, meaning the politician never published it and someone else had to create it. The second tweet was posted directly by the official account of a highly polarized political actor in the UK. In his tweet, the politician is without a doubt

disseminating misleading information, which has been identified by The Ferret, a Scottish fact checking organization. The third tweet shows a series of pictures of the House of Commons during multiple debates and accuses the members of parliament to only be present when they are debating MPs' payment and expenses. It was posted in a meme-like format on the official account of a Scottish writer. The Ferret's verdict is that this tweet is mostly false. Certain pictures are linked to the correct debate but the two photos where MPs are apparently debating their pay, are false. This tweet puts correct information out of context. Altogether, I examine three types of misinformation: A fake social media post, misleading information, and accurate information put out of context. Figure 2 displays the tweets studied in this dissertation.

Figure 2. Tweets containing political misinformation

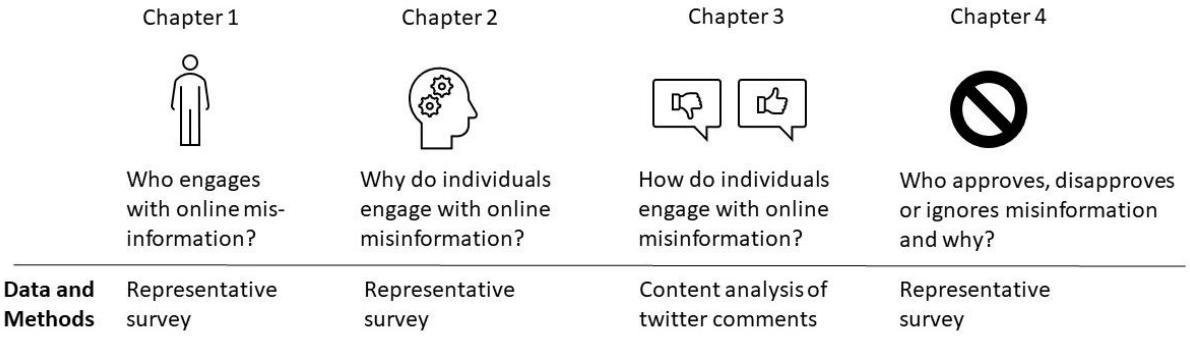


Just as with self-reported survey data, manual content analyses have their limits too. For instance, their scope and considered time period are somewhat limited. Further, developing and testing a codebook and instructing coders is extremely time-consuming. Hence manual content analyses are often limited in the number of items that are analyzed.

Structure of the Dissertation

In sum: the main goal of this dissertation is to strengthen and complement our understanding of the interplay between individuals and online misinformation. This complex relationship is studied in detail through four different chapters (see Figure 3 for an overview).

Figure 3. Overview of the dissertation



The **first chapter** of this dissertation focuses on the individual characteristics of individuals who engage with online misinformation and is titled: '*Identifying the Drivers Behind the Dissemination of Online Misinformation. A Study on Political Attitudes and Individual Characteristics in the Context of Engaging with Misinformation on Social Media'*. This is an essential first step in examining the 'ins and outs' of the relationship between individuals and misinformation on social media. My co-authors and I ask which individuals interact with this kind of information and what role attitudinal congruence plays in this context. To uncover shared individual factors by people who would willingly disseminate false and misleading content, we conducted a large-scale survey with social media users in six countries (BE, CH, DE, FR, UK, U.S.). We furthermore test three noncountry-specific topics (immigration, climate change protests, and COVID-19) to make our results more robust and to be able to control for issue salience. Our results uncover that besides issue attitudes and issue salience, personality traits, political orientation, and heavy social media use increase the willingness to disseminate misinformation on social media. We conclude that there exist specific user groups that are particularly susceptible to misinformation and possibly caught in social media 'fringe bubbles'.

Chapter 2 with the title *'To Convince, to Provoke or to Entertain? A Study on Individual Motivations behind Engaging with Conspiracy Theories Online'* sheds light on the reasons why individuals engage with online misinformation. It does so by focusing on the conspiratory aspects of misinformation and investigates the role of specific beliefs, but also individual factors such as personality traits in connection with the motivations behind sharing conspiracy theories online. Just as chapter 1, this chapter relies on the data we collected during an extensive survey in multiple countries. My co-authors and I investigate three possible motivations to engage with online misinformation based on the Uses and Gratifications Approach (conviction, entertainment, and reaction provocation). Our findings demonstrate that across the three different issues, individuals who indicated they would engage with conspiracy theories, do it mainly because they are convinced by the message. Political orientation and issue attitudes prove to be linked to people's engagement with online conspiracy content out of

conviction, while dark personality traits such as narcissism and psychopathy are compelling predictors for why individuals would disseminate conspiracy theories out of entertainment or to provoke reactions within their networks.

Because this dissertation studies the connection between individuals and misinformation, it is essential to know how people react to real online misinformation. In **chapter 3** I therefore closely examine the comment section linked to three different types of political online misinformation. The chapter is entitled '*Come on people, this is clearly fake. How Social Media Users React to Political Online Misinformation by Commenting'.* Little to no research has explored how people react to this content by commenting on it. To fill this void and to identify first patterns of reactions, I use a manual content analysis of user comments on three, as misinformation identified, tweets involving UK politicians. The main findings show that the majority of the user comments disagreed with or countered the misinforming tweets, and only a fraction expressed approval. Thus, the majority of Twitter users disapproves of the false or misleading content or even calls out the original poster of the claims. Furthermore, I find that a vast amount of the comments contained uncivil language.

Chapter 4 titled *'Who Approves, Disapproves or Ignores Online Misinformation? A Study on Individual Characteristics in Connection with Supporting and Renouncing Online Misinformation'* builds on the findings of chapter 1 and chapter 2 but looks the other way. As important as it is to know why and who disseminates online misinformation, it is equally important to elucidate who and why individuals counter misinformation. This chapter compares three groups of individuals based on their engagement with misinformation (no interaction, signaling agreement, signaling disagreement) and three different topics (climate change protest, immigration, COVID-19). We find striking similarities between individuals who agree and disagree with misinforming content. They are mostly male, lower educated and more interested in politics. However, social media users who generally support the misinforming messages showed stronger manifestations of psychopathy, narcissism, and Machiavellianism, indicating that these individuals have a darker personality. We deduce from our findings that there is not a clear profile of individuals who would renounce misinformation, but rather indications for a dynamic and adaptive social media user.

In the final chapter of this dissertation, the findings of the four empirical chapters are synthesized, the overall contributions are emphasized, and the limitations are acknowledged. To conclude, suggestions are made for future research in the field of online misinformation.

CHAPTER TWO

Identifying the Drivers Behind the Dissemination of Online Misinformation

A Study on Political Attitudes and Individual Characteristics in the Context of Engaging with Misinformation on Social Media

The increasing dissemination of online misinformation in recent years has raised the question which individuals interact with this kind of information and what role attitudinal congruence plays in this context. To answer these questions, we conduct surveys in six countries (BE, CH, DE, FR, UK, U.S.) and investigate the drivers of the dissemination of misinformation on three noncountry specific topics (immigration, climate change, and COVID-19). Our results show that besides issue attitudes and issue salience, political orientation, personality traits, and heavy social media use increase the willingness to disseminate misinformation online. We conclude that future research should not only consider individual's beliefs but also focus on specific user groups that are particularly susceptible to misinformation and possibly caught in social media 'fringe bubbles'.

Keywords: misinformation, dissemination, social media, political attitudes, personality traits

Reference: Morosoli, S., Van Aelst, P., Humprecht, E., Staender, A., & Esser, F. (2022). Identifying the Drivers Behind the Dissemination of Online Misinformation: A Study on Political Attitudes and Individual Characteristics in the Context of Engaging With Misinformation on Social Media. *American Behavioral Scientist*, *0*(0). https://doi.org/10.1177/00027642221118300

Introduction

Because a well-informed citizenry is considered a vital foundation for a healthy democracy, the dissemination of false and misleading information poses a major threat to it. The deliberate fabrication of false information to mislead individuals, generally labelled as disinformation, is nothing new, but today's high-choice media environment provides a fertile ground for its dissemination (Weeks & Gil de Zúñiga, 2021). After the 2016 U.S. presidential elections in particular, scholars began examining origins, causes and consequences of false and misleading messages. Ample research indicates that this false content is often deliberately and strategically created to reach certain segments of the public and influence their beliefs and political behavior (Bennett & Livingston, 2018; Wardle & Derakhshan, 2018). Moreover, ordinary citizens contribute to its dissemination and are consequently co-responsible for misinforming parts of the public (Talwar et al., 2019).

A logical starting point is that people share false information because they belief it is true. However, several authors argue that belief in false information is a sufficient, but not a necessary condition for the spread of misinformation and that misleading messages are often shared without the person necessarily believing it (Rossini et al., 2020; Van Bavel et al., 2021). According to Pennycook et al. (2021), citizens disseminate misinformation for different reasons, among others, because they care less about the truth and more about having their opinion confirmed or because they do not focus on the veracity but on other aspects of the message. By now, scholars have identified ample factors that play a central role in the dissemination of online misinformation, but a shared comprehensive framework of what drives the spread of false information is still missing. In a recent step towards an integrative theoretical model, Van Bavel et al. (2021) identified multiple psychological risk factors, which underlie the dissemination of misinformation. First of all, they include several political factors such as partisan bias, polarization, and political ideology (see also Calvillo et al., 2020; Guess et al., 2019). For instance, partisan bias can facilitate the spread of misinformation because it confirms people's own opinions even though the information might be incorrect. Moreover, in polarized environments people might be willing to spread misinformation for partisan gain, to communicate their political identity, and to encourage hostility towards the partisan out-group. In addition to these partisan and ideological factors, Van Bavel et al (2021) argue that several deeper personality traits, emotions, and cognitive skills influence the dissemination of misinformation.

Building on this strand of research, this study aims to identify the individual characteristics that drive the dissemination of online misinformation. In line with Van Bavel et al. (2021) we look both at political attitudes and personality traits (see also Hameleers et al., 2020; Petersen et al., 2018). Additionally, we include insights from the literature on social media and misinformation. Because false information

is spread primarily through social media, it has been argued that people who use social media intensively are more likely to come into contact with false information and are thus more likely to spread it (Humprecht et al., 2020). To structure these different risk factors, we use a funnel that moves from stable to more fluent individual characteristics and attitudes (see Figure 1).

Our study relies on extensive survey data from six Western democracies (Belgium, France, Germany, Switzerland, the United Kingdom and the United States). We test our assumptions on three issues that are highly salient, often polarizing, and relevant in all countries under study: climate change, immigration, and COVID-19. This focus on specific issues allows us to more fine-grained test the role of attitudinal congruence with the message of a post, rather than relying on more general proxies such as partisanship. We focus on the characteristics of people who are willing to like, share or comment on social media posts that contain false or unproven information. We are interested to what extent people are willing to engage with misinformation and what drives these behaviors. Since the intention of the producers of such information is often not apparent to citizens, we prefer, in the context of this specific study, the term misinformation instead of disinformation, often used to refer to the deliberate spread of false information (Wardle & Derakhshan, 2018).

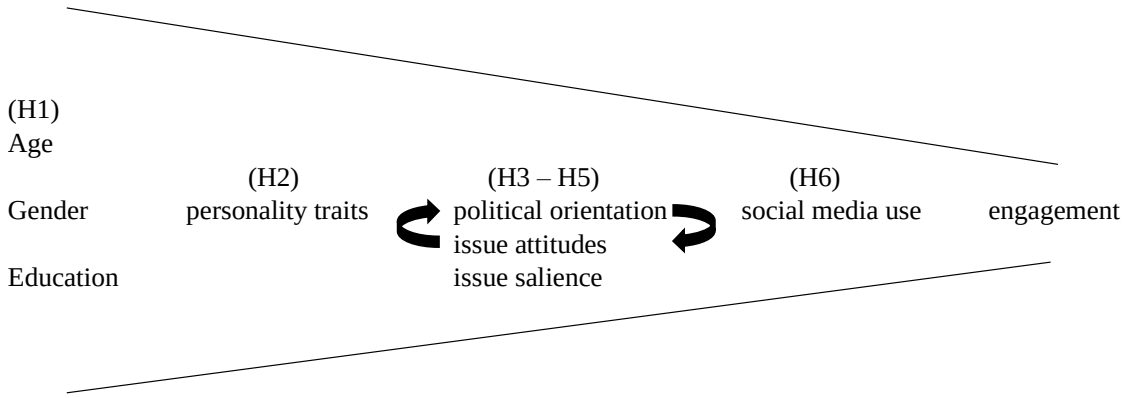
Since there has been a strong focus on the U.S. in the research on misinformation, we additionally included Western European democracies to test our assumptions across varying contexts and to increase the robustness and generalizability of our results (Esser & Vliegenthart, 2017). Subsequently, the strength of the following study lies in combining different individual factors and testing them on various issues across populations. Our results show both specific attitudes towards issues and general characteristics explain the individual dissemination of online misinformation, thereby contributing to a better understanding of the rapid and wide spread of false and misleading messages on social media.

The Dissemination of Online Misinformation

Social media play a crucial role in how individuals consume and share news with others. Spreading content on social media can, among other things, strengthen underrepresented opinions (Chadwick & Vaccari, 2019), but it becomes problematic when people share false or misleading messages because they can reach large audiences (Allcott & Gentzkow, 2017). One reason for liking or sharing specific social media posts is that citizens agree with the message and convince themselves that sharing the information is legitimate. Since the rise of social media, scholars have extended their interest from selective exposure to selective sharing of information. For instance, Johnson et al. (2020) showed that across various issues, people's existing political beliefs are even more important to explain the selective sharing of online information than to explain selective exposure. People are eager to share information that is in line with their views, even when that information is unproven or openly false

(Buchanan, 2020b; Guess et al., 2019). In addition, previous research has also shown that other overriding factors influence citizens' social media behavior regarding the dissemination of misinformation (Van Bavel et al., 2021). Overall, the current state of research suggests that people who spread deceptive or misleading content often share similar characteristics (Guess et al., 2019; Pennycook & Rand, 2019). This seems to imply that false information is shared online independent from a specific topic or message. To test whether this general expectation holds, we combine different factors in one comprehensive model, and at the same time, explore in a more detailed manner the role of specific issue attitudes in the spread of misinformation. More concretely, we examine the role of the individual-level factors like a funnel (see Figure 1). We move from fixed personal factors such as age, gender, and education to personality traits as well as political ideology, issue attitudes, issue salience, and social media use for the dissemination of misinformation. We argue that the factors closer to the narrow end of the funnel are less structurally determined and more directly related to our dependent variable (engagement with misinformation). Furthermore, we expect that the more factors we add, the more nuanced and comprehensive we can explain individuals' engagement with misinformation. Next, we elaborate on the different characteristics and formulate concrete hypotheses.

Figure 1. Model of the funnel-like order of the hypotheses



Sociodemographic Factors

Basic individual characteristics such as age, gender and education have been associated with the dissemination of misinformation. In previous studies older social media users have been found to be more likely to engage with possibly misleading content (Grinberg et al., 2019; Guess et al., 2019; Osmundsen et al., 2020). On the other hand, Chadwick and Vaccari (2019) found that rather younger British people (younger than 45) would share misinformation. Thus, research on the role of age in the

context of misinformation sharing is still inconclusive. Additionally, research has shown that men are more likely to engage with misinformation than women (Chadwick & Vaccari, 2019; Grinberg et al., 2019). This effect might be explained by the gender gap in general online behavior. According to recent research, women tend to be less likely to comment on political news online and men generally tend to share more news on social media (Chadwick & Vaccari, 2019; Van Duyn et al., 2021). Other findings suggest that next to age and gender, education plays a crucial role in the dissemination of misinformation on social media. Higher educated individuals are less prone to disseminating false information (Pop, 2019). Following the findings from misinformation research and research on general social media behavior, we assume that these individual factors have a significant impact on the engagement with misinformation across issues and various populations.

Hypothesis 1a: Men are more likely to disseminate misinformation on social media than women are.

Hypothesis 1b: Older individuals are more likely to disseminate misinformation on social media than younger individuals are.

Hypothesis 1c: Lower educated individuals are more likely to disseminate misinformation on social media than higher educated individuals are.

Personality Traits

Personality traits have shown to be an important predictor for studying general social media use (Chen, 2016), and very recently, scholars begun to study the influence of personality traits on propagating online misinformation (Buchanan, 2020b; Buchanan & Benson, 2019). To answer the question of what drives the dissemination of misinformation, it is important to understand what role aversive personality traits play in the dissemination of those messages. In particular, the Dark Triad of personality traits has been identified as a factor that explains why people express malicious online behavior (Buckels et al., 2014; Craker & March, 2016; Lopes & Yu, 2017). The Dark Triad consists of three components: psychopathy, narcissism, and Machiavellianism. Psychopathy describes personalities with a lack of empathy, remorse, and anxiety as well as high impulsivity and thrill-seeking behavior. It consists out of two main elements: deficits in self-control and affect. Narcissism is characterized by feelings of dominance, grandiosity, and superiority. The driver behind narcissistic behavior is ego-reinforcement, whereas psychopathy and Machiavellianism are driven by instrumental gain. Machiavellianism consists of the tendency to manipulate other people in a strategic, calculating way. The impulsivity attributed to psychopathy is central in distinguishing it from Machiavellianism. Machiavellians are, in contrast to psychopaths, concerned about their reputation and tend to plan ahead (Jones & Paulhus, 2014; Paulhus & Williams, 2002). Because the three constructs are slightly intercorrelated, they share some key aspects, such as aggressiveness, emotional coldness, self-

promotion, and general antisocial behavior (Jonason & Webster, 2010; Jones & Paulhus, 2014; Paulhus & Williams, 2002; Rauthmann & Kolar, 2013).

We assume that for individuals who exhibit a dark personality, the veracity of online information does not play a central role but the impact of possibly false and misleading information does. For example, these individuals intentionally disseminate this kind of information to attract attention or alienate others. Personality traits can therefore contribute to the understanding of why people would generally engage with misinformation.

Hypothesis 2a: Individuals with narcissistic personality traits are more likely to disseminate misinformation than individuals without narcissistic traits are.

Hypothesis 2b: Individuals with psychopathic personality traits are more likely to disseminate misinformation than individuals without psychopathic traits are.

Hypothesis 2c: Individuals with Machiavellian personality traits are more likely to disseminate misinformation than individuals without Machiavellian traits are.

Political Orientation

Previous studies have highlighted the role of party affiliation and political orientation in connection with misinformation sharing (e.g. Bennett & Livingston, 2018; Chadwick & Vaccari, 2019; Grinberg et al., 2019; Guess et al., 2019; Van Bavel et al., 2021). Osmundsen et al. (2020) found that individuals rather share false information on social media if it is in line with the positions of their party. This is also a possible explanation for the discovered party divides: Studies focusing on the U.S. indicate a significant asymmetry exists between Republicans and Democrats when it comes to the sharing of misinformation or fact checks. In general, conservatives were more likely to disseminate inaccurate information and were less likely to engage with fact checking messages (Guess et al., 2019; Osmundsen et al., 2020; Shin & Thorson, 2017). This is consistent with research in the United Kingdom, where conservatives were found to be more likely to spread misinformation, whereas left-leaning citizens rather corrected others (Chadwick & Vaccari, 2019).

These findings indicate that political orientation is an important predictor of online misinformation sharing. Previous research suggests that misinformation tactics are frequently connected to right-wing actors and that misinformation is not exclusively, but predominately right-leaning (Bennett & Livingston, 2018; Ognyanova et al., 2020). Although the fabricated social media posts used in this study do not all have a unidimensional right-wing political orientation (e.g., the one on the coronavirus and on climate protestors), it can be assumed that right-leaning individuals are more likely to support and disseminate this content.

Hypothesis 3: Individuals with a right-leaning political orientation are more likely to disseminate misinformation.

Attitudinal Congruence

Following the argument that social media users are more inclined to share misinformation if it is consistent with their political orientation, another central predictor is attitudinal congruence (e.g., Buchanan, 2020; Hameleers et al., 2020). The concept of attitudinal congruence dates back to the beginning of cognitive consistency theories and is related to confirmation bias (Festinger, 1957). Individuals actively or unconsciously tend to select content that is in line with their existing beliefs and avoid uncongenial information to reduce cognitive dissonance (Hameleers, 2019; Iyengar & Hahn, 2009). Sharing attitudinally congruent information can reduce "states of negative dissonance incurred by the existence of contradictory information, particularly as it pertains to one's in-group" (Hopp et al., 2020, p. 6). In general, this belief-based content selection on social media results in individuals mainly engaging with attitude-congruent social media posts. Even if people are not sure whether the information is true or partly true, they might uncritically decide to share it to support the views of their in-group or community (Clemm von Hohenberg, 2019; Hameleers, 2019). In contrast, the willingness to express support for a (misleading) social media post decreases when individuals assume incongruity within their network (Nekmat & Ismail, 2019). This finding indicates social media users decide what to disseminate and support based on their perceptions of the topic and on preexisting attitudes.

Against this backdrop, attitudinal congruence is of high importance when it comes to disseminating misinformation. We assume that people are more likely to disseminate false social media content when the message is in line with their preexisting attitudes:

Hypothesis 4: The stronger the attitudinal congruence regarding a specific issue, the higher the willingness to disseminate misinformation.

Issue Salience

Closely related to attitudinal congruence is the concept of issue salience or issue importance. Issue salience has been studied extensively as part of agenda-setting research, (e.g. McCombs & Shaw, 1972; Tan & Weaver, 2007; Wanta et al., 2004). In general, salient issues increase people's cognitive and emotional engagement while they process a message (Eberl et al., 2020). In the context of social media, the individually ascribed salience of an issue influences how users react toward related social media posts (Waddell, 2018).

Previous research has shown that a higher personal importance attributed to an issue leads to individuals having greater willingness to express their support and share corresponding information

on social media platforms (Nekmat & Ismail, 2019). The increased involvement triggered by the topic's importance thus influences user behavior on social media. We expect an effect of issue salience on the individual motivation to engage with false or misleading information on social media.

Hypothesis 5: The higher the perceived salience of an issue, the higher the willingness to disseminate misinformation.

Social Media Behavior

Because misinformation spreads easily through social media, it is crucial to take users' general social media behavior into account when investigating their propensity to disseminate false or misleading content. When it comes to individual social media use, previous research has postulated significant correlations between the scale of the social media use (e.g., in terms of liking and sharing news) and the likelihood of sharing false information (Buchanan, 2020). Because social media activity is seen as active rather than passive behavior, we expect that individuals who generally engage with social media content frequently are more inclined to like, share, or comment on online misinformation. Moreover, higher levels of trust in news on social media can lead to greater engagement with social media content in general (Sterrett et al., 2019). Aside from making people more active, research shows that online trust makes people less cautious in their sharing behavior. For instance, in their study on news sharing on WhatsApp, Talwar et al. (2019) found that those who trusted news and information on social media more were more likely to spread false news stories and less likely to authenticate the news story. Against this backdrop, we assume that active social media users and individuals who trust news on social media are more willing to spread misinformation.

Hypothesis 6a: Individuals who use social media more frequently and are more active on these platforms are more likely to disseminate misinformation than less active individuals are.

Hypothesis 6b: Individuals who trust news on social media more are more likely to disseminate misinformation than less trustful individuals are.

Methods

Design. To be able to make claims about individuals in different information environments, we conducted representative surveys in six Western democracies (Switzerland, Belgium, France, Germany, the United Kingdom and the United States.[4]) to investigate the dissemination of

[4] The data used in this study are connected to a large-scale research project. These countries were chosen to test different contextual factors that create opportunity structures for the dissemination of online misinformation. This study does not focus on these factors or on specific country differences, rather, we aimed to find generic factors that matter across (Western) countries.

misinformation related to three issues (immigration, climate change and COVID-19). We argue that these three topics are well suited for the study of misinformation, because they had been connected to false and misleading messages circulating widely on social media in the months and weeks prior to the survey (Mimikama, 2019; Petersen et al., 2018) and are highly polarized. These issue characteristics allowed us to examine the roles of political attitudes and issue salience in the context of the dissemination of such content. The polling company Respondi recruited representative samples of social media users (usage of at least once a month) in all six countries under study based on country-specific census data (see Appendix D[5]). A detailed description of the procedure and the vignettes used in this study can be found in Appendix F[6].

Sample. After removing careless respondents and other outliers based on response time and quality fail questions, we secured a sample of 7,006 respondents (quotas for age: *M* = 43.87, *SD* = 14.69; gender: female = 51.7%, male = 48.3%, education: lower = 27.7%, moderate = 40.1%, higher = 32.1%). We are interested in any type of engagement with misinformation and subsequently in any type of dissemination. However, we excluded respondents who indicated they would *comment* on the social media post to signal disagreement (climate change post, *n* = 735; immigration post, *n* = 486; COVID-19 post, *n* = 626) which resulted in a final sample size of *N* = 5,791. We see commenting to disagree with the post as a way of combatting misinformation (e.g., by warning others), whereas liking or sharing the post to signal disagreement is mostly well intended but can unwillingly contribute to the message's dissemination.

Dependent Variable. All participants were shown three fabricated social media posts, one for each issue. The created vignettes resembled news articles posted on Facebook by a fictional news outlet (news.com). The participants were told a cover story at the beginning of the survey stating that the aim of this study was to assess their opinions on different social media posts, political issues, and actors as well as their social media use. Before they were exposed to the fabricated posts, the participants were told to imagine that the following posts would appear on their newsfeed and to read them carefully.

Willingness to interact with the social media posts was measured with three items. On a 7-point Likert scale (1 = very unlikely, 7 = very likely), participants had to answer how they would usually react to the presented post, specifically whether they would "like the post," "share the post," or "leave a comment.". Because the three types of user engagement with the post were correlated strongly (Cronbach's α = .85), we built mean indices across all types of reactions (like, share, and comment) for

[5] The Online Appendix can be accessed here: http://www.disinformation-project.com/data
[6] The Online Appendix can be accessed here: http://www.disinformation-project.com/data

each topic (climate change: *M* = 2.43, *SD* = 1.67; immigration: *M* = 2.13, *SD* = 1.67; COVID-19: *M* = 2.28, *SD* = 1.74).

In our sample, between 10% and 37% of the participants in a specific country stated that they likely would engage with the presented post. Table 1 shows that, across countries, the post that contained the claim regarding climate protesters led to the highest willingness to engage, followed by the posts related to immigration and finally coronavirus (climate change: *n* = 1,930, immigration: *n* = 1,073, COVID-19: *n* = 1,563).

Across countries and types of interaction, about 12.1% of the participants signaled they would react to all three of the fabricated posts, 13.6% stated they would engage with two of the three posts and 19.4% would only engage with one of the three posts. This already indicates a substantial overlap between the interactions with the posts on all three issues.

Table 1. Distributions of people who indicated they would interact with the social media post in percentages

	Climate Change	Respondents (*n*)	Immigration	Respondents (*n*)	COVID-19	Respondents (*n*)
BE	32.0%	313	19.0%	186	23.3%	229
CH	22.8%	270	10.7%	127	16.3%	193
DE	23.8%	222	17.6%	164	21.2%	197
UK	26.3%	342	13.2%	171	21.5%	279
FR	37.4%	443	17.9%	212	29.1%	345
US	36.3%	340	22.7%	213	34.2%	320
N		1930		1073		1563

Note: Percentages in the table represent people who answered 5-7 on a seven-point scale on how likely it is that they would like, share or comment on the post.

Independent Variables

Personality Traits. Based on the literature on the Dark Triad of personality traits (see e.g. Jonason & Webster, 2010; Jones & Paulhus, 2014; Paulhus & Williams, 2002), we focus on its three components: narcissism, psychopathy, and Machiavellianism. The constructs were measured on 7-point Likert scales, based on the studies by Jonason and Webster (2010) and Jones and Paulhus (2014). Narcissism was measured with four statements (e.g., "I tend to want others to admire me" and "I tend to want others to pay attention to me"), Cronbach's α = .88, M = 2.96, SD = 1.55. Psychopathy was measured through three items (e.g., "I tend to lack remorse" and "I tend to be callous or insensitive"), Cronbach's

α = .67, M = 2.89, SD = 1.39. The Machiavellianism scale consisted of four items (e.g., "Make sure your plans benefit you, not others" and "Most people can be manipulated"), Cronbach's α = .68, M = 3.70, SD = 1.32. The complete scales can be found in Appendix B[7].

Political Orientation. The participants had to indicate where they would place themselves on an 11-point ideological scale (0 = extreme left, 10 = extreme right), M = 6.03, SD = 2.46.

Attitudinal Congruence. Attitudinal congruence was measured through agreement with statements about the three issues. For immigration and climate change we worked with existing Likert items that are often used in survey research (see Appendix C[8]). We built mean indices for each issue (climate change, Cronbach's α = .70; immigration, Cronbach's α = .81). For COVID-19 we used one statement stating that the weak response of the Chinese authorities has caused the coronavirus to become a worldwide pandemic. The respondents had to indicate on a 7-point scale (1 = strongly disagree, 7 = strongly agree) if they agreed with the statement shown (climate change: M = 3.33, SD = 1.40; immigration: M = 4.32, SD = 1.50; COVID-19: M = 5.29, SD = 1.70).

Issue Salience. To measure each issue's salience, the participants were asked, "How important do you consider the following issues in [country] at the moment?" (1 = not at all important, 7 = very important). We presented several issues but focus here on climate change (M = 5.27, SD = 1.66), immigration (M = 5.09, SD = 1.66), and finally the broader label of health and medical issues to measure the importance of the COVID-19 pandemic (M = 6.28, SD = 1.06).

Social Media Variables. To measure the social media use, we asked the respondents how frequently they used Twitter, Facebook or Instagram for reasons like entertainment, work, or information seeking on a 5-point scale (1= *never*, 2 = *less often*, 3 = *monthly*, 4 = *weekly*, 5 = *daily*), *M* = 2.73, *SD* = 2.33. Furthermore, we tested for general social media activity. The participants were asked how often they responded to personal posts of friends or family members via likes, shares or comments (1 = *never*, 7 = *very often*), *M* = 4.62, *SD* = 1.97. To measure the perceived trustworthiness of online news, we asked the participants to indicate if they thought they could trust the news on social media (1 = *strongly disagree*, 7 = *strongly agree*), *M* = 3.13, *SD* = 1.53. The means of all independent variables across the six countries can be found in Appendix E[9].

[7] The Online Appendix can be accessed here: http://www.disinformation-project.com/data
[8] The Online Appendix can be accessed here: http://www.disinformation-project.com/data
[9] The Online Appendix can be accessed here: http://www.disinformation-project.com/data

Results

To answer the question of what drives the dissemination of online misinformation, we ran multilevel regressions with random effects to account for the nested data structure. Each participant saw one post for each issue (three posts in total), so that individual willingness to engage with a certain post was nested within respondents. In a first step, we investigated the impact of sociodemographic factors on the willingness to engage with misinforming content (H1a-H1c). In a second step, we entered factors into our models that postulated a connection between the disseminating behavior and certain personality traits (H2a-H2c). Third, we measured the influence of political orientation, specific issue attitudes and the perceived issue salience on the willingness to engage with the fabricated posts (H3-H5). Last, we tested for general social media behavior (H6a-H6b). Although, we expected that our hypotheses would work across countries, we controlled for country differences.

In model I (Table 2) we observed that gender and education were significant predictors. Furthermore, the coefficients stay significant throughout all models, and age becomes significant in model II and model IV, although with a marginal effect (B = .00, SE = .00, p < .001). In support of H1, the findings showed that male, older, and less educated individuals were generally more likely to engage with the post (model IV: gender: B = .19, SE = .03, p < .001, age: B = .00, SE = .00, p < .001, education: B = -.19, SE = .02, p < .001). Results of the second model indicated a greater willingness to engage with misinformation if individuals showed darker personality traits (narcissism: B = .11, SE = .01, p < .001, psychopathy: B = .10, SE = .01, p < .001, Machiavellianism B = .03, SE = .01, p < .05). The significant effect of Machiavellianism disappeared in the following models, which implied correlations with the added predictors. People with stronger degrees of narcissism and psychopathy indicated more often that they would react to the post, in contrast to people without these traits. This result supported H2a and H2b but not H2c.

In the third model, attitudinal and socio-political predictors were included, namely political orientation, issue attitudes and issue salience. In line with our hypothesis (H3), the data showed that the more right-leaning individuals are, the more likely they are to engage with the misinforming content (B = .06, SE = .00, p < .001). Findings, however, also indicated that individuals were more likely to disseminate the posts when they were in line with preexisting issue attitudes (B = .04, SE = .00, p < .001). In support of Hypothesis 4, we found that the stronger the preexisting attitudes towards the presented topic were, the higher the chances were that individuals would like, share, or comment on the message. We additionally tested the influence of perceived issue salience on the engagement with misleading social media content, assuming that the more salient a topic was perceived to be, the higher

the motivation to disseminate the corresponding post. In model III, issue salience had a positive significant effect (*B* = .07, *SE* = .00, p < .001). This finding supported Hypothesis 5.

Model IV further included social media use and the activity on social media. The results led us to accept our sixth hypothesis as well (*B* = .06, *SE* = .01, p < .001). People who tended to engage with posts from friends and family were more willing to disseminate the post. The strongest effect was found for general social media use. Heavy social media users were generally more willing to engage with the misinforming post compared to people with a lower use (*B* = .18, *SE* = .01, p < .001). Furthermore, model IV showed that trust in social media news had a significant effect on the willingness to disseminate the posts (*B* = .12, *SE* = .01, p < .001). The data fit improved stepwise with each model though the inclusion of the different sets of variables, with an explained variance of 18% in the most comprehensive model IV.

Table 2. Random effect models predicting the willingness to disseminate misinformation

	Model I		Model II		Model III		Model IV	
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
(constant)	7.41***	.45	6.80***	.45	5.65***	.44	3.50***	.43
Sociodemographic factors								
Sex	.26***	.03	.16***	.03	.15***	.03	.19***	.03
Age	.00	.00	.00***	.00	.00	.00	.00***	.00
Education	-.25***	.02	-.26***	.02	-.24***	.02	-.19***	.02
Personality traits								
Narcissism			.11***	.01	.12***	.01	.05***	.01
Psychopathy			.10***	.01	.10***	.01	.09***	.01
Machiavellianism			.03**	.01	.02	.01	.02	.01
Political orientation					.06***	.00	.06***	.00
Issue attitudes					.04***	.00	.05***	.00
Issue salience					.07***	.00	.06***	.00
Social media variables								
Social media use							.18***	.01
Social media activity							.06***	.00
Trust in social media news							.12***	.01

Table 2. Random effect models predicting the willingness to disseminate misinformation – part II

Country controls								
CH	-.35***	.05	-.36***	.05	-.30***	.05	-.19***	.05
DE	-.28***	.05	-.27***	.05	-.23***	.05	-.18***	.05
UK	-.19***	.05	-.11*	.05	-.08	.05	-.06	.05
FR	-.01	.05	-.01	.05	.02	.05	-.00	.05
US	.20***	.05	.26***	.06	.27***	.05	.25***	.05
R^2		.05		.08		.10		.18
N		5791		5784		5775		5775

Note: Unstandardized regression weights (B). Reference category for county controls is Belgium. *p < .05. **p < .01. ***p < .001.

In the next step, we aimed to detect issue differences and additionally tested three separate interaction effects with political orientation, issue attitudes, and issue salience. The comparison among the three issues was relevant as not all of them could be connected tightly to a right-leaning ideology. This is especially the case with the post regarding COVID-19. Our survey indicates that individuals with different, even opposing, political ideologies was highly critical of the role China played in this health crisis[10]. Furthermore, the post regarding climate protesters leaving garbage behind in a public space can infuriate individuals from both sides of the political spectrum. Put differently, a climate change sceptic and a climate change believer might both be triggered by green protesters littering a park. One because of the hypocrisy and the other because of the sheer littering and its consequences for the environment. Based on this reasoning, it was relevant to explore these three issues in more detail by consulting interaction effects. Table 3 shows the complete models and the corresponding interaction effects. The two-way interaction effects of the different issues and political orientation (model I) illustrate that the issue of immigration differs significantly from the issue of climate change. A more outspoken right-leaning political orientation seemed to have a bigger impact on the willingness to engage with the immigration post than with the climate change post. Regarding the political orientation, we could not find any differences between the issues of COVID-19 and climate change. Second, the interaction effects of issues and issue attitudes were positive and significant for the topics of COVID-19 and immigration (model II). Both issues varied significantly from climate change with regards to preexisting attitudes, whereas the effect for immigration was stronger than for COVID-19 (B = .08, SE = .01, p < .001 & B = .04, SE = .01, p < .001). Last, we observe a negative significant interaction effect between issues and issue salience for COVID-19 (B = -.04, SE = .01, p < .05) in model III.

[10] We measured people's opinions toward the Chinese government regarding the pandemic by presenting the following statement: "The weak response of the Chinese authorities has caused the coronavirus to become a worldwide pandemic" (1 = strongly disagree, 7 = strongly agree). The statement had a high agreement overall (M = 5.29, SD = 1.70), and across the political spectrum (M_{left} = 4.84, SD = 1.81; M_{right} = 5.78, SD = 1.56; on an eleven-point scale 1-4 was considered left, 9-11 was considered right).

Table 3. Interaction effects for political orientation, issue attitudes and issue salience with issues

	Model I: Interaction effects political orientation		Model II: Interaction effects issue attitudes		Model III: Interaction effects issue salience	
	B	SE	B	SE	B	SE
(constant)	3.61***	.42	3.70***	.42	3.37***	.42
Sociodemographic factors						
Sex	.18***	.03	.18***	.03	.19***	.03
Age	.00***	.00	.00***	.00	.00	.00
Education	-.19***	.02	-.19***	.02	-.19***	.02
Personality traits						
Narcissism	.05***	.01	.05***	.01	.05***	.01
Psychopathy	.09***	.01	.09***	.01	.09***	.01
Machiavellianism	.02	.01	.02	.01	.02	.01
Political orientation	.04***	.00	.05***	.00	.05***	.00
Issue attitudes	.04***	.00	.00	.01	.05***	.00
Issue salience	.05***	.00	.04***	.00	.07***	.00
Social media variables						
Social media use	.18***	.01	.18***	.01	.18***	.01
Social media activity	.05***	.00	.06***	.00	.06***	.00
Trust in social media news	.12***	.00	.12***	.00	.12***	.00
Interaction effects						
COVID-19 x political orientation	.01	.00				
Immigration x political orientation	.03***	.00				
COVID-19 x issue attitudes			.04***	.01		
Immigration x issue attitudes			.08***	.01		
COVID-19 x issue salience					-.04*	.01
Immigration x issue salience					-.01	.01
R^2		.18		.18		.18
N		5775		5775		5775

Note: Unstandardized regression weights (B). Reference category and for the issue it is climate change. *p < .05. **p < .01. ***p < .001

Discussion

Although research on misinformation is on the rise, we know little about what is driving people to engage with this type of content on social media (Pennycook et al., 2021; Pennycook & Rand, 2019). As a key contribution to this strand of research, we aimed in this study to determine whether the process of disseminating online misinformation depends on specific attitudes toward issues or is determined by more general characteristics and behavior, such as personality traits, political orientation, and social media behavior. By combining several possible drivers, structured in a funnel of factors, we aimed to examine whether individuals with certain overriding characteristics contribute willingly or unwillingly to the dissemination of misinformation.

The results of surveys in six countries showed that sociodemographic characteristics matter and are hardly affected by other factors. In line with previous research (Chadwick & Vaccari, 2019; Guess et al., 2019), men and less educated individuals are most likely to engage with misinformation. Furthermore, our findings indicate that personality traits are of great importance in the context of misinformation dissemination. Individuals who showed stronger manifestations of narcissism and psychopathy were more prone to disseminate the misinforming social media post. Moreover, general social media use and activity is a focal predictor when it comes to the interaction with online misinformation suggesting that active social media users are more willing to disseminate misinformation. We found that participants who frequently use social media, generally like, share, or comment on the posts of friends and family members, and have higher levels of trust in news on social media were more likely to engage with possibly false or misleading content. These results are in line with previous studies showing that belief in the content of misinformation is less important for its spread than general social media behavior (Pennycook et al, 2021; Pennycook & Rand, 2019).

Further, in line with previous studies (e.g. Chadwick & Vaccari, 2019; Guess et al., 2019) we observed that political orientation is a strong predictor for spreading misinformation. Although individuals with a right-leaning worldview seem more willing to disseminate misleading messages, it makes sense to consider more specific attitudes related to the issue of the post. In particular, this study shows that attitudinal congruence and issue salience help to explain why people disseminate online misinformation. The more congruent the message was with preexisting attitudes, and the higher the personal importance of an issue, the more likely people were to engage with a post on that issue, even if the content is false or unproven.

By considering the three issues separately regarding political orientation, issue attitudes and salience, minor, but relevant differences came to light. On average, right-leaning individuals were more likely to engage with the misinforming post about immigration, than with the climate change post. In addition,

issue congruence mattered more for the issue of immigration, than for the other two issues. This means that specific attitudes towards the issue are more important for the willingness to engage with the post about immigration than for the other two topics. In line with previous studies on selective sharing (Johnson et al., 2020), this finding suggests that existing attitudes on a polarizing issue such as immigration are particularly important to explaining why people engage with misinformation on immigrants and refugees. However, we must also consider the specific elements of the presented posts on the different issues. For instance, the post on climate change blamed climate change protesters for leaving garbage behind, meaning that being willing to disseminate this post does not mean that individuals question climate change itself, but rather the irresponsible behavior of environmental activists. Thus, we conclude that specific issue attitudes play an important role regarding all issues but are most important for the topic of immigration.

Our study bears some limitations, which must be acknowledged. As a first key limitation, a precise and elaborate measure of attitudinal congruence is very challenging. In this study, we operationalized attitudinal congruence as the agreement with single items regarding the issues. These items were kept very general and did not necessarily reflect the message and specifics of the claims presented to the respondents, except for the post on COVID-19. For a more elaborate measure of attitudinal congruence, future researchers should consider specific stimuli-oriented attitudes toward issues. Second, we used a non-existent news outlet as the source of the fabricated posts to keep the conditions consistent in each country. However, by doing so we neglected the effect of trustworthy or doubtful sources. To avoid this situation, a variety of country specific news outlets should be tested in future research. Third, we used three claims in this study which have been disseminated in various countries. Future research, however, should examine a broader range of issues that bring different ideologies to the fore and appeal to different populations. Fourth, since we presented respondents with multiple cases, we did not ask credibility questions about each post to avoid priming respondents about the intentional misleading/false character of the message (Klar et al., 2020). However, we acknowledge that belief in misinformation is a valid factor for researching the willingness behind the dissemination of misinformation online. Therefore, future studies that focus on a single issue should delve deeper into the importance of the perceived truthfulness or accuracy of a message. Finally, we included individuals in our analysis who indicated they would like or share the posts to signal disagreement because we were interested in the dissemination of misinformation. For the sake of the interpretation of our results, we excluded individuals who indicated they would comment on the posts to signal disagreement. By commenting to disagree, individuals may try to correct the misinformation or warn others about the misleading message. Future research should therefore focus on this on every level of social media reaction.

Despite these limitations, this study provided important insights into which group of people are disseminating online misinformation and which characteristics they share. This is not only helpful for scholars working on this topic, but also for policy makers and organizations combatting misinformation. Weeks and Gil de Zúñiga (2021) suggest that corrective actions should have a persuasive approach to work. We argue that, in addition to more influential attempts at correction, it is also important to target groups of individuals who potentially spread misinformation. Our study shows that there is an overreaching 'profile' of people who can be considered more vulnerable to (un)consciously disseminating false (political) information. Moreover, we also find proof that specific issue attitudes and perceived issue salience explain why individuals would engage with misinforming information. This implies that efforts to limit the spread of misinformation need to target specific groups of individuals which might be caught in social media fringe bubbles (Barkun, 2017). Heavy social media users with strong political opinions and a distinct desire to express themselves form a vocal minority that contributes significantly to the spread of misinformation. Finally, the problem of growing misinformation is likely to continue if people's attitudes become more radical and polarized. With our study, we want to contribute to the literature on misinformation by identifying various drivers for the spread of misinformation, which is particularly relevant in a time of crisis, when untruth is a central threat to societies around the world.

CHAPTER THREE

To Convince, to Provoke or to Entertain?

A Study on Individual Motivations behind Engaging with Conspiracy Theories Online

The growing dissemination of conspiracy theories on social media has challenged the well-being of societies. This study aims to understand why individuals would engage with conspiracy theories and what role specific beliefs, but also individual factors such as personality traits play. To answer these questions, we conducted surveys in six countries (Belgium, Switzerland, Germany, France, the UK, and the U.S.) and investigate three motivations (conviction, entertainment, and reaction provocation) behind the dissemination of conspiracy content on social media. Our findings demonstrate that across issues individuals who indicated they would engage with conspiracy theories do it mainly because they are convinced by the message. Political orientation and issue attitudes proof to be connected to individual engagement with conspiracy theories out of conviction, while dark personality traits such as narcissism and psychopathy are valid predictors for why individuals would disseminate conspiracy theories out of entertainment reasons or to provoke reactions.

Keywords: conspiracy theories, motivations, dissemination, personality traits

Reference: Morosoli, S., Van Aelst, P., & van Erkel, P. (2022). To Convince, to Provoke or to Entertain? A Study on Individual Motivations behind Engaging with Conspiracy Theories Online. *Convergence*, *28*(4), 1030–1059. https://doi.org/10.1177/13548565221105792

Introduction

The spread of disinformation and the emergence of conspiracy theories on social media in recent years can be seen as a major threat to the wellbeing of a society. Recent events, such as the outbreak of the COVID-19 pandemic or election campaigns in different countries, have demonstrated that conspiracy theories and false information spread rapidly on social media platforms and can undercut basic trust in democratic institutions (e.g. Jensen et al., 2021; Mari et al., 2021). The growing scholarly attention for the belief in conspiracy theories has also raised questions about specific behavioral reasons why people would disseminate such problematic content. This is exactly what this study tries to answer, by exploring and identifying the reasons behind the willingness to spread misleading information that is related to conspiracy theories. Although the literature on conspiracy theories is growing rapidly, only few studies investigate *why* individuals engage with and spread conspiracy theories online. Part of the answer might be found in the literature on news sharing. Recent research on general online news sharing has shown that individuals spread news because they want to inform other users, express their point of views, or want to get an idea of other people's opinions (Chadwick & Vaccari, 2019; Thompson et al., 2019). More relevant for the following study, is the growing research on the dissemination of mis- and disinformation. Several studies have tried to explain why people engage with online misinformation (e.g. Chen et al., 2015; Van Bavel et al., 2021). As conspiracy theories can be understood as a specific form of disinformation, many of its driving forces might be applicable to this specific content as well.

Following the broader idea of the Uses and Gratifications Approach (Katz et al., 1974), we argue that people have very distinct reasons or personal motivations to interact with conspiracy theories. Based on this classical approach and more recent literature (e.g. Chadwick et al., 2018), we roughly differentiate between three types of motivations behind this social media behavior: (1) being convinced by the message, (2) entertainment reasons and (3) provoking reactions. The current study is, to our knowledge, one of the first to empirically and systematically compare three specific motivations that have previously been connected to sharing conspiracy theories online. However, these three motivations are by no means exhaustive, but can be seen as proxies for broader (online) behaviors that satisfy certain personal needs. For instance, being convinced by the message of a post can be connected to the individual desire to inform others and to share relevant information (Apuke & Omar, 2020; Thompson et al., 2019), which eventually leads to the spread of the message. Individuals who want to provoke reactions on social media want to trigger interactions and discussions within their network, which can be associated with the socializing aspect (of misinformation sharing) on social media (Chen & Sin, 2013). Sharing conspiracy theories for fun further functions as a gratification. Individuals might spread false information to amuse themselves, experience emotional

release, and anxiety relief, which is common social media sharing behavior (e.g. Islam et al., 2020; J. won Kim, 2014; C. S. Lee & Ma, 2012; Thompson et al., 2019).

In a next step, we aim to measure the influence of individual characteristics, such as personality traits and political orientation on these three types motivations, in order to understand whether different people engage with conspiracy theories out of different reasons. Identifying the motivations behind this problematic sharing behavior is an essential piece of the puzzle in combatting conspiracy theories on social media. We expect that the added individual factors will enable us to paint a more nuanced and comprehensive picture of why individuals would engage with conspiracy content on social media. Hence, we ask the following research questions:

RQ1: To what extent can the online dissemination of conspiracy theories on social media be explained by being convinced by the message, entertainment reasons or reaction provocation?

RQ2**:** Which individual factors determine the motivations behind online dissemination of conspiracy theories on social media?

This study relies on extensive survey data from six Western democracies (Belgium, France, Germany, Switzerland, the United Kingdom and the United States). We aim to answer the research questions by considering two issues that are highly salient and often related to existing conspiracy theories in all countries under study: immigration and COVID-19. Researching sharing behavior, we are aware that people often interact with conspiracy theories on social media without reflecting about their behavior (Buchanan, 2020b; Pennycook & Rand, 2019). Since we are mainly interested in *why* people interact with conspiracy theories these unconscious forms of sharing conspiracy theories fall outside the scope of our study. Subsequently, the strengths of the following study lie in investigating the understudied motivations behind the online dissemination of false information that can be linked to conspiracy theories circulating on social media and in testing different individual factors that can explain these motivations.

Motivations behind the Dissemination of Conspiracy Theories

A substantial amount of literature has discussed the definition of conspiracy theories (e.g. Clarke, 2002; Keeley, 1999; Sunstein & Vermeule, 2008). In this study, we adapt the definition of Keeley, which understands conspiracy theories as "a proposed explanation of some historical event (or events) in terms of the significant causal agency of a relatively small group of persons the conspirators acting in secret" (1999, p. 116). According to this definition, a group of people is trying to benefit politically or economically by acting in secret, harming the common people and violating certain rights (Uscinski et al., 2016). Individuals can be exposed to this kind of information explicitly by it being part of a (news)

story, or more implicitly suggested by the use of certain words and images. It becomes apparent and is easy to imagine that these narratives can be connected to the broader concepts of dis- and misinformation (Wardle & Derakhshan, 2018). Conspiracy theories are oftentimes named together with 'fake news', rumors, and deliberately misleading stories as being specific forms of misinformation (e.g. Halpern et al., 2019; Starbird, 2017; Van Bavel et al., 2021). Therefore, we rely additionally on literature covering misinformation and disinformation when researching the potential motivations behind the dissemination of conspiracy theories. As we often do not know whether people are aware that the stories they interact with are false or misleading stories, we use the term misinformation, instead of disinformation which suggests a clear intention to mislead people.

The underlying approach of our study is based on the classical Uses and Gratifications Theory (Katz et al., 1974). Scholars applying this approach, typically generate a list of functions that different media might have for satisfying a variety of citizens' needs (e.g. Katz et al., 1973). This mass communication work strongly contributed to a shift in thinking about media effects from "what the media do to people" to "what the people do with media". Such a functional approach turned out to be a fruitful way to study the motivations behind the use of 'new' media, that are often more personalized in comparison to the 'old' media (Ruggiero, 2000). In recent years this theory also turned out to be appropriate to study the sharing of false and misleading information (Chen et al., 2015), including conspiracy theories (Apuke & Omar, 2020; Balakrishnan et al., 2021).

We argue that at least three broad motivations in particular capture diverse rationales why individuals would engage with conspiracy theories online. A first gratification or need might be related to *conviction,* being able to show agreement with what you see online and to share these views or moral stances with others (Leiner et al., 2018). Spreading conspiracy theories because users are convinced by the conspiracy content can be connected to attitudinal congruence where the claim of a conspiracy theory matches the preexisting attitudes of an individual (e.g. Hameleers et al., 2021). Second, the very common need for *entertainment* has also been proven to be a valid predictor for why individuals engage with conspiracy theories or misinformation (e.g. Apuke and Omar, 2020; Islam et al., 2020). Individuals who strive for entertainment rather than informing themselves, are more prone to disregard checking the information for its truthfulness. Third, *provoking* reactions on social media as a main motivation to spread false information can be linked to the socialization gratification within the theory, where social media users desire to have conversations with other users and expand their network (Apuke & Omar, 2020). In the section below, we will explore each type of motivation in more detail.

Conviction. First, people might interact with conspiracy theories out of conviction. Drawing on the literature on attitudinal congruent misinformation dissemination, individuals actively or unconsciously tend to select content that is in line with their existing views and avoid incongruent information to reduce cognitive dissonance (Hameleers, 2019; Hopp et al., 2020). In general, this belief-based content selection on social media results in individuals mainly disseminating attitude-congruent social media posts. Even if people are not sure whether the information is true or partly true, they might uncritically decide to share it to support the views of their in-group or community (Clemm von Hohenberg, 2019; Van Bavel & Pereira, 2018). Thus, accuracy seems to be much less important and this can lead to the deliberate sharing of false information because individuals care more about the content being ideologically consistent with them (Pennycook et al., 2021). We believe that these findings can be adopted by the research on motivations behind conspiracy theory sharing on social media. If individuals are convinced by or agree with the message of a conspiracy social media post, they are in consequence more likely to support and spread this messages by liking, sharing or commenting on it. Hence, message agreement and personal relevance can be seen as a central motivation for why people engage with conspiracy theories on social media. We suspect that this is especially the case for liking and sharing. These two reactions can be seen as rather endorsing behavior. Liking is mostly linked to a quick emotional reaction towards a post, which requires little effort and sharing a post is can be connected to sharing one's believes and general self-presentation (C. Kim & Yang, 2017). Commenting on the other hand, is described as 'composed communication' (Burke & Kraut, 2016, p. 266), which demands more effort of users and can not only have an endorsing but also a criticizing purpose.

Entertainment. Second, engaging with conspiracy theories might be considered as a 'fun' thing to do. Social media has been described as a "hedonic information system" (Islam et al., 2020, p. 4), where pleasure such as, fun and entertainment explains some of the social media use. Against this backdrop, we claim that perceived entertainment plays a central role why individuals would disseminate conspiracy theories. Individuals do not only experience conspiracy theories as potentially harmful or misleading, but can also perceive them as funny and entertaining (Daniel & Harper, 2020; van Prooijen et al., 2021). Van Prooijen et al. (2021) observed that the triggered entertainment value through conspiracy theories fosters the belief in them. People experience conspiracy theories as fascinating and exciting. Thus, they can spark intense (positive) emotions, which can predict the belief in and the appeal of conspiracy theories (van Prooijen et al., 2021). Literature on misinformation sharing further suggests that perceived entertainment increases the likeliness to disseminate the false and misleading information (Islam et al., 2020; Metzger et al., 2021). Individuals who seek entertainment on social media platforms might not be bothered with the accuracy and authenticity of a post and therefore disseminate unverified information (Islam et al., 2020). Following these results, we assume that social

media users might therefore disseminate conspiracy theories for fun, even when people do not fully agree with them.

Reaction provocation. Third and last, we consider reaction provocation on social media as a third possible reason for the dissemination of conspiracy theories. Social media use is considered to be highly gratifying (van Koningsbruggen et al., 2017). Individuals post on social media to stimulate conversations and to get opinions from others (Zivnuska et al., 2019). Liking, sharing and commenting on social media posts can be seen as one way to satisfy these needs, for example, to get attention and reactions from others, and start conversations within the network. On a content level, research suggests that information that reinforced COVID-19 conspiracy theories has been proven to get more reactions and be more viral than neutral or debunking content (Papakyriakopoulos et al., 2020). Zhang et al. (2021) found similar results observing that conspiracy theories are more viral than scientific information and have a longer lifetime on social media. In addition, more users are involved with conspiracy theories online and they proliferated more than scientific content (Zhang et al., 2021). In general, conspiracy content generates more likes and shares (Bessi et al., 2015), indicating that this kind of content is attractive to spread for social media users that want to engage with other users. Thus, we argue that a third reason for why conspiracy theories achieve such high engagement on social media is that individuals wish to provoke reactions and gain attention.

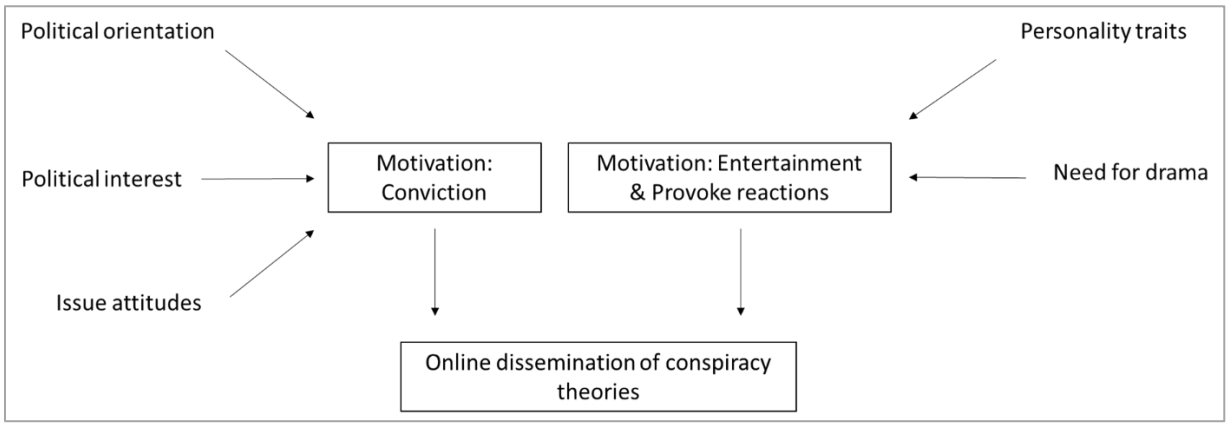
Individual Factors

In sum, we suggest that people interact with conspiracy content because of three types of motivations. These different reasons are not mutually exclusive, but some motivations might be more relevant for certain groups or types of people in society. Therefore, we argue that the three types of motivations to disseminate conspiracy theories online can be linked to a number of individual features and attitudes. This argument is in line with previous research where endorsement for conspiracy theories has been attributed to individual characteristics (van Mulukom et al., 2020). Against this background, we further aim to find out which individual factors can explain why some people share conspiracy theories out of political convictions, whereas others just do it for fun or to provoke reactions.

Based on a previous meta-analysis (Goreis & Voracek, 2019), we divide the individual predictors into two sets (see Figure 1). Goreis and Voracek identified a strand of research that links social and political variables to the belief in conspiracy theories. We therefore expect that political orientation, political interest, and attitudes towards the issue will mostly influence the motivation to disseminate conspiracy theories out of conviction. The authors (Goreis & Voracek, 2019) further found in their systematic literature review that a number of existing studies connect the belief in conspiracy theories to certain personality traits, that can be linked to psychological disorders (e.g. schizotypy). Thus, the

second set anticipates that psychological factors such as dark personality traits and the need for drama have an impact on the motivations to spread false and misleading content out of entertainment reasons or to provoke reactions on social media. In a next step, we elaborate on the individual factors in connection with the specific motivations in more detail.

Figure 1. Factors influencing the dissemination of conspiracy theories on social media



Political factors

Literature on conspiracy endorsement and the belief in conspiracy theories sheds light on the role of political orientation. According to Miller et al.(2016), liberals endorsed conspiracy theories that discredit conservatives and the other way round. Similarly, Min (2021) finds that white men who score high on conservatism are more likely to endorse conservative conspiracy theories. In line with these results, van Prooijen et al. (2015) observed that individuals with extreme political ideologies, were more likely to belief in conspiracy theories. Furthermore, political extremism has been connected to a higher susceptibility to conspiracy beliefs (van Prooijen et al., 2015). Other results suggest that conservatives were more likely to disseminate conspiracy theories than liberals (Mahl et al., 2021). Covering the COVID-19 pandemic, Romer and Jamieson (2020, p. 7) observed that conspiracy beliefs regarding the pandemic were mostly held by individuals with a conservative ideology or by "disadvantaged racial-ethnic groups". These results suggest that people with more conservative, radical or outspoken political ideas are more eager to interact with this type of content and that political orientation is an important predictor to engage with and share conspiracy theories online out of conviction.

Next to political orientation, political interest may also play an important role in explaining the sharing of conspiracy theories based on the conviction about the information. Conservative individuals who discuss politics frequently were stronger associated with the belief in conspiracy theories than individuals who discuss politics less frequently (Min, 2021). Ahmed (2021) linked higher political interest to sharing deepfakes more often, which means that the spread of false information is more

likely to be ascribed to politically interested social media users. In line with this, Chadwick and Vaccari (2019) found that people who purposely spread misleading content were more interested in politics.

Closely related to political orientation, and a potential driver of spreading conspiracy theories out of conviction is attitudinal congruence. The concept of attitudinal congruence dates back to the beginning of cognitive consistency theories and is related to confirmation bias (Festinger, 1957). People actively or unintentionally tend to choose content that is in line with their existing beliefs and avoid uncongenial information to decrease cognitive dissonance (Iyengar and Hahn, 2009; Hameleers, 2019). Broadly speaking, this belief-based content selection on social media results in individuals mainly interacting with attitude-congruent social media posts. Even if people are insecure whether the information is true or partly true, they might uncritically decide to disseminate it to promote the views of their in-group (Van Bavel and Pereira, 2018; Clemm von Hohenberg, 2019; Hameleers, 2019). Literature on misinformation indicates that it is increasingly disseminated within social networks if the message is in line with people's preexisting attitudes towards the issue (Buchanan, 2020). For instance, literature on conspiracy theories surrounding the coronavirus suggests that the belief in them is connected to negative attitudes towards vaccinations (van Mulukom et al., 2020). The more congruent individual's preexisting attitudes are with false information on the coronavirus, the more likely people are to perceive this false information as credible and agree with it (Hameleers, Humprecht, et al., 2021).

Following these findings, we assume that political orientation, political interest, and attitudinal congruence can explain why individuals would engage with conspiracy theories out of conviction. We refer to political orientation, political interest and issue attitudes as 'political factors' and will use this operationalization throughout this study. Thus, the following hypothesis is examined:

Hypothesis 1: The individual online dissemination of conspiracy theories out of conviction can be explained by political factors.

Personality Traits

Very recently, scholars begun to study the influence of certain personality traits on propagating conspiracy theories online (e.g. Hughes & Machan, 2021; March & Springer, 2019; Sternisko et al., 2020). In their systematic literature review, Goreis and Voracek (2019) reported that paranormal belief, narcissism, and desirability of control have all been linked to the belief in conspiracy theories. It is important to understand what role aversive personality traits play in connection to the motivations behind the dissemination of conspiracy theories. We are particularly interested in the so-called Dark Triad of personality traits (i.e. psychopathy, narcissism, and Machiavellianism), since it has been identified as a factor that explains why people would spread COVID-19 conspiracy theories (Hughes

and Machan, 2021). The Dark Triad consists of three elements: psychopathy, narcissism, and Machiavellianism. Psychopathy describes personalities with a lack of remorse, empathy, and anxiety as well as thrill-seeking behavior. It comprises two main elements: deficits in affect and self-control. Narcissism is defined by feelings of grandiosity, dominance, and superiority. The driver behind narcissistic behavior is ego-reinforcement, whereas psychopathy and Machiavellianism are motivated by instrumental gain. Machiavellianism consists of the tendency to manipulate other people in a strategic and calculating manner. The impulsivity associated with psychopathy is central in distinguishing it from Machiavellianism. Machiavellians are, in contrast to psychopaths, concerned about their reputation and tend to plan ahead (Paulhus and Williams, 2002; Jones and Paulhus, 2014). Machiavellians and individuals with narcissistic personality traits have been observed to be less skeptical of conspiracy theories on COVID-19 and by disseminating conspiracy theories on social media, narcissists obtain the attention they strive for (Ahadzadeh et al., 2021). Thus, provoking reactions online. Tang, Reer and Quandt (2022) found that the Dark Triad is linked to what the authors call 'social media disorder' – the disordered (addictive) use of social media ('social media disorder', SMD) and its harms on physical and psychological health (p.1). This relationship is mediated by the motivation to entertain oneself.

Another concept closely related to the Dark Triad is the need for drama. It contains elements of interpersonal manipulation, impulsive outspokenness and persistent perceived victimhood (Lerma et al., 2021, p. 3). Additionally, the concept has been connected to intense social media use (Lerma et al., 2021). Following these premises, we expect that the need for drama can be seen as another explanatory factor for the motivations why individuals would disseminate conspiracy theories for fun or to gain reactions on social media.

Dark personality traits and the need for drama can therefore contribute to the understanding of why individuals would disseminate conspiracy theories for fun or to provoke reactions on social media. We understand dark personality traits and the need for drama as 'psychological factors'. Against this background, the study tests the following hypothesis:

Hypothesis 2: The individual online dissemination of conspiracy theories out of entertainment and to provoke reactions can be explained by psychological factors.

Control Variables

Because conspiracy theories spread relentlessly through social media, it is crucial to take users' general social media activity as a control variable into account when investigating their motivations behind the dissemination of conspiracy theories. When it comes to individual social media use, previous research has postulated significant correlations between the social media use (e.g., in terms of liking, sharing,

and commenting on news) and the belief in conspiracy theories (e.g. Enders et al., 2021; Hall Jamieson & Albarracín, 2020; Romer & Jamieson, 2020). We also assume that (dis)trust influences the dissemination of conspiracy theories online. Conspiracy theories have been observed to undermine the trust in governmental institutions, especially in connection to terrorism and governmental counter strategies (Bartlett & Miller, 2010). Similar findings from Mari et al. (2021) showed that conspiracy beliefs triggered distrust against governmental institutions. The authors further argue that this general distrust can fuel the creation of new conspiracy theories. Bruder and Kunert (2021) found that trust in media was negatively correlated with the belief in conspiracy theories.

Method

Design. We conducted representative surveys in six Western democracies (Switzerland, Belgium, France, Germany, the United Kingdom and the United States[11]) to investigate the motivations behind the dissemination of conspiracy theories related to two issues: immigration and COVID-19. We argue that these topics are suitable for the study of conspiracy theories, because they have been investigated in previous studies in connection with conspiracies (Marchlewska et al., 2018; Papakyriakopoulos et al., 2020; Uscinski & Olivella, 2017) and are highly polarized.

Procedure. The polling company Respondi recruited representative samples of social media users in all six countries based on country-specific census data (see Appendix 2). The specific quotas on age, gender and education were provided by the polling company. In other words, Respondi recruited until all quotas were full and participants were not accepted if their quotas were already complete. Because we are interested in the motivations behind the dissemination of conspiracy theories on social media, the sample only consisted of social media users (usage of at least once a month). The data was collected over four weeks in April and May 2020, and the respondents were given an incentive by the polling company. The surveys were presented in the corresponding language of the (majority of the) country (in Dutch for Belgium and in German for Switzerland). After expressing their informed consent, participants completed the first section of the survey, which included standard demographics and news consumption. In a second step, each participant was exposed to two fictional social media posts containing conspiracy claims about immigration and COVID-19 (see Appendix 2.1). The claims contained blame attributions and institutional skepticism. After being exposed to the posts, the participants were forwarded to the second part of the survey, which included measures for the dependent variable.

[11] The data used in this study are connected to a large-scale research project. The countries were selected to test different contextual factors that create opportunity structures for the dissemination of online misinformation. The following study does not focus on these factors or on specific country differences, rather, we aimed to find generic factors that matter across (Western) countries.

Sample. After removing straightliners and other outliers based on quality fail questions and response time, we collected a sample of 7,009 respondents (quotas for age: *M* = 43.87, *SD* = 14.69; gender: female = 51.7%, male = 48.3%, education: lower = 27.7%, moderate = 40.1%, higher = 32.1%).

Dependent Variables and Vignettes. All of the respondents were shown two social media posts, one for each issue in a random order. The created vignettes resembled news items posted on Facebook by a fictional news outlet (news.com). To enhance the study's external validity, we used statements that had prior circulated on social media and were discussed on various fact checking websites from the six countries under study. We adjusted them to match the purpose of this study. The claim regarding immigration reads as follows: "The news media withholds information about dangerous immigrants – Well known national news media often deliberately don't report about crimes of illegal immigrants against innocent citizens". This claim mainly follows the narrative of the of the Great Replacement conspiracy theory (Cosentino, 2020; Obaidi et al., 2021). The term was coined by Renaud Camus, a disputed French philosopher. This white nationalist conspiracy theory states that the European identity and civilization is in danger because (Muslim) immigrants are plotting against the West, to eventually taking it over in a hostile manner (Bergmann, 2021). We argue that the fabricated post of this study can be connected to this conspiracy theory. In the sense that criminal non-European or non-Western immigrants might be plotting against Western countries, taking them over, and that the corrupt mass media are in cahoots with the masses of immigrants, by covering this fact up.

The COVID-19 post had the following statement: "The coronavirus might be a bioweapon from China – Report claims the Chinese government intended to use COVID-19 to damage the economy of other countries". This message not only calls out the Chinese government for being mainly responsible for the global pandemic, but also suggests that the virus was spread on purpose to harm other countries. Blame attribution and secretive plotting are common features of actual conspiracy theories that can be found on social media (AFP et al., 2020; Uscinski et al., 2016).

The participants were told a cover story at the start of the survey stating that the aim of this study was to measure their opinions on different social media posts, political issues, and actors as well as their social media use. Before they were presented with the fabricated posts, participants were told to envision that the following posts would appear on their newsfeed and to read them attentively.

After reading each vignette they were asked how likely it is that they would a) like the post, b) comment on the post, and c) share the post on a 7-point scale (1= very unlikely, 7= very likely). Each time respondents scored a 4 or higher, they were subsequently asked about their motivation behind that specific form of engagement with that specific conspiracy theory (e.g. liking the immigration post, sharing the COVID-19 post, etc.). These motivations to engage with the social media posts were

measured with four items, all on a 7-point Likert scale (1 = completely disagree, 7 = completely agree). They were asked the following items to tap into the motivation to engage (in a specific way) with the post: "I want to signal that I agree with the content of the post", "I want to provoke reactions", "For the giggles: I want to entertain others and like it for fun" [12]. Since we asked for the motivations for every type of reaction (like, comment, share) to every post we needed to limit the number of motivations. We selected these three specific options, relying heavily on a previous study on individuals' motivation to share news and misinformation online (Chadwick et al., 2018). The authors studied seven motivations which were grouped into three clear motivational clusters: (1) persuading/informing; (2) debating; (3) entertaining/trolling. Persuading/informing was measured through two items (to inform others & to influence others). Debating was measured with two items (to find out other people's opinions & to provoke discussions) and the motivation entertainment was captured through three items (to entertain others; to please others & to upset others). In this study, we chose to take one adapted item for each motivation. Figure 2 displays the mean distribution per reaction and issue, whereas Appendix 2.2 provides detailed tables of the mean distribution per reaction, issue, and country.

Independent Variables

Personality Traits. Based on the literature on the Dark Triad of personality traits (see e.g., Jonason and Webster, 2010; Jones and Paulhus, 2014; Paulhus and Williams, 2002), we rely on its three components: psychopathy, narcissism, and Machiavellianism. The components were measured on 7-point Likert scales, based on the studies by Jonason and Webster (2010) and Jones and Paulhus (2014). Narcissism was measured through four statements (e.g., "I tend to want others to admire me" and "I tend to want others to pay attention to me"), Cronbach's $\alpha = .88$, $M = 2.96$, $SD = 1.55$. Psychopathy was measured with three items (e.g., "I tend to lack remorse" and "I tend to be callous or insensitive"), Cronbach's $\alpha = .67$, $M = 2.89$, $SD = 1.39$. The Machiavellianism scale consisted of four items (e.g., "Make sure your plans benefit you, not others" and "Most people can be manipulated"), Cronbach's $\alpha = .68$, $M = 3.70$, $SD = 1.32$. The complete scales are in Appendix 2.3.

Need for Drama. Based on literature on the concept of need for drama (Frankowski et al., 2016), we measured the concept through three items ("Sometimes it's fun to get people riled up", "I say or do

[12] In the survey included is a fourth motivation stating; "I want to signal that I disagree with the content of the post". However, in this study we specifically focus on motivations behind conspiracy theory sharing, which endorse the spread of them on social media. For the sake of the framing of this study, we therefore excluded this motivation from the multilevel analysis (Table 2). By indicating disagreement, individuals may try to correct the conspiracy theory or warn others about the misleading message.

things just to see how others react" and "It's hard for me to hold back my opinion"), Cronbach's α = .67, M = 3.05, SD = 1.38.

Political Orientation. The participants had to indicate where they would place themselves on an 11-point ideological scale (0 = extreme left, 10 = extreme right), M = 6.04, SD = 2.46.

Political Interest. Political interest was measured on a 7-point scale (1 = not interested at all, 7 = extremely interested). Participants were asked how, generally speaking, interested they are in politics, M = 4.21, SD = 1.87.

Attitudinal Congruence. Attitudinal congruence was measured through agreement with statements about the three issues. For immigration we worked with existing Likert items that are often used in survey research (see Appendix 2.4) and built a mean index (immigration, Cronbach's α = .81). For COVID-19 we used one statement stating that the weak response of the Chinese authorities has caused the coronavirus to become a worldwide pandemic. The respondents had to indicate on a 7-point scale (1 = strongly disagree, 7 = strongly agree) if they agreed with the statement shown (immigration: M = 4.32, SD = 1.50; COVID-19: M = 5.29, SD = 1.70).

Social Media Variables. To measure social media use, we asked the respondents how frequently they used Twitter, Facebook or Instagram for reasons like entertainment, work, or information seeking on a 5-point scale (1= *never*, 2 = *less often*, 3 = *monthly*, 4 = *weekly*, 5 = *daily*), *M* = 2.73, *SD* = 2.33. Furthermore, we tested for general social media activity. The participants were asked how often they responded to personal posts of friends or family members via likes, shares or comments and how often they engage with political and societal news (1 = *never*, 7 = *very often*), Cronbach's α = .75, *M* = 3.38, *SD* = 1.56.

Trust Variables. To measure the perceived trust in news, we asked participants to indicate if they thought they could trust the news most of the time (1 = strongly disagree, 7 = strongly agree), M = 4.23, SD = 1.53. The same question was asked to assess the trust in news on social media (1 = strongly disagree, 7 = strongly agree), M = 3.20, SD = 1.57. Last, we measured the trust in the government by asking the participants how much they trust politicians in government (1 = strongly disagree, 7 = strongly agree), M = 3.40, SD = 1.69. The means of all independent variables across the six countries can be found in Appendix 2.5 and the correlations between all relevant variables are displayed in Appendix 2.6.

Analyses. To answer the question which individual motivations are behind the dissemination of conspiracy theories on social media, we first explore descriptively how many individuals were willing to disseminate the posts and based on which motivations. In a second step, we then explore which

individual factors are connected to each motivation. Within this second step the units of analysis are specific forms of engagement by a respondent (i.e. liking, sharing, and/or commenting) for a certain conspiracy theory (level 1), which are then nested within respondent-issue combinations (level 2), and ultimately respondents (level 3). This means that a respondent can be included in the nested dataset, one – if she indicated to only engage in one way with one of the posts – to maximum six times – if she indicate to like, share *and* comment on both posts. An example of the data structure can be found in Appendix 2.7. Our main dependent variables, the three motivation Likert scales are all situated at the first level. The variable measuring attitudinal congruence is situated at the second respondent-issue level, whereas the stable features of respondents (political orientation, political interest, the psychological factors and the controls) are all situated at the, third, respondent level. To account for the nested data structure with three levels we use multilevel regressions with random effects. In addition, fixed effects are added to account for the further nesting in the two issues (the reference category is the immigration conspiracy theory), and the three types of engagement (the reference category is liking).

Results

In our sample, between 15% and 38% of the social media users in a specific country indicated that they would interact with the presented conspiracy post. Table 1 shows that, across countries, the post that contained the claim regarding the coronavirus led to a slightly higher willingness to engage with it, than the post related to immigration (COVID-19: n = 1,947, immigration: n = 1,559). If we filter out the social media users that indicated that they would interact with the conspiracy post to signal disagreement, we still have between 10% and 25% of participants who would interact with these conspiracy theories.

Swiss citizens were the least likely to interact with the posts, followed by citizens from Germany and the United Kingdom. Participants in Belgium, France and in particular the United States were the most likely to disseminate the conspiracy theories. In the further analysis, we did not focus on the national differences but added country dummies as controls. They largely mirror the variation in Table 1, with United States citizens being significantly more, and German and Swiss citizens being less willing to engage with the posts.

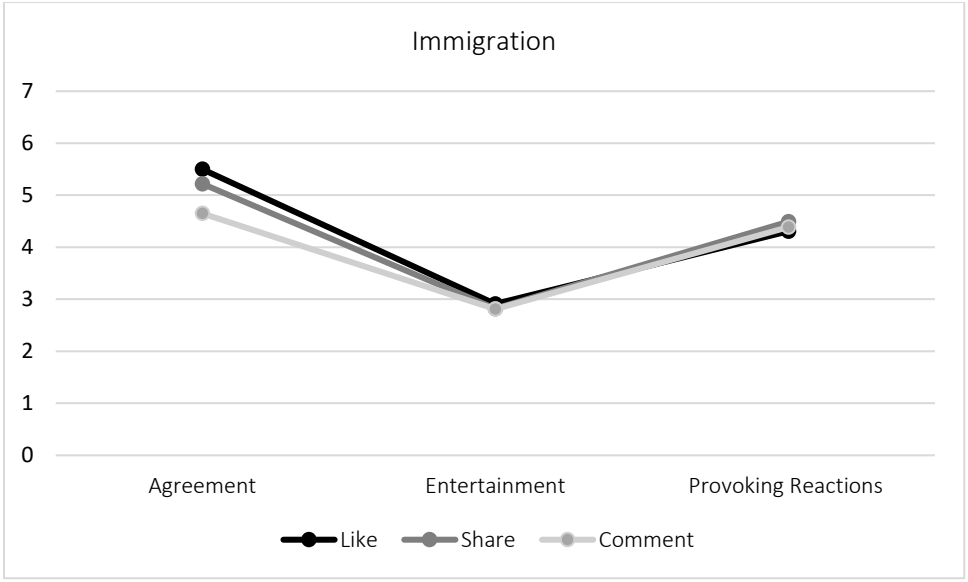
Table 1. Distributions of people who indicated they would interact with the social media post in percentages

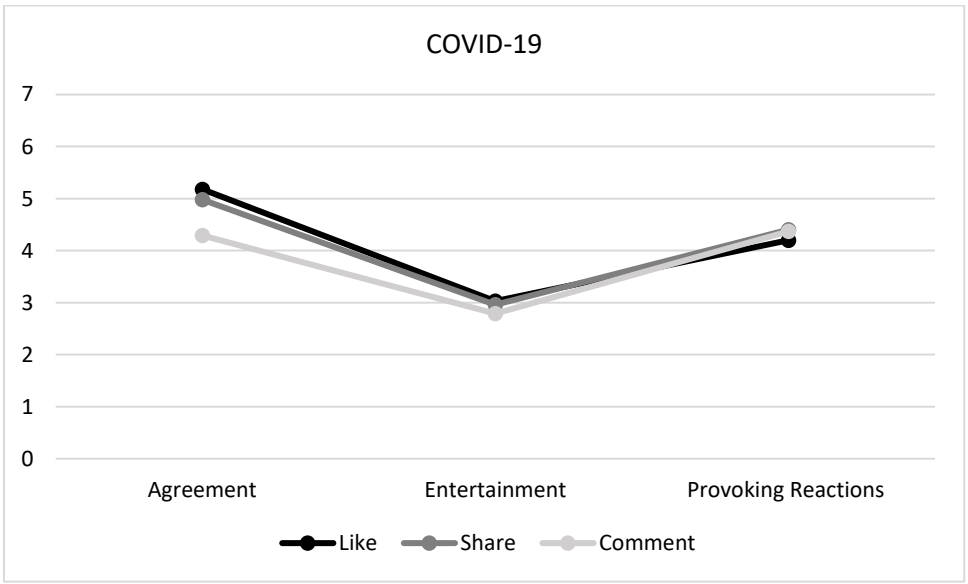
	Immigration	Respondents (*n*)	COVID-19	Respondents (*n*)
BE	25.4%	270	27.9%	297
CH	15.3%	191	18.7%	234
DE	24.7%	251	25.3%	258
UK	18.3%	253	25.6%	353
FR	22.3%	280	32.3%	405
US	30.3%	314	38.5%	400
N	22.3%	1559	27.8%	1947

Note: Percentages in the table represent people who answered 5-7 on a seven-point scale on how likely it is that they would like, share or comment on the post. Included in the analysis were also individuals who indicated they would interact with the post to signal disagreement.

In Figure 2 we separated the data into two groups along the two conspiracy theories under study. On a seven-point scale it shows the means of which motivation to engage with the conspiracy content is the most common and which social media reaction was most frequently indicated per different motivation. Overall, we see that, across issues and type of reactions, agreement with the message of the post was the most important reason for disseminating it (immigration: M_{like} = 5.50, M_{share} = 5.22, $M_{comment}$ = 4.65; COVID-19: M_{like} = 5.18, M_{share} = 4.98, $M_{comment}$ = 4.29). The second most common motivation to disseminate the different posts was to provoke reactions on social media (immigration: M_{like} = 4.31, M_{share} = 4.49, $M_{comment}$ = 4.39; COVID-19: M_{like} = 4.20, M_{share} = 4.40, $M_{comment}$ = 4.37). Entertainment was the least indicated motivation for engaging with the conspiracy theories (immigration: M_{like} = 2.91, M_{share} = 2.82, $M_{comment}$ = 2.81; COVID-19: M_{like} = 3.03, M_{share} = 2.96, $M_{comment}$ = 2.79). This pattern is rather similar for the two conspiracy posts under study. The main difference is that for the post on COVID-19 the difference between three motivations is smaller, indicating that it was relatively more likely to be disseminated for fun. In terms of reactions, we see few differences in liking, sharing and commenting. For both issues, commenting is relatively less done out of conviction, compared to sharing and liking. Overall, the results suggest a pattern where the same motivations to engage with conspiracy theories online dominate across issues and the majority of social media reactions.

Figure 2. Mean distributions of motivations for engaging with the different posts





In a next step, we aim to detect which individual factors influence the motivations behind the willingness to disseminate conspiracy theories online. In model 1 of Table 2 we first of all observe, in line with Figure 2, that the COVID-19 post was less likely to be disseminated out of conviction (*B* = -.15, *SE* = .05, p < .01), and more related to entertainment (model 2) than the immigration issue (*B* = .20 *SE* = .03, p < .01). Looking at the different types of reactions, the results indicate that liking, compared to commenting (model 1: comment: *B* = -.16, *SE* = .03, p < .01), is the form of engagement that is mostly

motivated by conviction about the conspiracy theory and by entertainment. For sharing we could not find any significant differences between the different motivations in comparison to liking.[13][14]

Next, we turn to our main hypotheses. We expected that political factors can be more associated with the dissemination of conspiracy theories out of conviction (H1) and that psychological factors can be more linked to the online dissemination out of entertainment reasons or to provoke reactions on social media (H2). In general, focusing on the influence of the political factors, we observe that having an attitude that is congruent with the post plays a central role for the motivation to engage with conspiracy theories out of conviction. Congruence with the issue of the post does, however, not matter for entertainment reasons. If we look at provoking reactions, we also see a significant effect of attitudinal congruence. However, this effect is much weaker compared to its effect on conviction (conviction: $B = .17$, $SE = .02$, $p < .01$, provoking: $B = .05$, $SE = .02$, $p < .01$).

If we next look at political orientation we find that it is a significant predictor for both the motivation conviction and entertainment. Especially right-leaning individuals are more likely to engage with conspiracy theories out of conviction and entertainment. However, again, this effect is much stronger for the motivation to spread the messages out of conviction or agreement (conviction: $B = .16$, $SE = .01$, $p < .01$, entertainment: $B = .03$, $SE = .02$, $p < .05$). Thirdly, political interest is mostly a significant negative predictor regarding the entertainment motivation. In sum, we conclude that the political factors are mostly and more strongly linked to the motivation of being convinced by the conspiracy theory, but that they do influence the motivations of provoking reactions and entertainment to a limited extent. Therefore, hypothesis 1 can only be partly confirmed.

Zeroing in on the second set of individual predictors, the psychological traits, we find that individuals with psychopathic traits more often indicate that they would disseminate the posts based on all three motivations. Nevertheless, it has a much stronger effect on the entertainment motivation (psychopathy: $B = .32$, $SE = .03$, $p < .01$) and a much more moderate effect on the conviction and provoke motivation. If we look at narcissism we see a more or less similar pattern. This trait especially has an influence on the entertainment motivation and a more moderate influence on the provoke motivation. As expected, narcissism is not behind the conviction motivation. A Dark Triad trait that does have a strong influence on the conviction motivation, however, is Machiavellianism ($B = .15$, $SE = .04$, $p < .01$). Finally, as expected, we find that individuals with a higher need for drama are more likely

[13] Note that if we change the reference category to sharing, we find similar significant differences between sharing and commenting, with commenting being less done out of conviction and entertainment than liking.

[14] This is further supported by the fact that the correlations between the motivations for liking and sharing are quite high (above .76). This suggests that future studies could combine these two types of reactions in one single question.

to engage with the conspiracy theories out of entertainment reasons and to provoke reactions (entertainment: *B* = .13, *SE* = .03, p < .01; provoking reactions: *B* = .31, *SE* = .03, p < .01). The need for drama is not related to sharing a post out of conviction. Thus, although most, but not all results are in line with our expectations, we can partly accept the second hypothesis.

Our control variables furthermore show that older individuals are generally more likely to disseminate the conspiracy posts because they agree with the message, whereas younger and male individuals are most likely to disseminate posts for entertainment reasons. Older and male participants are generally also more prone to like the posts for reasons of provoking reactions on social media reasons. If we look at social media activity, we find that more active social media users are more likely to engage with the conspiracy theories out of entertainment reasons and to provoke reactions, as do those who have a higher trust in social media, although for this group also conviction is an important motivation. Trust in traditional news on the other hand does not help to differentiate between the different motivations.

Table 2. Random effect models predicting the motivations for disseminating conspiracy theories

	Model 1 *Agreement/Conviction*	**Model 2** *Entertainment*	**Model 3** *Reaction provocation*
Ref = immigration			
- COVID-19	-.146**	.199**	.071
	(.051)	(.033)	(.039)
Ref = Like			
- Share	-.047	.024	.04
	(.028)	(.028)	(.03)
- Comment	-.161**	-.05*	-.029
	(.03)	(.029)	(.031)
Congruent attitudes	.166**	-.021	.045**
	(.02)	(.014)	(.017)
Political orientation	.158**	.028*	.021
	(.014)	(.012)	(.014)
Political interest	-.034	-.044*	.004
	(.023)	(.019)	(.023)
Narcissism	.015	.224**	.087**
	(.029)	(.024)	(.028)
Psychopathy	.109**	.324**	.09**
	(.032)	(.027)	(.031)
Machiavellianism	.149**	-.017	.03
	(.036)	(.03)	(.035)
Need for drama	-.014	.13**	.313**
	(.033)	(.027)	(.032)
Male	-.07	.149*	.161*
	(.077)	(.064)	(.075)
Age	.013**	-.015**	.011**
	(.003)	(.002)	(.003)
Education	-.106*	.034	.058
	(.049)	(.041)	(.048)
Social media activity	.052	.055*	.205**
	(.028)	(.023)	(.027)
Trust in news	-.064*	-.003	.01
	(.027)	(.022)	(.026)
Trust in social media	.164**	.188**	.078**
	(.028)	(.023)	(.027)
Ref = Belgium			
Switzerland	.097	.028	-.135
	(.141)	(.117)	(.137)
Germany	-.054	.229*	-.073
	(.136)	(.113)	(.132)
UK	.328*	.304**	-.04
	(.128)	(.106)	(.124)
France	.02	.018	.801*
	(.125)	(.104)	(.122)
US	.403**	.361**	.176
	(.126)	(.105)	(.123)
Constant	3.237**	-1.044	-1.652
	(1.106)	(.92)	(1.077)
$\sigma^2_{respondent}$	.313**	.25**	.409**
$\sigma^2_{respondent-issue}$	(.028)	(.02)	(.02)
$\sigma^2_{residual}$	.134**	-.622**	-.35**
N (observations)	**6531**	**6526**	**6523**
N (respondents)	**3160**	**3157**	**3157**

*Standard errors are in parentheses **p<.01,*p<.05*

Discussion

Despite the fact that research on the belief in conspiracy theories is growing, little is known about what motivates people to disseminate this type of content on social media. The present study sought to understand the motivations behind the dissemination of conspiracy theories online and whether these motivation can be related to certain individual factors such as personality traits and political orientation.

Our analysis found that being convinced by or agreeing with the message was the primary motivation for disseminating conspiracy theories online. This is in line with previous studies on selective sharing (Johnson et al., 2020), suggesting that existing attitudes on a polarizing issue such as immigration are particularly important to explain why people engage with misinformation on immigrants, and potentially start believing the broader conspiracy theory that immigrants are here to 'replace' us. Given the results of the surveys, we identified provoking reactions as the second most indicated motivation to engage with false and misleading content. Individuals were willing to like, share or comment on conspiracy posts in order to trigger reactions and gain attention within their social media network. Entertainment played a less prominent role on average, but was still mentioned by a significant group of people when it comes to the motivations behind the spread of conspiracy content. Thus, the three selected motivations can, to some extent, explain what drives people to engage with conspiracy theories on social media.

By considering two sets of individual factors, relevant differences between the three selected motivations came to light. Since we were interested in individuals who would engage with conspiracy theories, we filtered the sample accordingly to determine which characteristics match these individuals. First of all, we expected political attitudes to be strongly related to engaging with conspiracy theories out of conviction. This turned out to be largely the case. Attitudinal congruence with the issue of the post mattered clearly more for disseminating conspiracy theories based on conviction than for the other two motivations. In addition, right-leaning individuals were also more likely to engage with conspiracy content out of agreement. Finally, the more politically interested people were, the less was entertainment driving this type of behavior. In sum, more political outspoken and interested people are mainly engaging with conspiracy theories because they believe that the message has to be told. We believe this adds to previous studies, which postulated that political orientation is a strong predictor for disseminating misinformation. (e.g. Chadwick & Vaccari, 2019; Guess et al., 2019).

Our second main expectation, was that psychological factors help to explain the motivations to disseminate conspiracy theories based on perceived entertainment and reaction provocation. Here

our findings paint a more mixed image. We do show that narcissism and the need for drama are of great importance in the context of conspiracy sharing out of entertainment and provoking reactions. Individuals with narcissistic traits might be more willing to disseminate conspiracy theories to be in the spotlight and get recognition. This result is in line with previous research, where narcissistic individuals were found to share conspiracy theories to gain wanted attention online (Ahadzadeh et al., 2021). Individuals who show stronger manifestations of psychopathy are more prone to disseminate the conspiracy theories out of all three reasons but the effect is the most outspoken for entertainment reasons. For Machiavellianism, however, the effect goes in the opposite direction as this trait was mostly related to agreement and conviction. To some extent this makes sense, as those who score higher on this trait may be more inclined to share congruent posts out of instrumental and strategic (political) reasons to influence others.

Hypotheses 1 and 2 postulated relationships between political and psychological factors and the individual motivations to disseminate conspiracy theories online. The results largely confirm the hypotheses. Political orientation and congruent attitudes towards the issue influenced the motivation to disseminate conspiracy theories out of conviction or message agreement. Whereas psychological factors such psychopathy, narcissism and the need for drama are more related to the motivations to spread false and misleading content out of fun or to provoke reactions on social media. At the same time our results are less clear cut than our hypotheses suggested. The two set of factors have some influence on all three motivations. This is not surprising as our motivations are not mutually exclusive, and often co-determine the online dissemination of conspiracy theories.

An additional takeaway of this study manifests itself in the observed similarities between the motivations to engage with conspiracy theories and with misinformation. We did not find exceptional motivation patterns that are unique to disseminating conspiracy theories. We rather find similar mechanisms behind sharing misinformation and messages related to conspiracy theories. Nevertheless, we are aware that this might be the case because our design was limited to three non-exhaustive motivations on only two conspiracy theories and thus, were not able to adopt a more nuanced lens. We urge future research to test more diverse motivations directly connected to the Uses and Gratification Approach (e.g. spreading content to inform others or experience emotional release) and apply them separately to conspiracy theories and specific forms of misinformation. Doing so, enables to get a better idea whether conspiracy theories should be considered as ‘exceptional’ or not, when it comes to sharing false or misleading information.

The findings of this study should be considered in the context of several limitations. As a first key limitation, we selected three motivations and disregarded other possible reasons for the dissemination

of conspiracy theories online. Especially, qualitative studies could advance and deepen the research of different motivations behind this behavior. In this study, we also focused on political and psychological variables, but of course many more factors can be of importance. For instance our control variables, suggest that age and gender are very relevant factors when trying to explain why people engage with conspiracy theories.

Second, we asked the respondents after they indicated that they would engage with the conspiracy posts about their motivations to do so. This raises the concern that individuals rationalize their decision ex-post. Nevertheless, we argue that ex-post rationalization is unlikely to singularly drive our results. The fact that we find individuals who indicated that they would engage with conspiracy related messages also just for fun or to provoke reactions, proofs that individuals are also willing to present themselves as 'trolls' or attention-seeking and not only as reasoned individuals. However, we do admit that the sharing of false information is not necessarily a conscious process, but can also be seen as a rather automatic or spontaneous one (Buchanan, 2020) that requires little cognitive effort (Pennycook & Rand, 2019). However, our study focused on the individual conscious willingness to spread conspiracy theories online. By first asking respondents whether they would interact with these posts, and next why, this study comes closer to understanding the deliberate dissemination of conspiracy theories, rather than the fast and unconscious sharing, which would require a different methodological approach.

Third, we used a non-existent news outlet as the source of the conspiracy posts to keep the conditions consistent in each country. However, by doing so we disregarded the potential important effect of trustworthy or doubtful sources. To avoid this, a variety of country specific news outlets should be considered in future studies. Finally, since we presented respondents with multiple issues we did not ask for every claim whether they believed it, to avoid priming respondents about the false or conspiracy nature of the message (Klar et al., 2020). However, we acknowledge that the belief in conspiracy theories is a valid predictor for researching the motivations behind the dissemination of conspiracy theories online. Therefore, future research which focuses on single issues should delve deeper into the importance of the perceived truthfulness or accuracy of a message.

Fourth and finally, our results show that the tested motivations hold for individuals across different Western countries with different contexts. Nevertheless, we did not delve deeper into the differences between these countries and how our dependent variable may be influenced by factors such as a country-specific belief system, information environment or political system. We encourage future studies to take a more explicit comparative design with a larger sample of countries into account and to look into these factors as well.

Yet, despite these limitations, we contributed to a better understanding of why certain people interact with conspiracy theories online. These findings not only have implications for scholars but also for policy makers. Our study suggests that different policy responses or efforts are needed to combat the spread of conspiracy theories among different groups, as they do not always have the same motivation behind sharing the post. While some people need to be persuaded that a certain conspiracy theory is incorrect, this approach might be unsuccessful among people that believe that engaging with this misleading content is simply funny or a good way to get attention in their personal network. Finally, our study shows that the fight against conspiracy theories needs to be adjusted to the specific topic of the conspiracy. For instance, people might like posts about immigration because they are really convinced that the media hides stories about violent immigrants, but share a story about COVID-19 as a bioweapon of the Chinese government much more to provoke reactions. The ongoing global COVID-19 pandemic will likely keep fostering the growth of conspiracy theories on social media. With our research, we hope to have contributed to a more inclusive understanding of the motivations behind the dissemination of conspiracy theories to be able to mitigate and combat the spread in the future.

CHAPTER FOUR

“Come on people, this is clearly fake”

How Social Media Users React to Political Online Misinformation by Commenting

Abstract

Political misinformation is thriving on social media and many individuals are unwillingly confronted with it. The comment section offers an opportunity for users to freely engage with online misinformation. However, little to no research has explored how people react to this potentially harmful content by commenting on it. To fill this void and to identify patterns reactions, we use a manual content analysis of user comments on three, as misinformation recognized, tweets involving UK politicians (N = 600). Our main findings show that the majority of the user comments disapproved the misinforming tweets and only a fraction expressed support. Furthermore, a vast amount of the comments contained uncivil language. These findings help us to shine a light on how individuals respond online misinformation and alert citizens, who are willing to call out misleading content.

Keywords: online misinformation, user comments, Twitter, engagement

This chapter has been submitted to ‘Communications: The European Journal of Communication Research’ at the time of writing this manuscript and is awaiting reviewer scores. It is a single authored paper that has been presented at the 2022 IJPP conference.

Introduction

It is by now a well-established finding that social media platforms are a fertile ground for the spread of online misinformation (Duffy et al., 2019; Hameleers et al., 2020). Misinformed citizens pose a threat to democratic societies, which is especially true if they are ill-informed about politicians and political processes. Exposure to misinformation has been connected to an increase in political polarization, a decrease in media trust, and an increase in political cynicism (Allcott & Gentzkow, 2017; Jones-Jang et al., 2021; Stubenvoll et al., 2021). Therefore, the concern about the influence misinformation can have on citizens is not only present among policymakers, journalists, and scholars, but also among part of the public at large. Consequently, in an online environment where conspiracy theories, disinformation, and counterfactual narratives thrive, we raise the question how ordinary individuals react to online misinformation – in terms of support, warning, and disagreement but also in terms of discourse characteristics.

In the context of this study, we understand misinformation as false and misleading information that is spread or created without the intention to harm (Wardle & Derakhshan, 2018). Even though there is a possibility that the actors under study spread the false information deliberately to harm others (disinformation), it is impossible for us to prove that this was the case. Furthermore, in this study we focus on active social media users and not on passive users. We are interested in individuals who are commenting on and actively engaging with misinforming content.

Various studies on commenting on misinformation have focused on the correction, reduction, and detection of misinformation (e.g. Albahar, 2021; Jiang & Wilson, 2018; Majid & Pal, 2020; Masullo & Kim, 2021). However, little is known about broader reactions to political misinformation and little research has looked in detail and in a systematic way at individual user comments on misinformation. The work of Metzger and colleagues (2021) touches upon if people signal belief or disbelief in misinformation and comes closer to finding out if individuals support or disagree with misleading content. However, the authors do not pursue a more fine-grained approach by, for instance, measuring discursive elements such as incivility. The following study takes a step back and will capture the bigger picture of reacting to political misinformation by commenting on it. This study does not discriminate towards the direction of the comments, by for example only looking at comments that counter misinformation, but we also look at all other possible reactions. Additionally, we measure other essential characteristics of online comments, such as incivility and the direction of the comment (e.g., directed at the author or at the issue at hand).

Because of how the algorithms of social media platforms work, every comment can make false and misleading information more visible – no matter if the comment is supportive or disapproving of the

content. We therefore argue that social media users' reaction to misinformation can vary substantially and are an essential piece of the puzzle in understanding the success of online misinformation. Misinformation is especially successful if it has been shared by people's trusted sources (Metzger et al., 2021). So might misinformation that has been commented on by a trusted contact, reach audiences that blindly disseminate it because of this relationship of trust. The lack of research investigating these simple but essential statements of social media users is striking. In particular, analyzing the comment section allows us to explore how individuals react to online misinformation. Understanding what type of misinformation triggers the most comments, which leads to the amplification of the content, can help us to combat misinformation in a targeted manner. Against this background, we assess the following overarching question: How do individuals react to online misinformation on Twitter involving politicians?

To answer this question, we rely on a manual content analysis of user comments that were posted on three misinforming tweets involving UK politicians. We focus on three, by fact checkers as misinformation identified, tweets and analyzed the 200 most recent comments after the content was posted. We are specifically interested in two groups of individual reactions: supportive and disapproving and we will delve deeper into how these two groups differ in how they react (e.g., referring to other sources; expressing anger). The study will also take relevant concepts, such as online incivility into account to better understand the processes behind the consumption of online misinformation. Together, this paper aims to classify potential user reactions to contribute to a better understanding of how individuals react to online misinformation involving politicians.

Effects of User Comments

On social media platforms, citizens have the possibility to directly engage with the content posted by media outlets, other users or political actors. Depending on the platform, they can like it, share it, and comment on it[15]. Since we are focusing on commenting as the main reaction to misinformation, it is important to discuss the implications of this type of reaction and to distinguish it from liking and sharing. The three different reactions can be seen as a scale of effort. On the first level there is liking. Liking a post is mostly connected to quick emotional reactions towards social media content, which requires little cognitive effort. It is simply about consuming information and not creating a message (C. Kim & Yang, 2017). A single click is enough to engage with the content. Second, sharing is on a higher level when it comes to consuming a message. Sharing can be connected to sharing one's believes and general self-presentation (C. Kim & Yang, 2017). Commenting on the other hand, is described as

[15] On Twitter, users can like, retweet, which equals to sharing on Facebook, and reply to a tweet. Replying equals to commenting. In this study, we will use the term commenting instead of replying out of consistency reasons, even though we focus on Twitter.

'composed communication' (Burke & Kraut, 2016, p. 266), which demands more psychological effort of users.

Taking a step back from misinformation research, it is essential to highlight the general importance of user comments. Previous research suggests that commenting plays a crucial role in general news participation and information evaluation. On the one hand, comment sections enable users to voice their opinions and directly engage with news content (Stroud et al., 2016). On the other hand, user comments themselves have the ability to "highlight certain elements of the news, thereby guiding (i.e., priming) subsequent cognitive processes" (E.-J. Lee & Tandoc, 2017, p. 440). Specifically, previous research has observed that user comments can influence individuals' perceptions about the media's portrayal of presidential candidates and the third-person effect about to what extent online news influences the political attitudes of other individuals (Houston et al., 2011). While reading manipulated news articles with attached comments, participants expected the partisan comments to have a stronger influence on other individuals than on themselves in terms of supporting or disapproving of a presidential candidate (Houston et al., 2011). Focusing on social media comments, Gearhart et al. (2020) find similar results. With the help of an experiment, the authors observed that after being exposed to user comments with which participants disagreed, individuals had an increased perception of media bias (Gearhart et al., 2020). Suggesting, that being exposed to this type of comments makes social media users feel like the news article, the outlet or the writer are heavily biased against their views. Generally speaking, user comments have the ability to reduce the perceived quality of a news item (Prochazka et al., 2018). Not only were perceptions about the news quality, the media bias, and third-person effect affected by user comments, but also news credibility. A substantial number of studies has investigated the impact of uncivil or negative comments on news credibility (Naab et al., 2020; Thorson et al., 2010; Waddell, 2018, 2020). Findings reveal that critical or uncivil comments decrease the overall perceived credibility of news articles (Naab et al., 2020; Waddell, 2018). Against this backdrop, we argue that user comments can substantially affect how we perceive information. It becomes apparent that we know more about the effects of user comments on news perceptions than on what the comments actually entail. This study is set out to fill this empirical lacuna.

Commenting on Misinformation

On social media, commenting can normatively be seen as a contribution to collective representation and participation (Galpin & Trenz, 2019). Individuals make the conscious decision to substantially contribute to a discussion about the social media post by commenting on it. For instance, they can warn other users about the message or engage with the original poster of the content. Comment

sections are spaces of opinion formation and deliberation (Galpin & Trenz, 2019), which makes them relevant to study, especially in the context of online misinformation.

The above listed results suggest that studying user comments can be a fruitful tool measure individuals' engagement with misinformation. Previous research has predominantly focused on misinformation correction and detection when researching user comments (e.g. Albahar, 2021; Bode & Vraga, 2018; Majid & Pal, 2020; Masullo & Kim, 2021). An analysis of user comments in response to conspiracy corrections showed that more than half of the comments were counter-correction and only 11% were pro-correction. The content of the anti-correction comments consisted for example, of distrust in institutions and refusal of evidence (Majid & Pal, 2020). Majid and Pal (2020) conclude that corrections might not be the solution to mitigate the belief in conspiracies and rumors, since they do not seem to be accepted easily by users. Another study which examined user comments that correct misinformation found that uncivil comments posted on a news story on Facebook led users to evaluate the correcting comments and commenters less positive. The credibility of the news story however was not affected by these uncivil comments (Masullo & Kim, 2021). Colliander (2019) finds similar effects: after being exposed to correcting comments on misinformation, individuals were less prone to share the false content and likelihood for them to post a positive comment decreased as well. In general, correction on social media platforms can happen through users commenting on false and misleading posts pointing out to others that this content is incorrect (Bode & Vraga, 2018).

A second, less developed, strand of research recognizes user comments as a tool to detect false and misleading messages. Based on crowd intelligence, comments can signal the truthfulness or falseness of online content (e.g. Albahar, 2021; Guo et al., 2018; Shu et al., 2019). For instance, user comments can contain warnings or references to trustworthy sources to point out to other users that the content they see is false or misleading.

Despite the already existing literature on commenting on misinformation, there is, to the best of our knowledge, little to no research yet on how social media users react to political misinformation by voicing their opinion through user comments. Little is known about the supportiveness or the degree of disagreement of these user comments and research is still scarce on the disruptiveness or constructiveness of user replies. To investigate user comments on political misinformation, we therefore draw our inspiration from previous, similar research from other fields. We relied mainly on online incivility literature to measure uncivil comments, deliberation studies, and online political discussion literature. Hence, the main goal of this study is consequently, to discover patterns in the diversity of comments individuals post on possibly false and misleading tweets.

RQ1: How do individuals react to online misinformation on Twitter involving politicians?

In this study we are specifically interested in two groups of individuals: individuals who are supportive of the misinforming tweets and who are disapproving. We will delve deeper into how these two groups differ in how they react. We will each consider different levels of support and disagreement, for example, does the user comment question the content of the post or does it add further supporting details to the tweet. Based on this, we ask the following set of research questions:

RQ2a: How do users express their support for the misinforming posts?

RQ2b: How do users express their disagreement with the misinforming posts?

Emotionality and Incivility in Online Comments

Triggered emotions or sentiment in connection with online misinformation play a central role when studying user comments. Emotional false and misleading messages flourish on social media and gain a lot of traction. Zollo et al. (2015) measured the sentiment of comments on conspiracy posts and science posts. The authors find that user comments on conspiracy posts were a lot more negative (55%) than on the science posts (29%) (Zollo et al., 2015). Studying comments on false news stories on Weibo, Guo et al.(2019) observe that user comments contained more anger, sadness, and doubt than comments on a real news story. Similarly, Klimiuk et al. (2021) detected anger as the most commonly expressed emotion linked to conspiracy-theories and misinformation on vaccines on Polish social media. Ani-vaccine comments were generally angrier than pro-vaccine user comments (Klimiuk et al., 2021). Previous research has further shown an increase in the use of swear words in the comment section when the veracity of a post decreases (Jiang & Wilson, 2018). These findings are a possible indication that misinformation triggers rather negative emotions and it is of high interest to measure emotional language when individuals are expressing their opinions in the comment section.

Connected to triggered emotional expressions, incivility plays an essential role in user comments, as previously discussed (Prochazka et al., 2018; Thorson et al., 2010). Incivility has the ability to impair the online sphere and damage even certain perceptions of reality. Civil information about political actors can be seen as a representation of reality, whereas uncivil information can diverge from facts and hence from reality (Hameleers, van der Meer, et al., 2021). This can foster the already existing divide between political in- and out-groups and make people believe in untruths. There are various definitions of incivility and the concept is hard to grasp. One person might find a remark or debate uncivil but another person might see it as a heated but fair debate (Papacharissi, 2004). However, in this study, we understand online incivility as "features of discussion that convey an unnecessarily disrespectful tone" that do not "add anything of substance to the discussion"(Coe et al., 2014, p. 660). It entails any type of offensive claim that is harmful for deliberation and democratic communication

(e.g. Anderson et al., 2014; Papacharissi, 2004). Consequently, incivility can take on many shapes and forms, for instance name-calling, partisan attacks, and intimidation (Hameleers, van der Meer, et al., 2021; Papacharissi, 2004). Uncivil user comments can be divided into two categories: 'uncivil attack posts', such as attacks on media, other users, and 'uncivil language posts', for example capital letters (shouting) and profanities (Rösner et al., 2016, p. 462). The valid concerns over online incivility have increased the attention of the scientific community and resulted in a multitude of research on the effects of incivility. For instance, incivility has been proven to enhance polarization by leading people to feel hostile towards other parties other than their own and to strengthen favorability of their own party (e.g. Y. Kim & Kim, 2019). Online incivility has also been found to boost aggressiveness in users when exposed to aggressive comments (Rösner et al., 2016). Kim and Kim's (2019) findings suggested that being exposed to uncivil opposing user comments leads to more negative emotions, lower willingness to read more comments, and higher levels of polarization. Research on incivility and user comments connected to misinformation has shown that incivility in correcting comments can lead to negative evaluations of the comments or the commenter, even though the users try to correct false and misleading information (Masullo & Kim, 2021). Kim and Masullo Chen (2021) find in their study on the tone of user comments in response to misinformation that civil comments were perceived as more credible than uncivil comments. The authors also showed that the credibility of a news story was not affected by uncivil comments (J. W. Kim & Masullo Chen, 2021).

Against this backdrop, we expect that there will be uncivil and emotionally loaded language in user comments in connection with political online misinformation. In this study we take a closer look at the different forms of uncivil language, for example name calling or stereotyping and we measure the direction of the incivility in the user comments. We therefore pose the second set of questions:

RQ3: What role does incivility play in user comments on misinforming posts?

RQ4: What role does emotional language play in user comments on misinforming posts?

Method

To assess social media comments on political misinformation, we conducted a manual content analysis of the comments belonging to three misinforming posts involving UK politicians. Manual coding was necessary, as many comments contained subtle utterances of support or disagreement with the content as well as speech elements such as sarcasm and therefore would not be captured by automated coding.

Case

Since we were interested in comments on political misinformation, we applied a inductive approach. From previous research we know that the UK, next to the US, has the most fact checking websites which publish articles involving politicians (Humprecht, 2019). Based on this, we first consulted UK fact checking websites to look for misinformation involving politicians. We exclusively looked for online misinformation posted on social media platforms because we were interested in user reactions in the comment sections. In the following study, we selected three fact checked tweets involving UK politicians. According to the Digital News Report, Twitter is the second most frequently used social media platform for news consumption (13%) and therefore a valid case to study how individuals react to misinformation (Newman et al., 2022). The sample selection is based on two different types of misinformation, identified by previous literature (Hameleers, Humprecht, et al., 2021): "*partially false* or decontextualized information and *completely false* or fabricated information" (2021, p. 3).

The first tweet involves Nadine Dorries, the Secretary of State for Digital, Culture, Media and Sport since 2021. The supposed tweet reads: "I am fed up of the loony left moaning about the rising demand for food banks. Every family needs to live within their means. Food doesn't grow on trees!" and is posted by an account as a screenshot. The fact checking organization Full Fact proofed that the tweet is fake, meaning Nadine Dorries never posted it and someone else had to create it. It was posted by a fake profile that is known to post made up tweets involving Brexit (Ashworth-Hayes, 2022). This corresponds with the second type of misinformation: completely false and fabricated (Hameleers, Humprecht, et al., 2021). The second tweet was posted by Nigel Farage, a former politician and party leader of the UK Independence Party and Brexit Party until 2021. In his tweet, which is a video message he calls for "the EU Human Rights Act" to be scrapped. The Ferret, a fact checking organization based in Scotland confirmed that there is no such law as the EU Human Rights Act, but it is a UK law which was passed in 1998 (Ferret Journalists, 2021a). Their verdict goes as follows: "There is no such legislation as the EU Human Rights Act so it is impossible to call for it to be scrapped. The Human Rights Act is a piece of legislation drafted and passed by the UK Parliament that incorporates the provisions set out in the European Convention on Human Rights"(Ferret Journalists, 2021a). This means that in his tweet, Nigel Farage is evidently disseminating misleading information. Hence, this tweet can be linked to first type of misinformation, which entails partially false information that stays close to reality (Hameleers, Humprecht, et al., 2021). The third tweet shows a series of pictures of the House of Commons during multiple debates and accuses MPs to only show up in big numbers when they are debating MPs' payment and expenses. The picture, which has been circulating online since at least 2016, was posted by the Scottish writer Irvine Welsh. The Ferrets' verdict is that this tweet is mostly false (Ferret Journalists, 2021b). Certain pictures correspond with the correct debate but the two

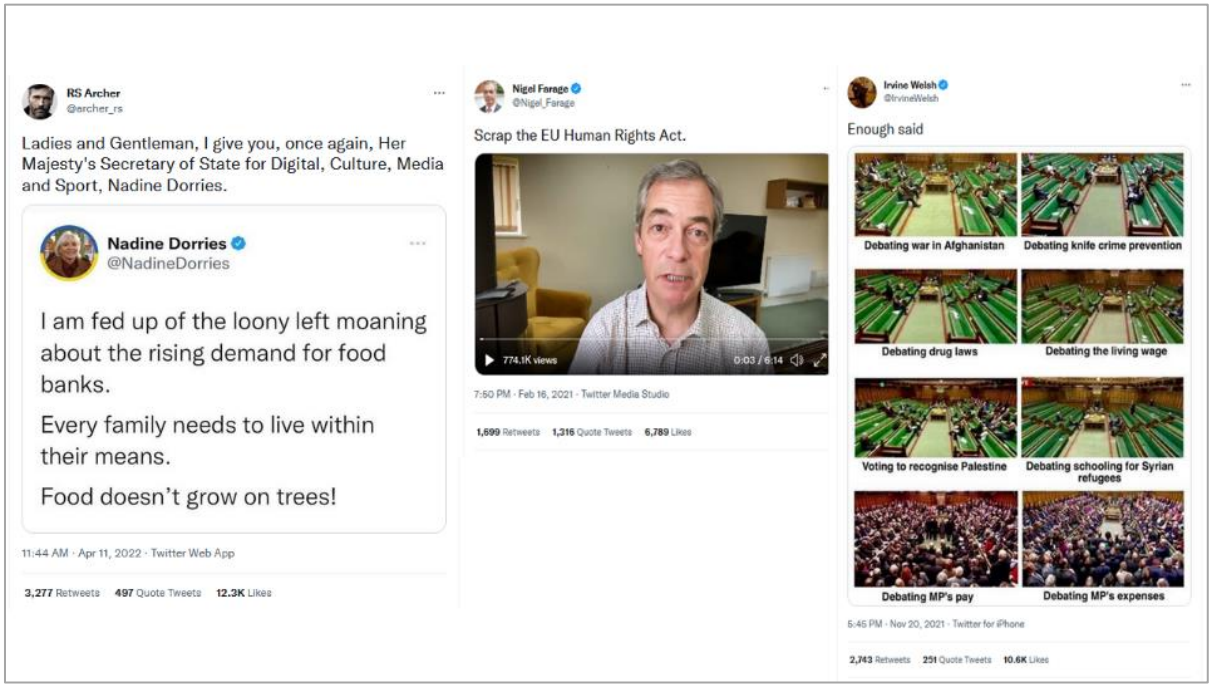
photos where MPs are apparently debating their pay are false. They match a debate in 2004 about tuition fees and the opening of the parliament in 2010 (Ferret Journalists, 2021b). This tweet puts correct information out of context and therefore corresponds with the first type of misinformation (Hameleers, Humprecht, et al., 2021). Table 1 shows an overview of the basic characteristics of the tweets.

Table 1. Characteristics of the Tweets

	Original Poster	**Fact Checker**	**Date**	**Likes**	**Retweets**	**Misinformation Type**
Nadine Dorries	RS Archer (fake profile)	Full Fact	11.04.2022	12.3K	3277	Fabricated / completely false information
Nigel Farage	Nigel Farage (politician)	The Ferret	16.02.2021	6789	1699	Partially false information
Parliament	Irvine Welsh (novelist)	The Ferret	20.11.2021	10.6K	2742	Decontextualized information

In addition to distinguishing between different types of misinformation, other criteria had to be fulfilled for the tweets to be in included in our sample. First, the tweets needed to clearly contain misinformation linked to UK politicians according to different fact checking organizations. We are interested whether the reactions in the comments vary, depending on which political actor is involved. The sample only contains three tweets because it was challenging to find other, clear examples of misinformation tweets which were recently published either by one of the political actors or about a political actor in the UK. Second, the posts should have gained a lot of traction after it published (more than 5000 likes). This ensured that the tweet was seen by a large audience, which increases the chance of more diversity and discourse in the comment section. Finally, the tweets had to involve different political actors based in the United Kingdom. Figure 1 displays the tweets used in this study.

Figure 1. Tweets under study



Sample and Coding

The sample comprised the 200 most recent comments after the respective tweets were posted (Total N = 600). To make the manual coding feasible we selected the first 200 comments. The advantage of this sample is the ability to capture the initial reaction of users in the first minutes or hours after the tweets were published. The downside of this limited sample is that we might miss the final development or the direction in which the complete set of comments might go. However, we believe that the first 200 comments give us a good idea what type of different reactions exist and how the initial reaction of the most active users looks like. Two of the comments were not available anymore because either the user profile that posted it does not exist anymore or because the comment was deleted, these two comments were replaced by the next ones available. For every comment we first assessed the readability or comprehensiveness of it. If a comment was incomprehensible it was eliminated from further coding (Brennan and Prediger coefficient = 0.97). The comprehensiveness was determined based on two existing codebooks (Naab & Küchler, 2021; Stromer-Galley, 2007). A central variable that was coded was whether a comment supports or disagrees with the misinforming tweet (Brennan and Prediger coefficient = 0.87).

In a second step we coded in detail what the supportive or disapproving comment contained. We each coded four different items. For supportive comments we measured whether users *agreed* with the statement that was made in the post (i.e. "I agree", "I also think that …is a good idea"), or whether the comments were expressing *author support* or appreciation for the individual who posted the

misinforming content. We further coded if the comments *developed* the claims of the posts (i.e. restate, clarify or exemplify the content of the post or add supporting details) or if they *tracked* the content. Tracking refers to attempts to get additional information in order to fully and correctly understand the content of the post. Disapproving comments were coded in a similar manner. First we identified whether users *disagree* with the content (i.e. signaling opposition, contradicting the statement, calling out that the post might be false) or if they express *author disapproval*. In addition, we measured if comments *question* the i) the relevance, ii) the legitimacy or iii) the veracity of the post. Finally, consistent with coding supportive comments, we were interested if user comments tracked the misinforming posts.

As affective language is an important factor when it comes to user comments, we coded six different emotions based on previous research on comments (G. H. Tran & Ngo, 2018): happiness, unhappiness, security, insecurity, satisfaction, dissatisfaction (Brennan and Prediger coefficient = between 0.60 and 0.93). Finally, we focused in detail on uncivil language and the direction of the incivility. Based on an overview of Esau (2021) we selected seven different forms of uncivil language: name calling, vulgarity, discriminatory language, political stereotyping, references to lying, sarcasm, and shouting (Brennan and Prediger coefficient = between 0.75 and 0.90). The direction of incivility was measured in four ways (Rowe, 2015; Seely, 2018): author-directed, political party or ideology directed, content-directed or directed at other options (Brennan and Prediger coefficient = between 0.60 and 1.0).

Much consideration went into the quality of the coding process. Two coders were trained extensively, and the codebook was pilot-tested multiple times. After the first pretest, the coding scheme was improved significantly resulting in a high intercoder reliability across items.

Results

Descriptively on the comment level[16] we find that most of the comments on Twitter disagree with the misinforming posts (see Table 2). In total 77% (407 out of 528) comments disapprove of the content of the misinforming tweets and only 16% of the replies are of supportive nature. A small fraction of the user comments could not be assigned to one of the three categories (other = 6%) and we found just two cases of trolling (0.2%). Focusing on the three different tweets, we observe that users are the most supportive in their commenting behavior linked to the tweet showing multiple misleading pictures of the House of Commons (30%). Nigel Farage receives about 18% of supportive comments, whereas the tweet involving Nadine Dorries almost exclusively gets disapproving comments (91.4%).

[16] We excluded comments from the analysis that were not comprehensive, n = 72.

We were also interested whether users used sources in their comments. Only 14 out of 494 comments mentioned a source, for example by posting a link to newspaper articles or official government websites. 13 out of these 14 sources are connected to disapproving comments.

Table 2. Number of supportive or disapproving comments per post

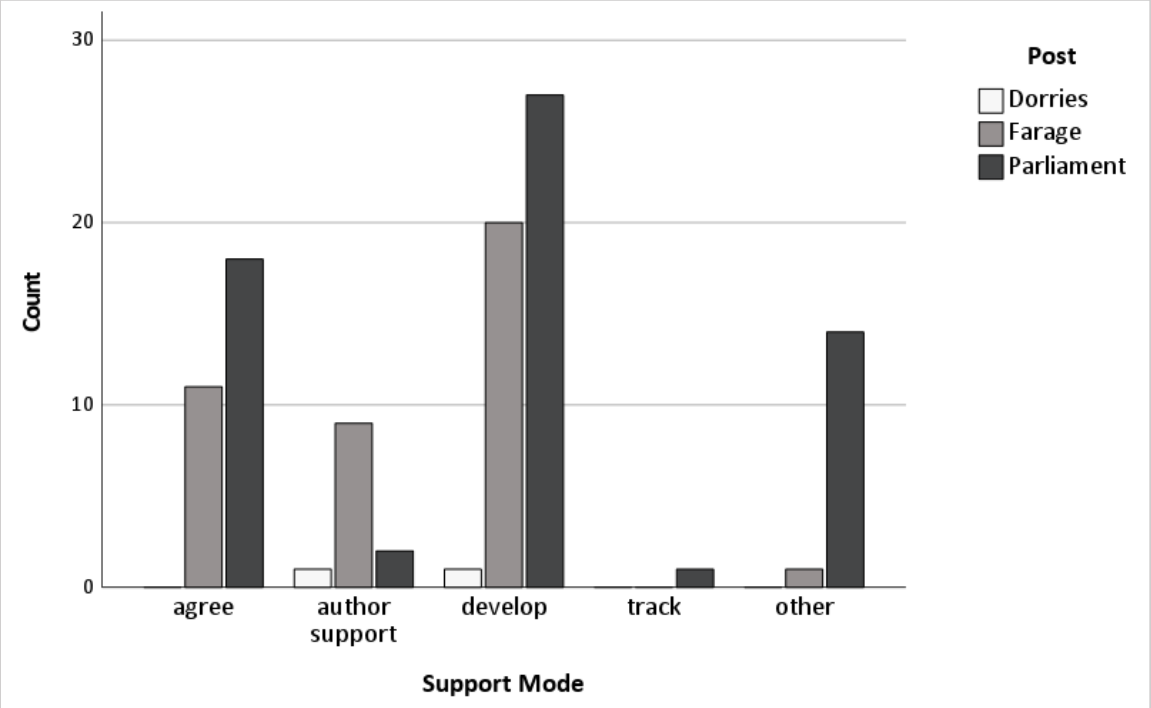
	Post							
	Dorries		Farage		Parliament		***Total***	
	N	%	*N*	%	*N*	%	*N*	%
support	2	1.1%	30	17.6%	55	30.1%	87	16.5%
disapprove	160	91.4%	128	75.3%	119	65.0%	407	77.1%
trolling	0	0.0%	1	0.6%	0	0.0%	1	0.2%
other	13	7.4%	11	6.5%	9	4.9%	33	6.3%
Total	175	100.0%	170	100.0%	183	100.0%	528	100.0%

χ^2=56.810, df=6, Cramer's V=0.232, p=0.000

Next, we measured how users expressed their support or disagreement for the misinforming tweets involving UK politicians (RQ2a & RQ2b). The main goal of this study is a first broad framework of user comments on misinformation. For this classification, we each consider five modes of support or disagreement. We base these five categories mainly on previous research on news comments on Facebook (G. H. Tran & Ngo, 2018), which adapted the conversation analysis framework and appraisal theory. More specifically, did we consider the different types of attitudinal language in the appraisal theory (Martin & White, 2005). Figure 3 shows the different ways of support for each post. We observe that the post regarding members of parliament debating in the House of Commons gets the most agreeing comments, followed with the tweet of Nigel Farage. Nigel Farage also receives the most support as an author of the tweet, since some users are supporters of Farage and his views. The most common supportive reaction overall are comments that reiterate or add (develop) to the content of the post. These are comments that restate, exemplify, and clarify the statement, or comments that add further supporting details to the content. For instance, “As much as that angers me, at least people voted for them. The House of Lords is where democracy dies in this so-called ’UK’” (ID143, parliament) or “I don’t know why people are upset. Was this not what people voted for? Try volunteering at any food bank around you and ask recipients who they voted for. You will be shocked that 80% voted for Tories.” (ID72, Dorries). The tweet concerning the parliament receives the most developing comments, followed by Nigel Farage’s tweet. Tracking, which refers to attempts to get additional information in order to fully and correctly understand the content of the post is the least frequent mode of support. The tweet involving the members of parliament also collects the most reactions that could not be

matched with a category. These comments are, among other things, exclamations such as: disgusting!; yikes; shameful.

Figure 2. Bar chart with mode of support expressed in comments per post

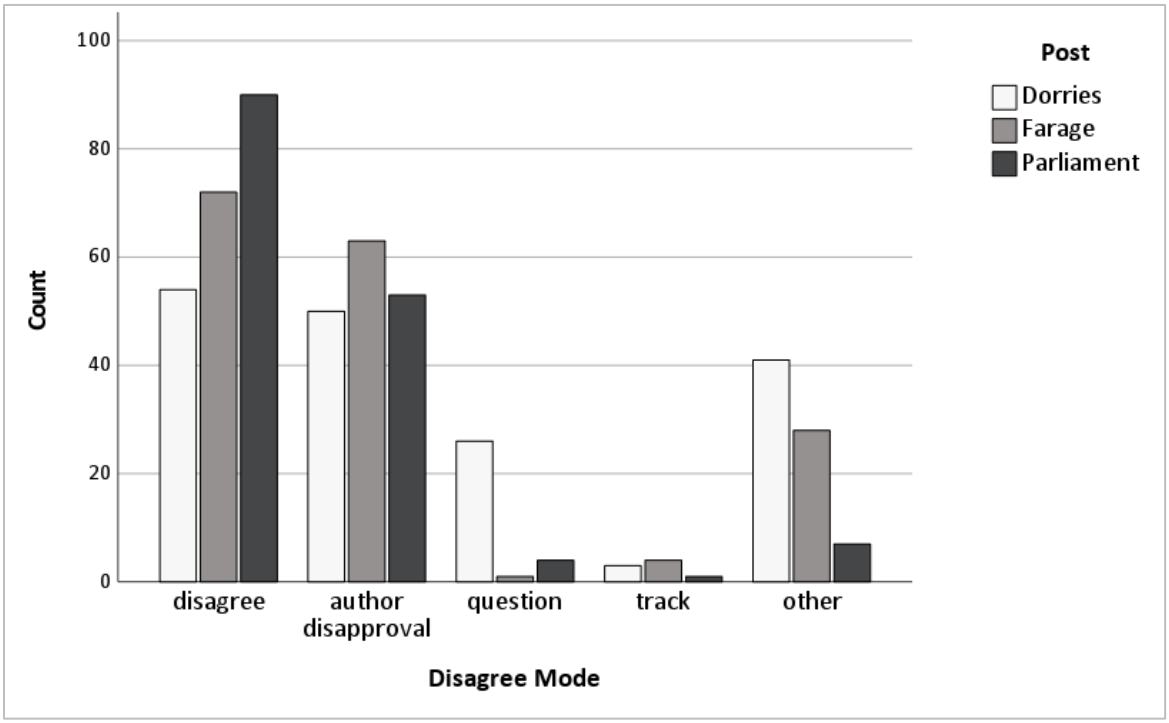


Note. The figure only considers comprehensive and supportive comments, *n* = 87.

Turning to the different ways of disapproving of misinformation, we see that the majority of the comments disagree with each of the three tweets and, consistent with supporting comments, tracking is also here the least frequent mode of disapproving (see Figure 4). Disagreeing comments also include warnings and remarks that the tweet might be fake. Interestingly, the most disagreement is triggered by the tweet involving the House of Commons. Thus, this tweet seems to attract the most agreeing but also the most disagreeing comments. Even though, 50% of the comments called the tweet out to be pictures taken during different debates than the post claimed, a large number of individuals still did agree on the general message of the post. For instance, “Correct sentiment but this is incorrect.” (ID15, parliament) or “Unfortunately it’s fake.” (ID161, parliament). After that, gain the tweets of Nigel Farage and the post concerning Nadine Dorries the most disagreeing or countering comments. The majority of the disagreeing comments linked to Nigel Farage’s tweet, are correcting him regarding the misinforming statement he made. For instance, “There is no such thing as an "EU Human Rights Act". There is, however, the Human Rights Act 1988, which is a BRITISH law giving effect to the human rights set out in the European Convention on Human Rights” (ID50, Farage). In total, 28.5% of the all the comments posted on Farage’s tweet are of corrective or counter active nature. Nigel Farage also

receives the strongest author disapproval. The original poster of the tweet concerning MP's gains the second most disapproving comments, followed by the author of the Nadine Dorries tweet. This tweet's authenticity is also questioned the most. A large number of users rely on others to confirm if the tweet is fake or not. Individuals asked: "I can't find this quote on her account. Anyone else managed it?" (ID37, Dorries) or seek help through comments like: "I'm not sure this is a genuine tweet. Can anyone confirm?" (ID45, Dorries). Related to that, 27 comments in total call out the tweet for being fake: "Come on people, this is clearly fake" (ID57, Dorries) or "Really not helpful to tweet fake stuff" (ID165, Dorries). The tweet involving Nadine Dorries also receives the most other disapproving comments, followed by Nigel Farage and the House of Commons tweet. These comments include sarcastic remarks and the request to post more information on the statement. For example: "Wait until she hears about apples." (ID194, Dorries) and "@Twitter why can I not report this post?" (ID84, Farage). We are also interested in whether users refer to fact checking organization or other sources that would debunk the misinforming tweets. We found only three comments that did this. One commentor posted a link to a source under the post of Nigel Farage, explaining the issue about the Human Rights Act. The two remaining comments were posted on the tweet involving MP's. One user commented on how to identify false information on Twitter and Full Fact, the fact checking organization itself referred to their article on that misleading picture.

Figure 3. Bar chart with mode of disagreement expressed in comments per post



Note. The figure only considers comprehensive and disapproving comments, *n* = 407.

We further examine if incivility appears in the user comments and in which direction it points (RQ3). We focus specifically in detail on incivility in connection with disapproving comments because we only find 17 mentions of incivility in supportive comments out of a total of 86 comments. Name calling is the most frequent incivility in supportive comments. In comments that disagree with the three posts, we observer that 37.3% of all the comments contain incivility. Table 4 shows the types of incivility per post. Overall, sarcasm (16.2%) was the most applied incivility in disapproving comments, followed by voiced vulgarities (7.4%), name calling (6.1%) and references to lying (5.2%). Zeroing in on the different tweets, we see that comments concerning the tweet about Nadine Dorries are consistent with the overall pattern: sarcasm (14.4%), vulgarity (8.1%) and name calling (4.4%), in that order, are used the most often. Nigel Farage's tweet triggers the most uncivil reactions. Sarcastic comments are also the most frequent (28.1%). Different than the reactions on Nadine Dorries, Nigel Farage receives the most name calling comments (7.8%), followed by vulgar comments (4.7%) and comments with accusation of lying (3.9%). The tweet involving members of parliament paints a different picture. The majority of the uncivil comments contain references to lying (12.6%). 9.2% of the disapproving comments use vulgar expressions and 6.7% are calling politicians names.

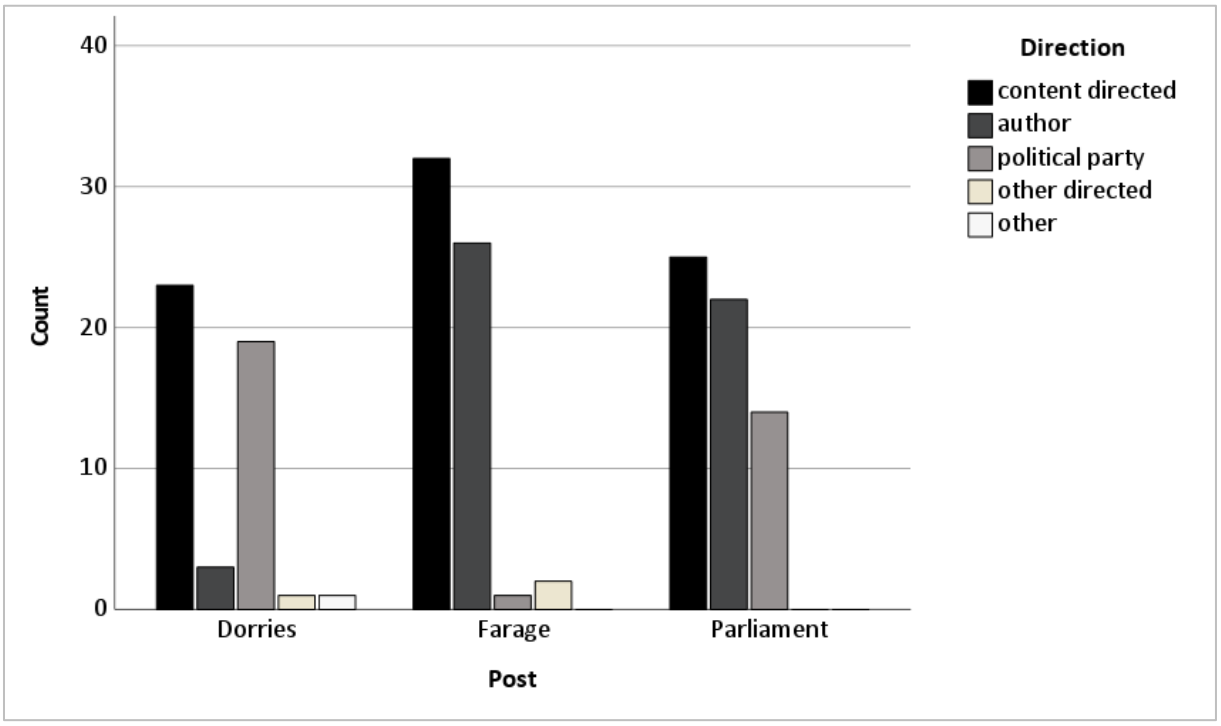
In a second step, we measure the direction of the uncivil comments (see Figure 5). The incivility is in all three cases most often directed towards the content (50%). A third (32.9%) of the uncivil comments are directed towards the author of the tweet and 14.5% at a specific political party or candidate. Concentrating on the single tweets, we notice that for the post involving Nadine Dorries, the incivility is mostly content directed and directed towards her political party/ideology or her as a candidate. Nigel Farage's tweet triggers the most uncivil comments which point towards the content or him as a creator of the tweet. The tweet regarding the members of the House of Commons generates the most uncivil comments that are directed towards the content, the author, and towards politicians.

Table 4. Type of incivility per post

	Dorries	Farage	Parliament	***Total***	
	%	%	%	N	%
None	70.6%	53.1%	62.2%	255	62.7%
Name calling	4.4%	7.8%	6.7%	25	6.1%
Vulgarity	8.1%	4.7%	9.2%	30	7.4%
Political stereotype	1.9%	0.8%	3.4%	8	2.0%
Lying	0.6%	3.9%	12.6%	21	5.2%
Sarcasm	14.4%	28.1%	5.9%	66	16.2%
Other	0.0%	1.6%	0.0%	2	0.4%
Total	100.0%	100.0%	100.0%	407	100.0%

Note. Only comprehensive cases who disapproved of the posts are considered (n = 407). χ^2*=52.257, df=14, Cramer's V=0.253, p=0.000.*

Figure 4. Direction of incivility per post



Note. The figure only considers incivility in comprehensive and disapproving comments, n = 152.

Finally, turning to the emotional language used in comments on political misinformation (RQ4), we observe that in almost half of the cases, affective language in supportive comments is not present (47.1%). For example: “I don’t know why people are upset. Was this not what people voted for? Try

volunteering at any food bank around you and ask recipients who they voted for. You will be shocked that 80% voted for Tories." (ID72, Dorries) or "Everywhere is just like that!" (ID22, parliament). Almost half of the disapproving comments do not display affective language or they cannot be linked to one of the categories (49%). These comments include statements like: "No rights for: migrants, the victims of British war crimes. Got it." (ID80, Farage) or "Did she actually tweet this?!" (ID56, Dorries).

In Table 5, comparing the different tweets, we see that the tweet involving MP's provokes the most comments with affective language (n = 174), followed by the tweet about Nadine Dorries (n = 161) and Nigel Farage (n = 158). Within the supportive comments on Nigel Farage's tweet, dissatisfaction is the most expressed feeling (7 out of 30 comments) and only 4 comments utter happiness. For the tweet involving the members of parliaments, unhappiness is the most used emotion (n = 17) and then follows dissatisfaction (n = 12) and satisfaction (n = 2). Disapproving comments only display negative emotions. For every tweet, dissatisfaction is the most common emotion expressed in the comments and unhappiness the second most frequently used emotion.

Table 5. Contingency table of affective language in supportive and disapproving comments

		Post					
		Dorries		Farage		Parliament	
		N	%	*N*	%	*N*	%
supportive	Happiness	0	0.0%	4	13.3%	0	0.0%
	Unhappiness	0	0.0%	1	3.3%	17	30.9%
	Security	0	0.0%	1	3.3%	0	0.0%
	Insecurity	0	0.0%	1	3.3%	1	1.8%
	Satisfaction	0	0.0%	0	0.0%	2	3.6%
	Dissatisfaction	0	0.0%	7	23.3%	12	21.8%
	None	2	100.0%	16	53.3%	23	41.8%
	Total	2	100.0%	30	100.0%	55	100.0%
disapproving	Unhappiness	16	10.1%	11	8.6%	24	20.2%
	Insecurity	0	0.0%	1	0.8%	0	0.0%
	Dissatisfaction	39	24.5%	66	51.6%	42	35.3%
	Other	5	3.1%	1	0.8%	2	1.7%
	None	99	62.3%	49	38.3%	51	42.9%
	Total	159	100.0%	128	100.0%	119	100.0%

Note. χ^2*=57.2796, df=12, Cramer's V=0.242, p=0.000.*

Discussion

Against the setting of misinformation being heavily and easily disseminated on social media, it is relevant to understand how individuals react to misinformation. It is especially important to disentangle user comments on political misinformation to grasp the wide spectrum of individual reactions. We made a first attempt at capturing supportive and disapproving user comments on misinforming tweets. We filled an existing research lacuna by trying to classify the characteristics of supportive or disapproving comments on proven false and misleading messages involving politicians.

We showed that the majority of the comments did disapprove or disagree with the three misinforming tweets and only a fraction of twitter users supported the messages in these posts (RQ1). This result can be seen as mostly consistent with previous research, which found that commenting is a way of correcting misleading messages and pointing out to other that the content is false (Bode & Vraga,

2018). However, we do see a big difference in supportive comments, comparing the three different tweets. Nigel Farage's and the tweet concerning the House of Commons, received a lot more supporting comments than the tweet about Nadine Dorries. We believe that context matters. Nigel Farage personally posted the tweet and his supporters, did of course comment in his favor. The tweet involving Nadine Dorries on the other hand, was posted by a fake account, was made up and had a very disrespecting quote in it. We think that even her supporters or the supporters of her party, were not willing or too skeptical to express their support. We suspect that the tweet on the MPs', which was posted by a Scottish writer, found support on both sides of the political spectrum because it accused MPs' in general.

We further delved deeper into the question *how* individuals express their agreement or disagreement with misinforming tweets (RQ2a & RQ2b). The most common supportive reaction overall are comments that developed the content of the post. These are comments that restate, exemplify, and clarify the statement, or comments that add further supporting details to the content. We did find some differences in supporting comments per tweet. We observed that the post regarding members of parliament debating in the House of Commons gets the most agreeing comments, followed by the tweet of Nigel Farage. Nigel Farage's tweet additionally received the most author support. This again can presumably be connected to the fact that he himself posted the tweet and that he has quite a big following on Twitter. Turning to the different ways of disagreeing with misinformation, we found that the majority of the comments disagree with each of the three tweets. Disagreeing also took the shape of calling out the content or the author of the tweets to warn others about it. Even though Nigel Farage received the most author supportive comments, he also does receive the strongest author disapproval. We believe this is an indication that Nigel Farage is a highly polarized public figure. Furthermore, the tweet on the members of parliament found the biggest agree- but also disagreement. Hence, even though a big part of the comments called the tweet out to be false, a large number of people still did agree on the general message of the post. It appears that the two comment groups (support/disapprove) mirror each other regarding these two items.

Literature suggests that incivility is a common appearance in user comments and we measured if this is also the case for comments on political misinformation (RQ3). The majority of the comments did not express uncivil language. However, depending on the tweet, this was just a little over 50%. This indicates that still a great number of comments were uncivil. These comments were mostly sarcastic and displayed nasty language such as vulgar expressions or name calling. Consistent with previous research, we found two categories of uncivil user comments: attacking comments and comments with simply uncivil language, such as profanities (Rösner et al., 2016). We did find small differences between the incivility in the different tweets that can be traced back to the content of the posts. The tweet

involving MP's has the most uncivil comments with reference to lying, which were mostly connected to the fact that politicians are liars and do not care about important societal issues but only about their salaries. The comments about the tweet involving Nadine Dorries showed the most vulgarities. These can be traced back to the derogatory remark in the fake tweet and users were insulting Dorries. The direction of these incivilities was slightly different for each tweet. They were mostly content directed. As the following direction, the comments on Nigel Farage's tweet and the tweet about the House of Commons aimed at the authors but the comments on Nadine Dorries addressed her as a politician or her party.

Finally, we found that the selected misinforming tweets triggered negative emotions in user comments, which is conform with previous research (e.g. Klimiuk et al., 2021 & Jiang & Wilson, 2018). Based on the Appraisal Theory, we selected emotions that describe the emotional state of individuals (affective language) (G. H. Tran & Ngo, 2018). In both of the commenting groups (support/disapprove), were negative emotions more prevalent. Unhappiness and dissatisfaction were the most expressed emotions. Even though both groups show the same negative emotions, the target of these comments is different. For example, unhappiness in supportive comments on the tweet about members of parliament was referring to how 'shameful' or 'sad' the situation is that MP's are just showing up in parliament when their budget is being discussed. Disapproving comments which expressed unhappiness were mostly directed towards the veracity of the tweet. Users were unhappy that the author posted this misleading pictures with a false context and asked him to remove the tweet. These results again point towards the fact that content and context matter when it comes to reactions towards political misinformation.

This study bears certain limitations. First, we only considered Twitter as a social media platform where misinformation is being disseminated and only the textual elements of the comments were measured in detail. Our findings can thus not be transferred to other platforms such as Facebook or Instagram. Because of the big majority of the comments being disapproving of misinformation, we suspect that comments on other platforms such as Facebook might paint a different picture. One possible reason for different results might be how associations are made. On Twitter, the connections can be one sided which means that individuals can follow public accounts without the other person's consent. On Facebook, to become friends, both individuals have to approve (Vraga & Bode, 2018). This is an important difference when investigating the individual spread of online misinformation, as previous research has shown that users are more likely to share false or misleading content that comes from a trusted source, such as a friend on Facebook (Buchanan & Benson, 2019). In addition, different audiences are on different platforms (Auxier & Anderson, 2021). For example, more women indicated that they use Facebook rather than men and on Twitter the opposite is the case. The average U.S.

Twitter user is further a lot younger than the average Facebook user and platforms also differ from each other regarding partisanship. According to the Pew Research Center, Democrats are generally using Twitter more often than Republicans, whereas on Facebook this difference is not present (Auxier & Anderson, 2021; Vogels et al., 2021). Hence, we encourage scholars to look into other platforms and to also consider other forms of content, for instance, visual content on Instagram or private messages on WhatsApp. An in-depth analysis of visual content would lead to a better understanding how individuals express themselves through memes or videos, since this is important information which was now not considered. Connected to that, the generalizability of our results should be reviewed. Based on our findings, we can only make claims about three specific types of misinformation in one political context. Although these types of misinformation are present on many topics in a wide range of countries, we cannot simply transfer our findings to other false and misleading content on Twitter. Non-political misinformation, for instance related to COVID-19 , might trigger other reactions in the comment section. Possibly with less incivility and less polarization. Therefore, we urge scholars to expand these results by considering different issues or actors, in other contexts.

Second, we chose the UK as a case and are fully aware that other national contexts are of great importance when researching how people react to political misinformation. Different political settings, different media systems, and different economic environments possibly lead to different types of political misinformation and to different individual reactions to it. Future research should therefore compare multiple national contexts to see whether our findings hold.

Third, one might ask how this commenting behavior differs from general online behavior when consuming political information? Could these patterns not be considered as normal online behavior? To some extent, we believe, that we did observe general online behavior. Individuals were eager to engage with the content and with each other and eager to support or disapprove of the statements made. However, we find a significant number of individuals who specifically express that they believe that the tweets contain false or misleading information. Thus, it is not just about expressing one's political views. We recommend that future studies test the direct differences between ordinary online engagement with political information and the engagement with misinformation in an experimental setting. Additionally, the impact of the sender of misinformation should be taken into account. For instance, how would individuals react to known misinformation when it is disseminated by, for example, a politician with whom they share their political views?

Finally, methodologically, we selected a manual content analysis as the applied method. We argue that a manual content analysis has the advantage to thoroughly investigate the contents of the comment section linked to political misinformation. Automated content analysis with a dictionary approach

could analyze bigger amounts of data and could lead to a more stable outcome. However, automated and machine learning approaches are still somewhat limited when it comes to detecting certain forms of speech like sarcasm or irony (Sykora et al., 2020). We encourage future research to combine the benefits from automated and manual content analyses to improve measuring how individuals engage with online misinformation.

Despite these limitations, we identified patterns of the different ways individuals react to political misinformation in the comment section. The following study descriptively measures different user comments with a focus on supportive and disapproving comments as a reaction to misleading messages. Our findings are a first attempt to uncover what messages these types comments convey. The results point towards the fact different types of political misinformation can lead to different types of reactions in the comment section. For instance, the misleading information spread by Nigel Farage lead to more author disapproval than the completely fake tweet involving Nadine Dorries. Or that the pictures taken out of context of the members of the House of Commons triggered the most comments, which added to or reiterated the statement of the tweet. Because this is a first attempt, we urge future research to develop a typology of user reactions on misinformation. On a positive note, we do see that the majority of the twitter users is disagreeing with and even trying to fight political misinformation. However, the dark side of our results show that an overwhelming amount of user comments contain nasty language and thus do not foster a respectful conversation. Unfortunately, if correcting misinformation is done in an uncivil way, it might backfire because uncivil correcting comments can lead to an unfavorable evaluation of the commenter or even the comment itself (Masullo & Kim, 2021). On the contrary, incivility can also be seen as simply being part of a polarized and heated debate in the comment section without any negative consequences. After all, as overwhelming as the spread of online misinformation may seem, our findings point towards alert citizens, who are willing to call out authors and misleading content.

CHAPTER FIVE

Who Approves, Disapproves or Ignores Online Misinformation?

A Study on Individual Characteristics in Connection with Supporting and Renouncing Online Misinformation

This study searches to understand who would engage with online misinformation to either signal support or disagreement and what role individual characteristics such as triggered emotions and dark personality traits play. It compares three groups of individuals based on their engagement with misinformation (no interaction, signaling agreement, signaling disagreement) and three different issues (climate change protest, immigration, COVID-19). Our findings show that both, individuals who agree and disagree with misinformation, experienced more negative emotions. Moreover, we find striking similarities between individuals who agree and disagree with misinforming content. They are predominantly male, lower educated and more interested in politics. However, social media users who generally agree with misinformation showed stronger manifestations of psychopathy, narcissism, and Machiavellianism, indicating that these individuals have a darker personality. We conclude that there is not a clear profile of individuals who would renounce misinformation but rather indications for a dynamic and adaptive social media user.

Keywords: misinformation, counteracting, engagement, social media

This chapter has been submitted to 'New Media & Society' and is at the time of writing awaiting reviewers' scores. An earlier version of this single authored paper has been presented as 'Who Counters Online Misinformation? A Study on Individual Characteristics in Connection with Renouncing Online Misinformation' at the 2022 ICA conference.

Introduction

In today's high-choice political information environment, social media platforms have become a central communication channel where citizens get their information from (e.g. Newman et al., 2021; Van Aelst et al., 2017). The platforms enable individuals to directly engage with (news) content. However, social media is nowadays also being utilized to manipulate and to shape public opinion (Bradshaw & Howard, 2017; McGregor, 2020). As a consequence an abundance of false and misleading information circulates on social media making it difficult for individuals to judge what is true and what is false. In the following study we use the term misinformation to refer to false and misleading messages without the intention to cause harm because we do not measure the intentionality of spreading and creating harmful content (referred as disinformation) (Wardle & Derakhshan, 2018).

In the most recent Digital News Report (2021), more than four out of ten people (43%) indicated that they have been exposed to misinformation and 58% are concerned about misleading messages in their information environment (Newman et al., 2021). This concern led U.S. citizens to increasingly wish for governmental support in restricting online misinformation (Pew Research Center, 2021). Even if they would lose certain levels of access or freedom to publish content, almost half of the surveyed U.S. adults want the government to take action against false information online (Pew Research Center, 2021). It becomes apparent that citizens are worried about their information environments on social media.

There are many different ways how individuals can deal with misinformation on social media and why they even engage with it in the first place. For instance, they can ignore it, try to counter-act or share it further (Coluccia, 2020). Research has found two prominent answers to what social media users do, when they encounter misleading messages: On the one hand, individuals would try to counter it (e.g. flagging the post) and on the other hand, the majority indicated that they would ignore the content (Tandoc et al., 2020). Individuals who said they would go against misinformation, did so because of the ascribed personal issue relevance or because of the personal relationship the person has with the user who shared misinformation (e.g. a close friend) (Tandoc et al., 2020). Why individuals would engage with misinformation at all has many different reasons too. For example, Talwar and colleagues (2019) found that misinformation sharing on social media is connected to the third-person effect, that individuals share false and misleading content out of social media fatigue, fear of missing out, and online trust. Van Bavel et al. (2021) identified multiple risk factors, which influence the dissemination of misinformation. According to the authors, people are willing to spread misinformation to display their political identity, for partisan gain, and to promote hostility towards the partisan out-group. Singaporean students engaged with misinformation because they perceived the messages as

entertaining, and found that misinformation dissemination is connected to a socializing aspect (e.g. interacting with peers) and self-expression (e.g., expressing one's opinion by sharing) (Chen et al., 2015).

This has raised the question whether individuals rather resist online misinformation or if they give in to it. Identifying and understanding which individuals disagree or agree with false or misleading content is crucial to mitigate the spread of misinformation on social media. Hence, we ask the following research question:

RQ: Which overriding factors can explain why individuals would support or go against online misinformation?

Research is still scarce on this issue. The aim of this study is therefore to better understand who comments on or shares misleading messages to go against it and who does it to show support or appreciation for the content.

This study relies on extensive survey data from six Western democracies (Belgium, France, Germany, Switzerland, the United Kingdom and the United States). We aim to answer the research question by comparing the profiles of people that would engage with misinforming social media posts to signal disagreement with two other groups: people who indicated they would disseminate misinforming posts out of agreement, and the group of people who indicated they would in no way engage with misinformation. This will allow us to test to what extent people, who react to misinformation for opposite reasons, also differ from each other in terms of socio-demographic background, emotional reactions towards false or misleading messages, personality traits, and political orientation. Based on this, we will be able to understand the more general individual drivers that lie behind and explain the opposing ways of engaging with online misinformation.

Misinformation Engagement

From previous research we know that people who engage[17] with online misinformation share certain overriding characteristics. Men have been found to be more likely to engage with misinformation than women (Buchanan, 2020b; Chadwick & Vaccari, 2019; Grinberg et al., 2019). Other findings suggest that, education plays an important role in the dissemination of misinformation. Lower education has been linked to the agreement with misinforming content (Pickles et al., 2020). Next to sociodemographic variables, are also other factors of high importance when it comes to

[17] We understand misinformation engagement as any reaction social media users can have towards false or misleading content. This includes liking, sharing or commenting, which consequently leads to the spread of misinformation

misinformation sharing. Past studies have found that conservatives were more likely to disseminate misleading information and were less likely to engage with fact checking messages, suggesting that political ideology matters (Guess et al., 2019; Osmundsen et al., 2020; Shin & Thorson, 2017).

Decreasing trust in news media has manifested itself over the past years in many different countries (Newman et al., 2017). Connected to that, previous research suggests a link between lower trust in mainstream media and the exposure to misinformation (Ognyanova et al., 2020). Looking at the trust in social media, research shows that users are more inclined to trust and spread content posted by other users and perceive the information to be more accurate (Anspach & Carlson, 2020). Furthermore, once social media users trust a fabricated news item they will not try to verify its content and in consequence share it with large audiences (Kalsnes, 2018).

Next to trust, emotions play also a crucial role in how individuals process misinformation. For instance, Han et al. (2020) observed that anger is supporting the dissemination of COVID-19 related misinformation. Freiling et al. (2021) further linked the feeling of anxiety to being more susceptible to spread both correct information and misinformation about COVID-19. It becomes apparent that emotions such as worry, anger, and anxiety triggered by false and misleading social media posts could be a valid predictor for the different reasons why individuals would engage with misinformation. An additional overriding factor which influences the engagement with misinformation are negative personality traits, more specifically the Dark Triad of personality traits (Hughes & Machan, 2021). The triad consists of three elements: psychopathy, narcissism, and Machiavellianism. Psychopathy describes personalities with a lack of empathy, remorse, and anxiety as well as thrill-seeking behavior. It contains two main elements: deficits in affect and self-control. Narcissism is defined by feelings of grandiosity, dominance, and superiority. Machiavellianism is described by the tendency to manipulate other people in a strategic and calculating manner. Machiavellians are, in contrast to psychopaths, concerned about their reputation and tend to plan ahead (Jones & Paulhus, 2014; Paulhus & Williams, 2002). Machiavellians and individuals with narcissistic personality traits have been observed to be less skeptical of conspiracy theories on COVID-19 and by disseminating conspiracy theories on social media, narcissists obtain the attention they strive for (Ahadzadeh et al., 2021). Tang et al. (2022) found that the Dark Triad is linked to what the authors call 'social media disorder' – the disordered (addictive) use of social media ('social media disorder', SMD) and its harms on physical and psychological health (p.1).

We believe that previous research has mostly explored individual factors behind the spread of online misinformation out of supporting reasons. However, we know next to nothing about the individuals who *disagree* with it by, for instance, commenting on the content to warn other users. Or about the

big silent majority who would simply ignore false and misleading messages. The following study shines a light on these individuals.

Misinformation and Counteracting Efforts

Together with the spread of online misinformation also the research on debunking and misinformation correction has been flourishing (e.g. Bode & Vraga, 2018; Chan et al., 2017; Hunt et al., 2020; Vraga et al., 2020). There are many different ways for counteracting misinformation, for instance on a policy level, holding social media platforms accountable for harmful content by implementing specific laws and fining the platforms if they violate them (Alemanno, 2018). Or one of the most prominent interventions, flagging or attaching warnings to possibly false and misleading messages on social media (Pennycook et al., 2020). One strand of research suggests that if people are warned about the falseness of information, they would be less likely to believe it and misinformation is consequently being corrected (e.g. Chan et al., 2017; Lewandowsky et al., 2012). Related to that, three debunking possibilities have been determined by previous research: (1) alarming other users about misinformation at the moment of the first exposure; (2) repeating the warning, and (3) offer compelling corrections which are coherent and easy stories (Lewandowsky et al., 2012). However, several studies found that debunking might even increase the belief in misinformation (e.g. Berinsky, 2017; Nyhan & Reifler, 2010). Individuals might be resistant to these warnings, when they go against their political beliefs or ideology (Flynn et al., 2017; Garrett & Weeks, 2013). Thus research seems to be still somewhat inconclusive regarding debunking measures.

Even though interventions by policy makers and flagging false content by social media companies are relevant steps in the fight against misinformation, we believe, focusing on the actions of ordinary individuals against online misinformation is essential. As individuals warning other individuals about false or misleading social media content is a central piece of the puzzle when combatting misinformation (Bode, 2020).

Individual-based correction on social media platforms can happen through users commenting on misinforming posts pointing out that this content is incorrect (Bode & Vraga, 2018; Tandoc et al., 2020). Individuals have been observed correcting or adapting their attitudes after seeing other users being corrected (Vraga & Bode, 2017). This is especially effective if a number of users present credible corrective sources (Vraga & Bode, 2018). Furthermore, three factors have been identified that help to explain why individuals would correct misinformation: First, individuals would counter misinformation if they have a personal relationship with the user who posted the false content. Second, if the issue of the false or misleading message is personally relevant to them, and third, the user's personal efficacy (Tandoc et al., 2020).

Connected to these three factors, Arif et al. (2017, p. 160) identified through interviews three rumor correcting actions: “correcting oneself, correcting the information space, and correcting another person (or organization)”. Correcting oneself includes, for instance, posting a correction of a false post one has posted beforehand. To correct the information space, interviewees indicated they wanted to keep the information on the whole platform as accurate as possible. So an individual mistake such as posting a misinforming post was not their main concern but rather the greater good of the information quality of the social media platform itself. Lastly, correcting another user by directly tagging them was the third action against misleading content (Arif et al., 2017). In line with these results, Borah et al. (2021) found that calling out users who spread misinformation was an essential strategy to combat false and misleading content on social media.

But who are the people that correct other social media users? Research is still scarce on this issue. Chadwick and Vaccari (2019) found that individuals who would call out or correct others were mostly male, younger, higher educated and more interested in politics. Other results suggest that rather older people are willing to correct others and themselves (Koo et al., 2021). This was also the case for individuals who perceived their personal exposure to misinformation as higher (Koo et al., 2021). Related to that, Bode and Vraga (2021) found that older individuals value it more than younger people if social media users correct others and call out misinformation.

Even though calling someone out seems to be an often mentioned action by individuals, some results suggest that only a fraction of users correct others. In the UK only 8.5% of the social media users called someone out for sharing fabricated news (Chadwick & Vaccari, 2019). It becomes apparent that individuals also take a passive role in the fight against misinformation by, for example, simply ignoring corrective approaches (Borah et al., 2021; Tandoc et al., 2020; Young, 2021). Interviews with young adults gave insights that they ignore misinformation out of the feeling of getting angry or frustrated (Borah et al., 2021). Tandoc et al. (2020) observed that almost three quarters (73%) of their respondents would ignore false and misleading posts and not go against them by posting a comment or directly message the user who posted the misinforming message. Supporting the results mentioned above, women were more likely to ignore the post in comparison to male social media users (Tandoc et al., 2020). One possible reason for ignoring possibly misleading messages is the perceived accuracy of the content. It is challenging to assess accuracy of misinformation (Pennycook et al., 2020). So, individuals might be reluctant to engage with certain suspicious social media content and just avoid it.

It becomes clear that individuals interact in different, even opposing ways, with false or misleading content. However, little is known about who would go against it. In the following study, we are therefore specifically interested in individuals who would share or comment on misinforming social

media posts to signal disagreement or warn others. Is there a certain profile that is shared by individuals who go against misinformation and how does it differ from people who agree with this content or do not react to it at all?

Methods

Design. We conducted representative surveys in six Western democracies (Switzerland, Belgium, France, Germany, the United Kingdom and the United States [18]) to investigate the different engagement with misinformation connected to three topics (immigration, climate change protests and COVID-19). We argue that these three issues are well suited for studying misinformation, because they had been connected to false and misleading messages circulating widely on social media in the months and weeks prior to the survey (Mimikama, 2019; Petersen et al., 2018; Poynter, 2020).

Procedure. A polling company (Respondi) recruited representative samples of social media users in all six countries based on country-specific census data (see Appendix 3). The country specific quotas were provided by the company. Based on the interest of this study, the sample only consisted of social media users (usage of at least once a month). The data were collected during four weeks in April and May 2020, and the participants were given an incentive. After giving their informed consent, respondents completed the first part of the survey, which included standard demographics and news consumption. In a second step, each participant was exposed to three fictional social media posts containing false and misleading claims about climate change protests, immigration, and COVID-19 (see Appendix 3.1). After being exposed to the posts, the participants were forwarded to the second part of the questionnaire, which included measures for the dependent variable. At the end of the survey, the participants were debriefed and informed about the aim of the study and the manipulation of the shown social media posts.

Sample. After removing straightliners and other outliers based on response time and quality fail questions, we secured a sample of 7,009 respondents (quotas for age: *M* = 43.87, *SD* = 14.69; gender: female = 51.7%, male = 48.3%, education: lower = 27.7%, moderate = 40.1%, higher = 32.1%).

Dependent variable and vignettes. Before they were exposed to the fabricated posts, the participants were told to imagine that the posts would appear on their social media newsfeed and to read them carefully. The created vignettes looked like news articles posted on Facebook by a fictional source (news.com). To enhance the study's external validity, we used claims that had circulated on social

[18] The data used in this study are linked to a large-scale research project. The countries were selected based on research on resilience to disinformation (Humprecht et al., 2020) to test different contextual factors that create opportunity structures to engage with online misinformation. The following study does not focus on these factors or on specific country differences, rather, we aimed to find generic factors that matter across (Western) countries.

media and were discussed on various fact checking websites. We adjusted them to match the purposes of this study. The COVID-19 post had the following statement: “The coronavirus might be a bioweapon from China – Report claims the Chinese government intended to use COVID-19 to damage the economy of other countries”. The claim regarding immigration read as follows: “The news media withholds information about dangerous immigrants – Well known national news media often deliberately don’t report about crimes of illegal immigrants against innocent citizens”. The post regarding climate change had the following message: “Green protesters leave public places scattered with their garbage – After a recent climate change demonstration, protesters leave their signs and rubbish behind in public parks”.

All three statements blamed actors for things that are failing in society, which is very common in online misinformation (AFP et al., 2020). After reading each post they were asked how likely it is that they would a) like the post, b) comment on the post, and c) share the post on a 7-point scale (1= very unlikely, 7= very likely).

In our sample between 8% and 27% of the respondents indicated that they would likely engage with the presented post. Our data (see Table 2 with distributions in Appendix 3.2) shows that, across countries, the post on climate change protests led to most engagement, followed by the posts related to immigration and the coronavirus (climate change protest: *n* = 2318, immigration: *n* = 1559, COVID-19: *n* = 1947). Swiss citizens were the least likely to engage with the posts, followed by Germany and the UK. Belgium takes a middle position regarding the willingness to interact with online misinformation. Participants in the U.S. and France were the most inclined to engage with the misleading social media posts. Across countries and types of engagement, about 12.1% of the participants signaled they would react to all three of the manipulated posts, 13.6 % stated they would engage with two of the three post and 19.4% would only engage with one of the three posts. This already indicates that there is a substantial overlap between the interactions with the posts on three different issues.

If respondents scored a 4 or higher, they were subsequently asked about their reasons for that type of engagement. Participants were asked the following items to tap into their motivations to engage (in a specific way) with the post: “I want to signal that I agree with the content of the post” and “I want to signal that I disagree with the content of the post”. The motivations were measured on a 7-point Likert scale (1 = completely disagree, 7 = completely agree). If the participants scored lower than 4 when being asked if they would interact with the posts, we considered them as being not willing to engage with the content and consequently formed the third group of individuals we are interested in. Since we are interested in individuals who engage with misinformation because they signal disagreement,

we excluded the reaction *liking* because this reaction cannot be used in a comprehensive way to signal disagreement (climate change protest, n = 1201; immigration, n = 948; COVID-19, n = 1125). In case respondents indicated that they would share a post to signal agreement and to comment on the same post to signal disagreement, these inconsistent answers were also excluded from the analysis and (n = 51).

Independent Variables. With our independent variables we want to capture individual features of the three different groups of individuals regarding their engagement with misinformation online. Political orientation, political interest and trust are crucial factors to get an idea about who engages with misinformation out of which reasons. They form an important part of the individual profile.

Political Orientation. The participants had to indicate where they would place themselves on an 11-point ideological scale (0 = extreme left, 10 = extreme right), M = 6.04, SD = .03.

Political Interest. Political interest was measured on a 7-point scale (1 = not interested at all, 7 = extremely interested). Participants were asked how, generally speaking, interested they are in politics, M = 4.20, SD = .02.

Trust Variables. To measure the perceived trust in news, we asked the participants to indicate if they thought they could trust the news most of the time (1 = strongly disagree, 7 = strongly agree), M = 4.23, SD = .02. The same question was asked to assess the trust in news on social media (1 = strongly disagree, 7 = strongly agree), M = 3.20, SD = .02.

Emotions. Negative emotions have been connected to the spread of misinformation (see e.g. Han et al., 2020). It is essential to see how these emotions are displayed throughout the different groups of individuals under study to paint a detailed picture of them. The triggered negative emotions were measured each time after participants saw the posts. They were asked on a seven point scale (1 = completely disagree, 7 = completely agree) how they felt when reading the post. This is the only independent variable which is directly connected to the stimuli. Based on previous literature (Freiling et al., 2021; Han et al., 2020), we selected three negative emotions: worry (climate change protest: M = 3.72, SD = .02; immigration: M = 3.55, SD = .03; COVID-19: M = 3.88, SD = .03), anger (climate change protest: M = 4.48, SD = .02; immigration: M = 4.13, SD = .02; COVID-19: M = 4.27, SD = .03), and anxiety (climate change protest: M = 2.35, SD = .02; immigration: M = 2.75, SD = .02; COVID-19: M = 3.10, SD = .02).

Personality Traits. Next to emotions, are personality traits another crucial variable to test why individuals with certain dark traits engage with misinformation. Based on the literature on the Dark Triad of personality traits (see e.g., Jonason & Webster, 2010; Jones & Paulhus, 2014; Paulhus &

Williams, 2002), we test its three components: narcissism, psychopathy, and Machiavellianism. The constructs were measured on 7-point Likert scales, based on the studies by Jonason and Webster (2010) and Jones and Paulhus (2014). Narcissism was measured with four statements (e.g., "I tend to want others to admire me" and "I tend to want others to pay attention to me"), Cronbach's α = .88, M = 2.96, SD = 1.55. Psychopathy was measured through three items (e.g., "I tend to lack remorse" and "I tend to be callous or insensitive"), Cronbach's α = .67, M = 2.89, SD = 1.39. The Machiavellianism scale consisted of four items (e.g., "Make sure your plans benefit you, not others" and "Most people can be manipulated"), Cronbach's α = .68, M = 3.70, SD = 1.32.

Concern about misinformation. To measure the perceived individual concern about misinformation, we asked the participants if they think misinformation in general is a problem (1 = no problem, 7 = serious problem), M = 5.29, SD = .02. This was asked independently from the stimuli to capture general attitudes towards misinformation. We included this variable to be able to see whether individuals who differ in their concern about misinformation also engage differently with it.

Judging Information Veracity. We asked the participants how easy or hard they find it to distinguish between true and false when they are following the news (1 = I find it very hard to judge, 7 = I find it easy to judge), M = 4.36, SD = .02. We believe that this variable is of high importance when it comes to researching the engagement with misinformation. The (perceived) ability to judge the veracity of a message also determines whether individuals would engage with it or not.

Reaction towards Misinformation. Based on previous research (Borah et al., 2021; Tandoc et al., 2020) we selected two reactions towards misinformation. The participants were asked how they react when they come across information they suspect to be false: *ignore the information* (1 = never, 7 = very often), M = 5.38, SD = .02 and *point out to others that information might be false information* (1 = never, 7 = very often), M = 4.40, SD = .02. We are aware that these variables are somewhat connected to our dependent variable (engagement with misinformation). However, these reactions towards misinformation are not directly connected to our stimuli, unlike our dependent variable. These two reactions here capture the general individual coping with misinformation and function as a robustness check. It enables us to test whether individuals behave like they say they generally would versus if they are confronted with actual misinforming content.

Results

To first give an overview of the three different groups under study, we explore descriptively the different categories of our dependent variable per country (Table 1). The respondents who indicate they would engage with the misleading posts to show disagreement are on average the smallest group per country and issue, between 3% and 13%. Interestingly, this group of people is larger in countries

where also more people would interact with the posts out of agreement (U.S., France). This is especially the case for the climate change protest and COVID-19 post, indicating that this behavior is more prominent in countries where misinformation is more present. Citizens in Switzerland were overall the most likely to show no reaction towards the posts, followed by citizens from the UK, Germany and Belgium. Looking at the different issues, we see that, across countries, the post that contained the false claim about the climate change protest resulted in the highest willingness to engage with it out of agree- and disagreement, followed by the post on the coronavirus (climate change protest: *n* = 649/486, COVID-19: *n* = 595/414, immigration: *n* = 548/281). In the further analysis, we do not focus on the country differences but added country dummies as controls. They largely reflect the variation in Table 1, with French and U.S. citizens being overall significantly more inclined to engage with misinformation, be it to signal agreement or disagreement.

Table 1. Distributions of respondents per group and issue in percentages

	No interaction			React to agree			React to disagree		
	Climate Change	Immig ration	Covid-19	Climate Change	Immigra tion	Covid-19	Climate Change	Immig ration	Covid-19
BE	77.9%	83.8%	82.7%	13.3%	10.0%	10.0%	8.8%	6.1%	7.4%
CH	84.4%	90.0%	88.1%	9.3%	6.1%	6.9%	6.3%	3.9%	5.0%
DE	81.4%	83.2%	82.8%	7.6%	9.6%	8.4%	11.1%	7.2%	8.8%
UK	82.8%	89.2%	84.5%	11.1%	7.9%	9.9%	6.1%	3.0%	5.7%
FR	71.7%	85.0%	77.5%	15.0%	10.7%	13.6%	13.3%	4.3%	8.9%
US	73.0%	80.0%	74.6%	17.2%	14.3%	15.9%	9.8%	5.7%	9.6%
N	4251	4941	4600	649	548	595	486	281	414

Note: Percentages in the table represent people who answered 4-7 on a seven-point scale that they would share or comment on the post to signal agreement and disagreement. And it includes individuals who scored lower than 4 on the misinformation engagement scale (no interaction group).

We aim to detect different individual profiles of individuals who engage with misinformation out of agree- or disagreement or who would not interact with misinformation. We performed multivariate multinominal logistic regressions to account for the categories of the dependent variable (individuals who signaled agreement, disagreement or no interaction). Table 2 shows the different components of the individual characteristics predicting the different motivations behind the spread of online misinformation. The reference category was the 'no interaction' group.

We find remarkable similarities looking at the socio-demographic variables. They show that individuals who would engage with the misleading social media posts to either signal agreement or disagreement

are both predominantly male and have a lower education than the individuals who would not interact with the misinforming posts. Looking at the age of the individuals, we find that individuals who agree with the climate change protest and immigration post were older than the individuals who would not interact with the posts. The disagreement group shows mixed results. Individuals of that group were older when engaging with the climate change protest post and younger when they disagreed with the immigration post. This suggests that who goes against misinformation is driven by a specific topic. Individuals in the agreement group and disagreement group furthermore have less trust in news and more trust in news they see on social media than individuals who indicated they would not interact with the posts. Social media users who agreed with the posts have, across all issues, rather a right-leaning ideology than individuals from the other two groups. Individuals from the agreement and disagreement group were additionally more interested in politics than people who were not willing to engage with the three posts.

Model 1, 2, and 3 show that people who react to misinformation for opposite reasons both have stronger emotions towards the post than people that do not react. This is, across issues, especially the case for the emotions anger and concern (worry). A main difference between the agreement and disagreement group is that the disagreement group is much less driven by fear. In general, were the negative emotions much more triggered in individuals who signaled agreement with the posts.

Next, focusing on dark personality traits, we find that people who agree with the post on climate change protests are more narcissistic than individuals who would not interact with the post (B = .11, SE = .04, $p < .05$). Individuals who were signaling agreement with the immigration post, showed stronger psychopathic and Machiavellian traits than the no interaction group (psychopathy: B = .10, SE = .05, $p < .05$; Machiavellianism: B = .14, SE = .05, $p < .01$). Looking at the COVID-19 issue, we see that the agreement group has a stronger manifestation of the Machiavellian personality trait than the group of individuals who would not interact with the post on the coronavirus (B = .13, SE = .05, $p < .01$).

Next, we bring in the general reactions or attitudes towards misinformation. As explained earlier, these reactions were not asked in relation to the presented posts, but rather towards misinformation in general. We could not find evidence that being concerned about misinformation and the ability to judge whether a message is true or false plays a role when assessing the three different groups of individuals. However, we observe that, across issues, the agreement and disagreement group are more likely, when confronted with false or misleading messages 'to warn others' in comparison to individuals who would not interact with the posts. Although this effect is always stronger for people who would engage with the posts to signal disagreement, it remains striking that people who are willing to engage

with the presented misleading posts out of agreement have a similar need to point out content they consider as misinformation. As a robustness check we tested whether 'ignoring the information' would be a good strategy when confronted with false information. Unsurprisingly, we find that the agreement and disagreement group are less inclined to ignore possibly false and misleading messages than the people who indicated they would not interact with the three posts.

Table 2. Multivariate multinominal regressions predicting the motives for (not) engaging with online misinformation

	Model 1 Climate Change		**Model 2 Immigration**		**Model 3 COVID-19**	
	Agree	*Disagree*	*Agree*	*Disagree*	*Agree*	*Disagree*
Sex	.23*	.14	.47**	.33*	.45**	.33**
	(.10)	(.11)	(.11)	(.14)	(.11)	(.12)
Age	.01**	.01*	.01**	-.01*	-.00	-.00
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Education	-.29**	-.50**	-.31**	-.24**	-.40**	-.21**
	(.06)	(.07)	(.08)	(.09)	(.07)	(.07)
Trust in news	-.07	-.13**	-.24**	-.05	-.13**	-.01
	(.04)	(.04)	(.04)	(.05)	(.04)	(.04)
Trust in social media	.24**	.22**	.29**	.17**	.28**	.13**
	(.04)	(.04)	(.04)	(.05)	(.04)	(.04)
Political orientation	.06**	-03	.20**	-.08**	.10**	-.03
	(.02)	(.02)	(.02)	(.03)	(.02)	(.04)
Political interest	.17**	.14**	.15**	.18**	.03	.12**
	(.03)	(.03)	(.03)	(.04)	(.03)	(.02)
Emotions						
Angry	.37**	.46**	.41**	.36**	.14**	.27**
	(.03)	(.04)	(.03)	(.04)	(.03)	(.03)
Worried	.13**	.10**	.25**	.13**	.37**	.05
	(.03)	(.03)	(.03)	(.04)	(.04)	(.03)
Afraid	-.03	-.10	.08**	.01	.18**	.03
	(.03)	(.03)	(.03)	(.04)	(.03)	(.03)
Personality traits						
Psychopathy	.04	.01	.10*	.05	.00	.03
	(.04)	(.05)	(.05)	(06)	(.05)	(.05)
Narcissism	.11*	.05	-.04	.05	.02	.08
	(.04)	(.04)	(.04)	(.05)	(.04)	(.04)
Machiavellianism	.07	.02	.14**	.00	.13**	-.04
	(.04)	(.05)	(.05)	(.06)	(.05)	(.05)

Table 2. Multivariate multinominal regressions predicting the motives for (not) engaging with online misinformation – part II

Misinformation variables						
Concern over misinformation	.03	-.00	-.04	.01	-.06	.07
	(.04)	(.04)	(.04)	(.05)	(.04)	(.04)
Judging message veracity	-.00	.03	-.00	.06	-.03	.03
	(.03)	(.04)	(.04)	(.05)	(.04)	(.04)
Warn others	.18**	.25**	.15**	.32**	.17**	.28**
	(.03)	(.03)	(.03)	(.04)	(.03)	(.03)
Ignore the information	-.13**	-.10**	-.11**	-.17**	-.11**	-.17**
	(.02)	(.03)	(.03)	(.04)	(.03)	(.03)
Countries						
CH	-.29	-.23	-.18	-.41	.04	-.22
	(.17)	(.20)	(.02)	(.23)	(.20)	(.21)
DE	-.47*	.29	.18	.02	.21	.20
	(.19)	(.20)	(.02)	(.22)	(.20)	(.20)
UK	.08	-.28	.15	-.62*	.51**	-.19
	(.17)	(.20)	(.20)	(.25)	(.19)	(.20)
FR	.34*	.59**	.15	-.13	.59**	.38
	(.16)	(.19)	(.19)	(.23)	(.18)	(.19)
US	.47**	.20	.77**	-.12	.94**	.23
	(.17)	(.20)	(.20)	(.24)	(.19)	(.20)
Constant	-1.34	4.05*	-2.78	-1.45	1.38	-1.47
	(1.50)	(1.71)	(1.72_	(2.08)	(1.62)	(1.72)
Chi^2 (df=38)	**1191.32 ****		**1416.04****		**1316.63****	
Pseudo R^2	**.17**		**.26**		**.21**	

Notes: Reference category is the no interaction group and Belgium for the country variables.

Standard errors are in parentheses. **p<.01,*p<.05

Discussion

The aim of this study was to identify overriding characteristics of three different groups of individuals regarding their interactions with online misinformation. Previous research focused mainly on individuals who would support misinforming content and found that these individuals are male, lower educated and mostly politically right-leaning (Buchanan, 2020b; Chadwick & Vaccari, 2019; Grinberg et al., 2019; Pickles et al., 2020).

However, little is known about the group of individuals who interacts with online misinformation to show disagreement or about individuals who would in no case interact with misinformation and the role of triggered emotions and dark personality traits connected to it. First, we find, consistent with Tandoc et al. (2020), that most people do not interact with online misinformation at all. Taking a closer look at the individual factors, we observe that individuals who would agree or disagree with the misleading posts experienced more anger and concern than individuals who would not interact with any of the content. Social media users who generally agree with misinforming content showed stronger manifestations of psychopathy, narcissism and Machiavellianism, indicating that these individuals have darker personalities than individuals who would not interact with this content. Surprisingly, the three groups did not differ in their concern about misinformation and perceived ability to judge whether a message is true or false. Individuals who disagree with the misinforming posts were not more worried about misinformation and did not think they were better at judging whether information is correct than the no interaction group. Both of the groups, agreement and disagreement, were more willing to point out to others that certain content might be false compared to the group of individuals who would not interact with the posts. We believe that this the case because, even though regarding the shown posts the groups might differ in their motives (agreeing, disagreeing), these individuals are generally more active and aware social media users than the group that would not interact with any of the shown content. Another reason could be that these individuals have very different or opposing views about what kind of content is false.

A surprising finding of the analysis is that the agreement and disagreement groups are fairly similar when it comes to socio-demographics. Even though, social media users engage with the posts out of opposite reasons, they are all predominantly male, lower educated and more interested in politics. It seems that higher educated individuals do not go against online misinformation. One reason for this behavior could be that this group does not want to spread the content any further, making it possible for the message to reach bigger audiences. It may be that these individuals rather engage in personal conversation and get loud about the false information in private messages.

Naturally, this study bares several limitations. First, we only selected two groups of possible forms of engagement with misinformation and disregarded others. Future research, however, should consider further motivations for engaging with misinformation such as expressing one's opinion, making the issue visible or informing users and look for overriding individual characteristics. Second, we find striking similarities between the agreement and disagreement group, which we cannot fully explain with our analysis. This interesting result should be addressed as a further avenue of future research. Finally, since we presented respondents with multiple issues we did not ask for every claim whether they believed it, to avoid priming respondents about the false nature of the message (Klar et al., 2020). However, we acknowledge that the belief in misinformation is a valid predictor for researching individual profiles behind the different ways of engaging with misinformation. Therefore, future research which focuses on single issues should delve deeper into the importance of the perceived accuracy of social media content.

Despite these limitations, this study also has a number of implications for policymakers and other relevant social institutions. The results showed that mostly male, lower educated and older individuals would interact with online misinformation to signal agreement. This suggests that specific policy responses or efforts to combat online misinformation can be tailored to this specific group of individuals. These results have also raised the question if it makes sense to go against possibly misleading content on social media by sharing or commenting on it or if it should just be ignored. Clearly, more research is needed to establish which strategy is most effective to combat misinformation, but it is clear that together with the rise of this problem also the eagerness of people to go against it increases. This demands for a research agenda that focusses concretely on when and how people should react to misinformation. With our research, we hope to have contributed to a more inclusive understanding of the overriding characteristics of individuals who would engage with misinformation to signal support or disagreement and be able to mitigate and combat the spread in a more targeted manner.

CHAPTER SIX
CONCLUSION

Aim and Setup

With misinformation spreading relentlessly on social media, it is key to know what individuals 'do' with it. Are they actively and consciously contributing to the spread of this type of content? Are they trying to counter misinformation or simply ignoring it? The aim of this dissertation was to shed light on individuals and their relationship with online misinformation. In particular, the goal was (1) to understand who engages with misinformation on social media platforms and (2) to describe how and why individuals react to misinforming content.

While ample research has already focused on certain microlevel-aspects in connection with misinformation on social media, little is known about the interplay of multiple individual factors and the dissemination of misinformation. To uncover this interplay, we first conducted three empirical studies that looked at individual characteristics and people's motivations to spread or counter false and misleading messages. This enabled us not only to characterize why people are willing to spread misinformation, but also under what circumstances misinformation is not disseminated. Knowing which individuals do not spread misinforming content is in that sense of practical utility that interventions could be employed that foster this type of behavior. In a fourth study, this thesis investigated how individuals react to real misinformation by studying the comment section on Twitter, to better understand how people engage with online misinformation. Understanding how individuals react to misinformation can be relevant to identify misinformation (e.g. if the comments suggest that the information is false or manipulated) or to measure what type of misinformation receives the most traction via user engagement and what type polarizes the audience the most. Furthermore, by applying different theoretical approaches in the four empirical chapters, this dissertation added to the pursuit of identifying appropriate theories for the field of misinformation research.

In the remainder of this concluding chapter, the results of the empirical parts of this dissertation are synthesized, the contributions of this thesis are discussed, limitations are identified, and recommendations for future research are highlighted.

Key Findings

In four empirical chapters, using two different methods (survey and content analysis), the relationship between individuals and online misinformation was investigated. In the following paragraphs, I will integrate the findings of the different chapters by presenting three key findings.

1) Individual Characteristics as a Strong Predictor for Misinformation Dissemination

By means of surveys with citizens, my co-authors and I unraveled the drivers behind the dissemination of online misinformation and shed light on the role of political attitudes and individual characteristics

connected to the spread of false and misleading content (chapter 1). First, and in line with previous literature, we show that sociodemographic characteristics are of high significance when investigating who spreads online misinformation. Men, less educated, and right-leaning individuals are more prone to engage with misinforming content, no matter the issue. Going one step further than previous misinformation studies, we considered dark personality traits as another predictor for why people disseminate misinformation. Stronger manifestations of narcissism and psychopathy explained the likeliness to spread false and misleading information online. Furthermore, general social media use is a central predictor when it comes to the engagement with online misinformation, implying that active social media users are more willing to spread problematic content online. Also important, is that attitudinal congruence can explain why people disseminate misinformation. The more congruent the content is with preexisting beliefs, and the higher the personal relevance of an issue, the more likely individuals were to engage with a post on that topic, even if the content might be inaccurate. We conclude, heavy social media users with strong political opinions, darker personalities, and an eagerness to express themselves form a 'loud' minority that contributes significantly to the reach of misinformation.

By looking at who counters, supports, or ignores online misinformation, I not only complement the debunking and misinformation dissemination literature, but also extend it by considering individuals who simply ignore false and misleading information. It is clear that individual characteristics matter as a predictor for these three types of behaviors (chapter 4). I repeatedly find that individuals who generally agree with the misinforming content, display stronger manifestations of dark personality traits. Zooming in on sociodemographic factors, strong similarities between the social media users who agree and disagree with the misinforming posts are shown. Even though individuals interact with the content out of opposite reasons, they are all predominantly male, have a higher interest in politics, and are lower educated. It seems that higher education is not a predictor to go against online misinformation. Even though, the findings of this dissertation display a vast amount of similarities, I do find some small differences between the identical explanatory factors in chapter 1 and chapter 4. First, age matters when explaining who counters, supports, or ignores online misinformation in comparison to chapter 1 where the willingness to disseminate misinformation was measured. Age could not explain the willingness to engage with information (chapter 1). However, in chapter 4, I found an effect of age. Individuals who supported the posts were mostly older. Second, if people disagreed with the misinforming content, political orientation did not matter either or these individuals were rather left- than right-leaning. Finally, individuals who disagreed with misinformation did not display any dark personality traits, whereas individuals who agreed with the misinforming messages and who were willing to disseminate misinformation, had higher levels of narcissism and psychopathy. Furthermore,

both groups, disagreeing and agreeing, are more willing to warn others that certain information might be inaccurate, compared to the group of individuals who would simply ignore the messages. I argue, that even though regarding the shown content the two groups might differ in their opinions, these people are generally more active and aware social media users than the group that would not interact with misinformation.

As already quickly touched upon when describing the importance of other individual characteristics as predictors for engaging with misinformation, ideology and political orientation matter – a lot. Complementing existing research, we identified political orientation as a central individual characteristic to predict the likelihood to disseminate online misinformation. On average, individuals with a right-leaning ideology were more likely to like, share or comment on misinforming social media posts.

Looking at the people's motivations to spread conspiracy theories on social media, the findings suggest the same relationship. A central reason for why individuals would disseminate conspiracy theories is out of conviction. If people are convinced by the message of a conspiracy post, they are in return more likely to support and spread it. Being convinced by a message is strongly linked to one's ideology or political attitudes. Thus, political orientation and issue attitudes can explain why people engage with conspiracy theories, especially when they are convinced by the content of the conspiratory social media post. We again observe that a right-leaning political ideology predicts the likelihood to disseminate conspiratory information out of agreement or conviction. Later on I will elaborate on the extent to which that result might have be driven by the selection of the three presented issues.

This observation is further confirmed. In a similar vein, I found strong similarities in individual characteristics between supporters and opponents of online misinformation. However, political ideology is what seems to make the difference. Because people who signaled agreement with the three presented misinforming posts displayed a rather right-leaning ideology, compared to individuals who disagreed or who were not willing to engage with the content. Thus, ideology appears to determine the direction of the engagement with online misinformation.

In a nutshell, ideology is one of the most important drivers why individuals engage with social media content the way they do. The concern about the veracity of attitudinal congruent content seems to fade into the background. The 'closer' the message to one's own ideology, the more likely are people to engage with social media posts on that topic, even if the content is false or misleading. In sum, this first key finding suggests that individual characteristics such as political orientation, personality traits or social media behavior can predict why individuals would further spread or act against misinformation on social media. This finding implies, that there is an overarching 'profile' of individuals

who can be considered more vulnerable to (un)consciously spreading online misinformation. By spending actively a lot of time on social media, this group of individuals is more likely to be exposed to misinformation. Our findings suggest that participants who frequently use social media and generally like, share, or comment on the posts of friends and family members, and who have higher levels of trust in news on social media were more likely to engage with possibly false or misleading content. Hence, some individuals of this group, which share the same political orientation, gender, level of education, and certain personality traits, are more likely to be affected by false and misleading information.

2) Three Motivations Behind the Spread of Online Misinformation

A second key finding of this dissertation is that people have multiple and different motivations to engage with online misinformation. Complementing the Uses and Gratifications Approach literature and adapting it to misinformation dissemination, my co-authors and I tested three distinct types of motivations (1) conviction, (2) entertainment, and (3) the need to provoke reactions on social media. In chapter 2, focusing on conspiracy theories related to immigration and COVID-19[19], we show that being convinced by a false or misleading message was the main motivation to further spread this type of information. Our survey results indicate that provoking reactions is the second most indicated reason to engage with conspiracy theories online. In order to gain attention and trigger reactions within their social media networks, people were willing to like, share, or comment on a false or misleading posts. Overall, entertainment played only a secondary role in the spread of conspiracy theories. However, it was still mentioned by a considerable amount of individuals that indicated they would engage with misinformation.

Building on the first key finding of this dissertation, the three chosen motivations vary from each other when linked to two sets of individual factors (political variables and psychological traits). Looking at the first set, we find strong evidence that political attitudes and political orientation are connected to spreading conspiracy theories out of conviction. An attitudinally congruent conspiracy post mattered a lot more for disseminating it based on people's conviction than for the other two motivations. In a second step, we show that certain – but not all – psychological traits can explain why individuals would disseminate conspiracy theories out of fun or to provoke reactions on social media. Narcissists and individuals with the need for drama both share conspiratory content out of entertainment reasons and to provoke reactions. People with rather narcissistic characteristics might be more inclined to disseminate conspiracy theories to be in the center of attention and to get appreciation online. For the

[19] Because chapter 2 is about conspiracy theories, only two out of the three posts that we used in the survey were considered. The post about climate change protesters did not fit the definition of conspiracy theories and was therefore excluded from the analysis.

third element of the dark triad of personality traits, Machiavellianism, however, the effect points into the opposite direction as this trait was mostly related to agreement and conviction. To a certain extent this is explicable, as those who score higher on Machiavellianism may be more prone to engage with congruent information out of instrumental and strategic reasons to influence others.

To sum this key finding up, my co-authors and I find that reasons why individuals engage with misinformation are diverse and that the three selected motivations can, to some extent, explain what drives individuals to spread conspiracy theories or misinformation on social media.

3) Individuals React to Misinformation out of Disagreement in Different Ways

Going further than existing literature on commenting on misinformation, studying Twitter comments posted on different types misinformation involving politicians (see chapter 3), the results reveal a number of patterns of how individuals react to misinformation. First, I find that the majority of Twitter users in the UK react to misinformation out of disagreement. This type of comment mainly consisted out of disapproving reactions, for example, calling out the author of the tweet or warning other users about the possible inaccuracy of the post. The second most common disagreeing reaction was expressing author disapproval. Thus, users expressed that the author of the tweet is spreading false information or demanded for the author to take down the tweet. Second, only a minority of comments were of supportive nature. The most common supportive reaction were comments that exemplify, restate and clarify the content. Following this main supportive reaction, users further support the misinforming posts by making general agreeing statements. For instance, they supported the claims made by the authors of the tweets. Looking at both sides of the spectrum of the comments, I observe big differences that can be tied to the three different tweets under study. Depending on who posted the tweet (e.g. the political actor personally vs. a fake account), whether it was known false content, or depending on what the exact claim of the tweet was, the comments vary heavily. Thus, different types of misinformation trigger different reactions.

Finally, I test whether incivility plays a role when commenting on misinforming posts. The majority of the reactions did not contain uncivil language. Nevertheless, depending on the selected tweet, almost 50% of the comments expressed at least one type of incivility. The comments mostly used nasty language (e.g. vulgarities) or were of sarcastic nature. Taken together, I find comments with simply uncivil language and comments that attacked the author or politician involved.

Overall, I aimed to classify the characteristics of disapproving and supportive comments on different types of online misinformation involving politicians. I find that the majority of the comments is attempting to counter the misinforming posts or expresses disagreement with the claims. If comments are supportive of the misinforming content, they mostly restate the statement or add supportive

details. However, I conclude that context matters when investigating user comments connected to misinformation and individuals might react entirely different when confronted with other types of political misinformation.

Contributions and Societal Implications

Misinformation spreads like wildfire on social media, and it is harmful. Through deliberate sharing or simple reactions such as 'liking', the reach of misinforming messages can increase exponentially (Buchanan, 2020b). Misinformation has the power to polarize (online) communities, reduce trust in democratic institutions and can promote incivility in social networks (Council of Europe, n.d.; Masullo & Kim, 2021; van der Linden et al., 2021). It became apparent that this type of harmful content can have real life consequences. For instance, during the COVID-19 pandemic, misinformation regarding the virus was putting the public health at risk. It is therefore vital to investigate how misinformation works and what the overall effects on individuals are. In this paragraph, the results of this dissertation are put into a broader perspective. Hence, what do the findings of this dissertation contribute to the field of misinformation research, and how do they matter for society?

As a key contribution to the field, this dissertation is able to characterize individuals who are willing to spread online misinformation. My co-authors and I advanced previous research by testing existing variables connected to the dissemination of misinformation, but also by linking them to others. We find that these individuals were (1) mostly male, (2) lower educated, (3) rather right-leaning, (4) heavy and active social media users, and (5) show stronger manifestations of narcissism and psychopathy. Identifying and understanding the dynamics between these individual characteristics is a first – but vital – step in developing solutions to target individuals who are less resilient to online misinformation.

This thesis shows that overall, only a minority of people is willing to spread false and misleading messages. The results of our large-scale survey show this, and the user comments on real online misinformation are mostly disapproving the statements made in the misinforming posts or even warn others. This means that, in the countries under study, we are dealing either with social media users who are mostly ignoring misinforming messages or with very alert citizens who are calling out questionable content. However, this does not mean that these individuals could not be affected by it. The results of the content analysis show how individuals react to real misinformation on Twitter and add to the findings of the survey. The majority of the comments disagreed with the misinforming content, and users warned others about the inaccuracy of the statements or questioned their veracity. At first sight, this is positive news. Unfortunately, with online misinformation, this is not enough. Because if only a fraction of individuals is willing to disseminate it – even only by a simple 'like' – it can still reach millions of people. Additionally, users are drawn to shocking, outrageous information, which

leads to false content being disseminated faster and farther than accurate content (Vosoughi et al., 2018). On top of that, social media platforms' algorithms are facilitating the spread of misinformation because attention equals money and algorithms are designed to feed into that logic. Because of the relentless spread of false and misleading information on social media, we have recently observed governments putting growing pressure on social platforms to take responsibility for their actions. However, we are still far away from pervasive policy regulations, such as algorithmic transparency and increased data access to hold platforms more accountable.

Furthermore, we find evidence that this group of individuals mostly spreads misinformation because they are convinced by the message or because they want to provoke reactions within their social networks. It is one thing to share a misinforming post because you believe in it. It is a completely different thing if individuals spread misinformation out of entertainment reasons or to provoke reactions. This insight is highly relevant. Because in the fight against misinformation, it clearly requires a different strategy to reach individuals who spread misinformation for fun or to provoke reactions. Fact checks, for instance, will not be a suitable remedy for this group of individuals. Rather, should warnings be attached to possibly misleading content which inform users about the consequences of disseminating misinformation for entertainment reasons or to gain attention within their network (e.g. how misinformation spreads exponentially just by liking it and how many individuals can be reached by a single click).

This dissertation set out to create a comprehensive understanding of the relationship between individuals and online misinformation. Three different polarized issues, which were salient in all six countries (BE, CH, DE, FR, UK, U.S.) under study were put to the test. The manipulated social media posts with the three topics (climate change protest, COVID-19, immigration) shared some similarities but also differentiated from each other. Polarized issues, like the ones studied in this dissertation, are known to pose easy targets for the producers of mis- and disinformation (Allcott & Gentzkow, 2017; Shin & Thorson, 2017). In addition, polarized topics encourage the dissemination of false and misleading content (Allcott & Gentzkow, 2017; Shin & Thorson, 2017). Thus, the stimuli we used display realistic examples of misinformation and are highly likely to trigger individuals' reactions. Two out of the three issues can be linked to existing conspiracy theories. The post on immigration is in line with the Great Replacement conspiracy theory where the white European identity and civilization is under threat because (Muslim) immigrants are plotting against the West, to eventually taking it over in a hostile manner (Bergmann, 2021; Cosentino, 2020; Obaidi et al., 2021). The claim of the COVID-19 post suggests that the Chinese government is to be blamed for the planned outbreak of the virus. The post on climate change protests cannot be clearly linked to an existing conspiracy theory. Thus, this dissertation tested different forms of misinformation with recognizable conspiratory narratives.

Furthermore, context matters. We conducted the survey during the COVID-19 pandemic and confronted the participants with a misinforming post on this issue. It is highly likely that this issue dominated people's lives at that point in time, and they might have processed this information differently than the other two issues. Despite a few differences between the topics, this dissertation finds that the more salient a polarized issue was to a person, the more likely it was for individuals to engage with it, even though the veracity of the presented information was questionable. Better understanding what type of issues are vulnerable to misinformation, can be a chance for governments and other societally relevant institutions to contain the spread of online misinformation. By monitoring highly salient or polarized issues, fact checking interventions or warnings can be directly linked to this type of content that is often abused to misinform citizens.

Lastly, this dissertation used different theoretical approaches and concepts. In the realm of mass communication research, as a contribution to the Uses and Gratifications Approach, I consider classic gratifications such as 'entertainment' and enrich it with other gratifications that are more specific to misinformation dissemination, for instance, reaction provocation. I further combine this communication science framework with a crucial concept of psychology: the Dark Triad of personality traits. Building on political science, I test whether there is an interplay between political orientation, political interest and attitudinal congruence. Thus, this thesis shows that different theoretical frameworks from different scientific disciplines can be combined to investigate misinformation on the micro-level.

In sum, based on the findings of this thesis, I was able (1) to characterize the individuals who engage with misinformation on social media platforms and (2) to describe how and why individuals react to misinforming content. The research conducted in this dissertation is crucial to inform entities that want to combat online misinformation. Knowing what we know from this thesis, specific solutions tailored to the most vulnerable individuals can be created. These solutions, such as enhancing media literacy or fact checking interventions on social media, would be the most fruitful if designed along political ideology rather than based on individual characteristics such as personality (Buchanan, 2020a). In addition, targeting smaller groups will work better than targeting an entire population. Fighting online misinformation needs to happen step by step with concrete measures and not with catch-them-all solutions. Countering online misinformation in a fine-grained manner can, however, only work if platforms start sharing data with misinformation scholars (Pasquetto et al., 2020). Only like this, we will be able to investigate misinformation on a large scale in real-life settings and test what types of interventions can limit the dissemination of online misinformation (Pasquetto et al., 2020).

Limitations & Avenues for Future Research

Naturally, this dissertation has limitations that have to be addressed. The following paragraphs discuss the limitations of the choices made in this thesis and how future research can overcome these shortcomings.

Three central limitations are directly related to the methods used and require carefulness when generalizing the findings. First, this study uses a largescale survey in which respondents were presented with social media posts on different societal issues; immigration, climate change protests and COVID-19. I believe that these three stimuli are well suited to investigate the individual spread of misinformation because they have been linked to false messages that have been diffused on social media prior to the survey (Kumar, 2020; Mimikama, 2019). Yet, the position reported in each of these posts was formulated in a right-wing direction. This comes with the obvious limitation that I can only make claims for individuals with right-leaning attitudes. I do not know if left-leaning citizens would react in the same manner as right-leaning individuals, if they were confronted with attitudinal congruent misinformation. I encourage scholars to conduct similar experiments using a broader set of issues, both formulated in a left- and right-wing direction. Further, the survey was taken at one point in time, and we do not have access to panel data. Because we presented the participants with multiple manipulated posts, we did not ask credibility questions about each stimulus to avoid priming respondents about the intentional misleading or false character of the claims (Klar et al., 2020). Future research should study online misinformation longitudinally to see whether repeated exposure to misinformation changes individuals' behavior.

A second concern that relates to the generalizability of my findings is the country selection. The sample consisted out of Belgium, Germany, France, Switzerland, the UK, and the U.S. The selection is based on specific macrolevel factors: Trust in media, news media reach, strength of public broadcasting, model of democracy, and party diversity. The selected countries display various levels of each contextual factor. For example, the U.S. has the weakest public broadcaster, whereas Germany has the strongest. Social media as a news source has the highest reach in Switzerland and the U.S. and the lowest in Germany and the UK. Due to the diversity in terms of media and political systems, the sample of countries allows for a robust test of the drivers behind misinformation dissemination. However, there is a clear Western bias in the country selection. Thus, even though we selected six different democracies as our sample, which have different macrolevel characteristics, we absolutely miss other global regions (e.g., Global South) and other forms of government (e.g., authoritarian regimes). Since this thesis did not explore country differences, but used the country variation as robustness checks, comparative research on different countries and political systems could further advance global

misinformation research and clarify the validity of our findings in different contexts. This is of importance because the spread of misinformation in these countries might be higher and state-sponsored. Furthermore, might individuals not be allowed or less daring to publicly question state-sponsored misinformation. For instance, in Brazil misinformation is being instrumentalized as a political weapon and governmental institutions are organizing the spread of false and misleading information (Ricard & Medeiros, 2020). Hence, the dynamics of misinformation creation and dissemination are different in these political and national contexts.

The third main limitation that is linked to the research design refers to the performed manual content analysis. The content analysis focused on the comment section on Twitter and investigated in detail comments on misinformation involving UK politicians. Zeroing in on one country and one platform, enabled me to conduct a thorough manual content analysis. However, this again hampers the generalizability of the findings. Considering a different platform, coding more comments, or focusing on a different country, could lead to different results. Probably, the findings would be substantially different if Facebook comments were studied (see further). As an avenue for future research, I recommend larger investigations into the comment section of different platforms (e.g. Instagram or TikTok), to consider different national contexts, and to look at different types of misinformation. However, even though the content analysis in this dissertation is limited in its scope, my approach shows that a manual content analysis has the advantage to thoroughly investigate the contents of the comment section linked to political misinformation. Automated and machine learning approaches are still limited when it comes to detecting certain forms of speech like sarcasm or irony (Sykora et al., 2020). The way forward for future work is to combine the manual approach with the benefits of automated analyses. Metzger et al. (2021), for instance, used such a mixed approach. In a first step, they identified misinforming posts, collected the connected comments and then manually coded each comment for expressing belief or disbelief in misinformation. In a second step, the authors applied a lexicon-based approach and natural language processing models to see if there are language differences in the comments and to create a classifier that automatically labels comments as 'dis-or belief'. Overall, mixed method approaches are something to strive for, and misinformation research would profit from the benefits from different methodological approaches. If future research combines automated and manual content analysis, it will bring forward a better understanding of the scope of the spread of online misinformation and also advance efficiently measuring how individuals engage with misinforming content.

A limitation which cannot go unaddressed are possible platform differences. This dissertation focused on Facebook (survey) and on Twitter (manual content analysis). After all, these are two of the most studied platforms in the field of misinformation research and Facebook is the most frequently used

platform within all the selected countries (Newman et al., 2022). Consequently, misinformation is spread commonly on Facebook and Twitter and can therefore reach large audiences. Focusing on these two platforms is therefore a valid choice when investigating the individual spread of misinformation online. The two platforms, however, do differ from each other. Twitter is a very public platform, whereas content on Facebook is mostly shared within a selected group or network. Hence, one central difference between the two platforms is how associations are made. On Twitter, the connections can be one-sided, which means that individuals can follow public accounts without the other person's consent. On Facebook, to become friends, both individuals have to approve (Vraga & Bode, 2018). This is an important difference when investigating the individual spread of online misinformation, as previous research has shown that users are more likely to share false or misleading content that comes from a trusted source, such as a friend on Facebook (Buchanan & Benson, 2019). Based on a report of the Pew Research Center (2021), we know that the audiences on the two platforms differ. For instance, more women indicated that they use Facebook than men and on Twitter the opposite is the case. In addition, is the average U.S. Twitter user a lot younger than the average Facebook user (Auxier & Anderson, 2021). The platforms also differ from each other regarding partisanship. According to the Pew Research Center, Democrats are generally using Twitter more often than Republicans, whereas on Facebook this difference is not present (Vogels et al., 2021). What implications do these platform differences have for the findings of this dissertation? One major implication could be that individuals who engage with misinformation that resembles Facebook posts, have completely different characteristics than individuals who would engage with misinformation on Twitter. Different platform audiences react differently to online misinformation (see also Poell et al. (2019)). A certain misleading message could be successful on one platform and not gain any traction on the other. For instance, the findings of this dissertation suggest that on Twitter there seems to be more political debating present. The majority of the users goes against misinformation and does not support it. On Facebook, congruent misinformation seems to gain a lot of support by followers. Future research should delve deeper into investigating systematic platform differences by conducting the same research across platforms. It is not only crucial to compare the platforms with each other, but also to conquer new platforms. We know that, for instance, on WhatsApp misinformation is thriving too (Herrero-Diz et al., 2020). However, accessing the encrypted (personal) messages is only possible if individuals give their consent. We recommend for future research to work together with users to gain access to valuable real-world user data based on confirmed consent to share personal data. In addition to messaging applications, more audiovisual social media platforms (e.g. TikTok) should be considered when researching the spread of misinformation in the future. We know little about the role of visuals or sound in misinformation. In line with Hameleers et al. (2020), I expect that misinformation containing visuals will seem more credible and accurate to individuals, which makes this type of misinformation even

more ‘successful’. I hope in my follow-up research on this topic to also explore audiovisual social media content to broaden the little existing research on audiovisual misinformation.

Lastly, one more shortcoming has to be addressed. This dissertation identified individuals who engage with misinformation for multiple reasons based on their sociodemographics, political variables and personality traits. One aspect that was only partially taken into account, and that might be relevant in this regard, is emotionality or emotions. One reason why misinformation is so successful, some scholars argue (Ecker et al., 2020; Pasquetto et al., 2020; Zhang et al., 2021), is because of its emotionality or because it generates strong emotions. Because social media users react to emotional false and misleading content a lot more and faster (Pasquetto et al., 2020). It seems that information that evokes and contains emotions is appealing to spread on social media. Yet, we know little about the exact influence of emotions on the belief in misinformation or the engagement with misinforming content. In this thesis, I find some evidence that political misinformation triggers more negative sentiment in the comment section and individuals express dissatisfaction and unhappiness. But future research could elaborate on this, with the help of transparent data from platforms, by investigating what types of emotions are triggered by misinformation and what effect emotional misinformation has on, for instance, network polarization or the belief in false and misleading content.

The Prospect of Misinformation Research

The incredible and continuous surge of misinformation research since the 2016 U.S. presidential election across scientific disciplines has proven that misinformation is a ‘hot topic’. The field is moving fast and has also been under some criticism. In this final paragraph, I would therefore like to reflect on where the field might be going in the near future. One thing is certain, the field will stay relevant as dis- and misinformation is still on the rise. Crises like the global COVID-19 pandemic or the war in Ukraine trigger and facilitate the dissemination of misinformation (T. Tran et al., 2020). As a consequence, not only different scientific disciplines are interested in the topic but also policymakers, governments and supranational institutions have been investing recourses to study misinformation. This makes misinformation a unique case. However, the field is also facing some serious challenges.

First, studies on misinformation have been observed to lack consensus on definitions. Lacking a shared understanding of the term makes it difficult to investigate the same research object. Other critiques state that misinformation research has highlighted media effects too much, underestimated the influence of funding institutions on the research, and have an immense influence on policymaking (Camargo & Simon, 2022). What do the listed challenges mean for misinformation research? Of course, they can have serious consequences for the quality of the research. However, in my view Camargo and Simon (2022, pp. 5–6) have put it perfectly: “As such, the field is too big to fail—and cannot be allowed

to. Therefore, it is vital that the community active in this space engages with the shortcomings of the field and tries to find ways to address them".

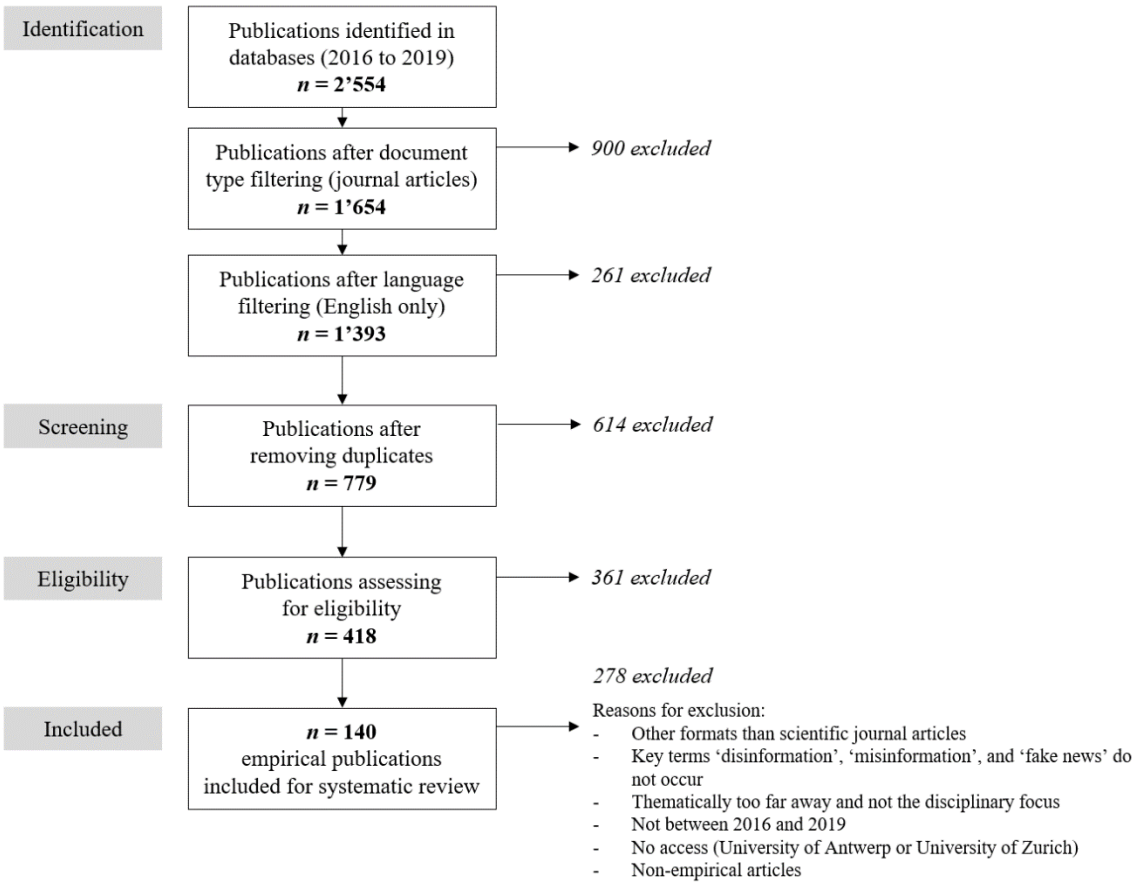
And what about the future? In the near future, at least three things gain a lot more attention within the field of misinformation research. First, I think that voices will be getting louder for social media platforms to cooperate with the scientific community and governments (Pasquetto et al., 2020). Platforms have to start sharing complete datasets and make the logic of their algorithms transparent. We can already observe some developments in that direction. So did the European Commission implement the European Centre for Algorithmic Transparency. The main goal of this institution is to "contribute to a safer, more predictable and trusted online environment for people and business" (European Commission, 2022). In the U.S., the Algorithmic Justice and Online Platform Transparency Act was introduced in Congress in 2021 and aims to "prohibit the discriminatory use of personal information by online platforms in any algorithmic process, to require transparency in the use of algorithmic processes and content moderation, and for other purposes" (H.R.3611 - Algorithmic Justice and Online Platform Transparency Act, 2021). The bill is currently awaiting to be passing the House. Second, artificial intelligence will get much more important when researching the creation and spread of online misinformation. AI has been instrumentalized to create deepfakes to harm citizens and democracies. But it can also offer solutions to mitigate the spread of online misinformation. Tools involving AI, can for instance, reverse-engineer videos or images that have been manipulated (Cassauwers, 2019). Furthermore, the idea of the Web 3.0 and blockchain could help to eliminate mis- and disinformation (e.g. Tee & Murugesan, 2018). Through blockchain, users will be able to influence the algorithms of their information flows and blockchain can track the origin of information to check whether it has been altered or not. And finally, I believe that the end user will be even more in the spotlight. Because people will always have reasons and very different reasons to spread misinformation and create misinformation – but they also have the power to fight it. I hope my dissertation can inspire scholars, practitioners and ordinary citizens to improve and broaden that fight. There is a lot at stake.

Appendices

Appendix 1

In order to systematize the terms and definitions used, we rely on a content analysis of published empirical articles from 2016 to 2019 listed on Web of Science, Scopus and EBSCO's databases Communication & Mass Media Complete and Political Science. The assessment of the articles follows the PRISMA guidelines and includes four steps of filtering (identification, screening, eligibility, inclusion) (Moher et al., 2009). After removing duplicates, the publications were screened for their eligibility. We excluded articles which had other formats than scientific articles (e.g. conference reports or book reviews) and that did not display any of the relevant search terms (disinformation, misinformation, fake news). By consulting the abstracts, we removed articles that were either thematically too distant from researching dis- and misinformation or did not have the disciplinary focus. Lastly, were all non-empirical articles excluded. The assessment lead to *n* = 140 eligible empirical publications that were analyzed in full-text (see Figure 1). This accounts for 5.6% of the originally recovered articles.

Figure 1. Flow diagram of the selection process



In order to detect the potential inconsistencies in the use of terms and definitions, we measure if the authors provided a definition of the concept. Furthermore, we investigate how specific terms (e.g.

"fake news") are defined, and vice versa if and which various terms were used for specific concepts (e.g. fake news is defined as disinformation).

Results

Starting by measuring if authors provided definitions of the terms and which terms were used, the results indicate an interesting, but concerning situation: Only 54% (*n* = 76) provided a definition of the phenomenon, whereas 46% (*n* = 64) did not define the concept explicitly. Considering this result by year, we observe that the amount of studies which did not provide a definition decreased from 88% in 2016, 53% in 2017, 46% in 2018, to 38% in 2019. This trend is at least promising, but the number of over 40% of studies not providing a proper definition for the concept connected to disinformation, misinformation and "fake news" is still substantial.

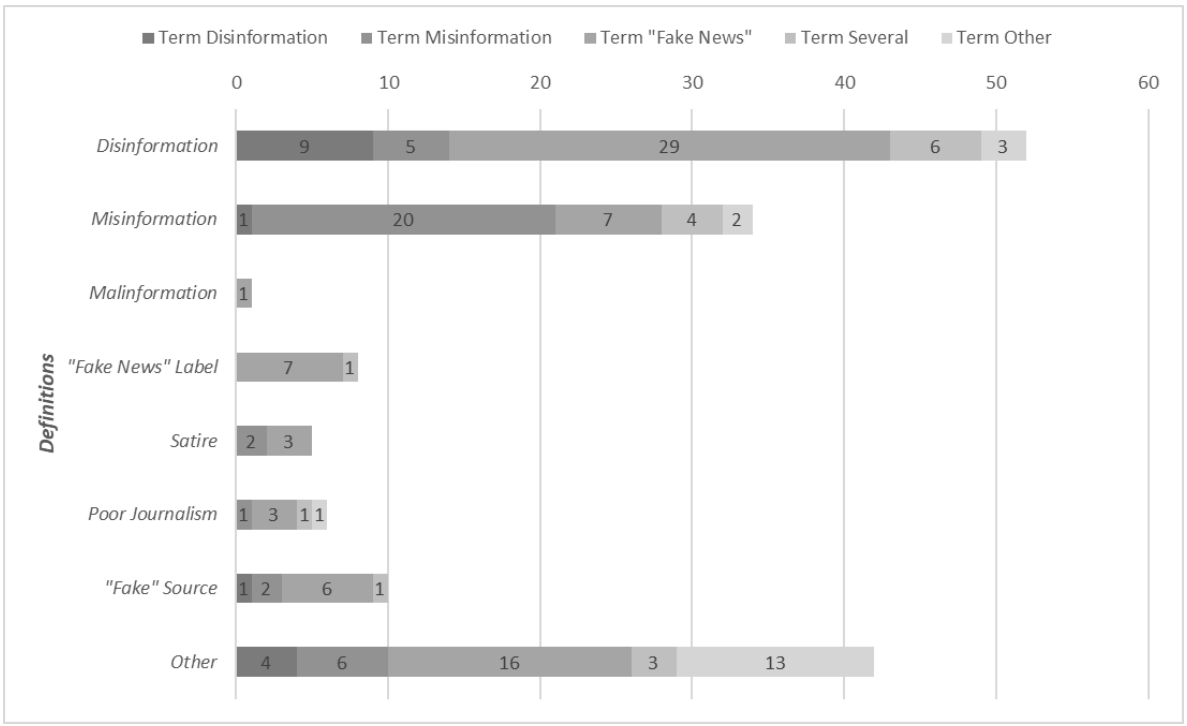
Within those studies, which did not define the concept (*n* = 64), the following terms were used most frequently: 50% (*n* =32) used the term misinformation, 30% (*n* = 19) used the term "fake news", and 8% (*n* = 5) used the term disinformation. This finding points either to a supposedly clear understanding of the terms "misinformation" and "fake news" and the assumption of a uniform definition, or to the fact that these terms were used as buzzwords within a research trend without deeper examination.

Within the studies that did provide a definition (*n* = 76), 49% used the term "fake news" (*n* = 37), in 17% (*n* = 13) of the cases the term misinformation (*n* =13) or other terms (*n* = 13) were used (e.g., rumors or conspiracy theories), 9% used the term disinformation (*n* = 7) and 8% used several of the key terms (*n* = 6). The distribution of those terms reflects only the terms that the authors decided to use, but they do not directly reflect how the authors then defined the concepts. Therefore, we aimed to combine the terms that were used with the definitions that were provided in order to analyze if scholars have the same understanding of those supposedly similar terms or whether their understanding of the terms differs from each other.

Figure 2 shows the distribution of terms and definitions. The term misinformation is to a small degree defined as intentionally false content, satire, poor journalism, fake sources and other definitions, however, the majority of studies that used the term misinformation also defined it as misinformation (n = 20). A similar pattern appears for the term disinformation (n = 15), which is predominantly defined as false content created or spread with the intention to harm (n = 9). Zeroing in on the definitions, we see that for the definition of disinformation (*n* = 52) several other terms are used: It is striking that the most dominant term is "fake news" with 55% (*n* = 29), only 17% of the studies (*n* = 9) used the term disinformation, and some used the term misinformation (*n* = 5, 10%). This means that the term fake news is often defined as false content created or spread with the intention to harm. Further, the results show that the term "fake news" occurs in all definition-categories and can therefore be interpreted as

the most unclear term from a conceptual perspective in political and communication science until 2019.

Figure 2. Terms and definitions



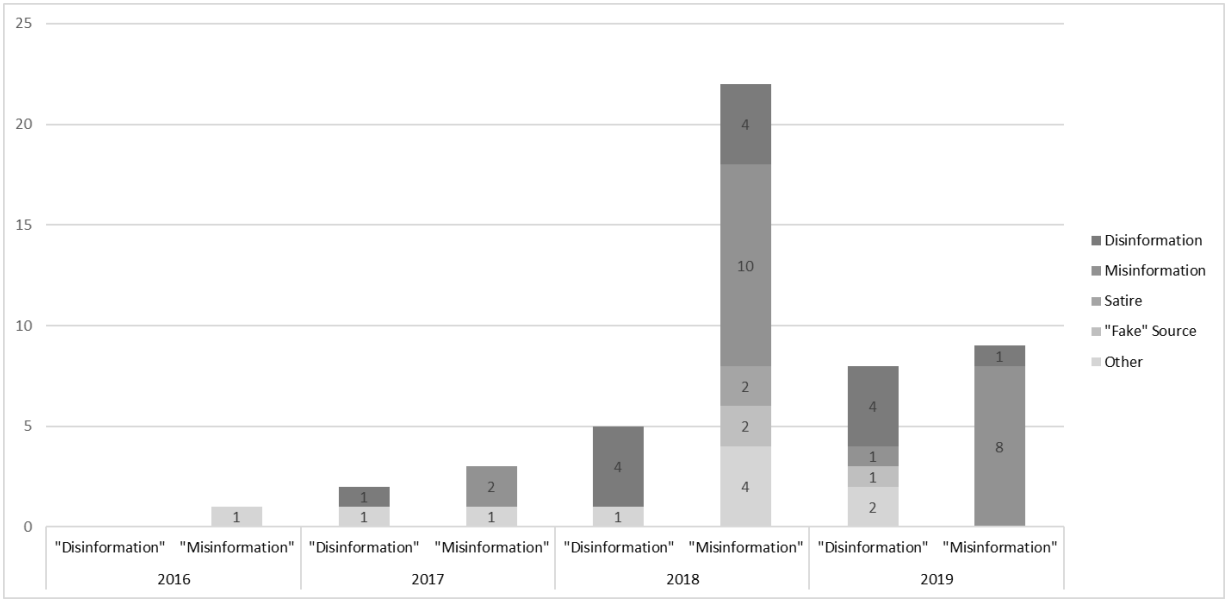
Note. The definitions were measured via a dummy coding, which means that within one study several definitions might have appeared. This leads to a sum of frequencies that exceed 100%.

In a next step, we look at the evolution over time of the use of definitions for the terms dis- and misinformation (2016 – 2019). First, the results in Figure 3 show that misinformation was more common as a term than disinformation. Furthermore, the term disinformation first appeared in publications in 2017 and its use slightly increased over the years (2017: n = 2, 2018: n = 5, 2019: n = 8). In comparison, the term misinformation was used between 2016 and 2017 very little but had a severe increase in 2018 (n = 22). In 2019 the usage of the term decreased again (n = 9). Linking the terms with the definitions, our findings show that the term disinformation, over the years, was predominantly defined as disinformation, but in 2019 it was also defined as misinformation, as fake sources, or as other concepts such as propaganda or rumors. Even though there is no unified or consistent definition for disinformation, the definitions point at a similar direction: For example, the definitions of fake sources (Grinberg et al., 2019) matches the definition of intentionally false content to the degree that the producers of content have certain intents and/or aim to push a specific political interest or agenda. However, fake sources or senders of propaganda do not only publish content that is clearly false: While fake sources often mix false stories with correct information on their websites or social media accounts, propaganda messages are in itself not all false but emotional or pointed in a way to lead or manipulate recipients in a certain intended direction (Tucker et al., 2018). Thus, those definitions do

follow the definition of disinformation regarding (harmful) intentions, but not that clearly regarding falseness.

Similarly to the term disinformation, studies which used the term misinformation define it mostly as misinformation. In 2018 however, different understandings of the term appeared. For instance, misinformation was defined as disinformation (n = 4) or as other concepts such as conspiracy theories (n = 4). This means that the terms seem to still be understood in relatively broad and inconclusive conceptual perspectives.

Figure 3. Definitions for the terms “disinformation” and “misinformation” from 2016 to 2019



Appendix 1.1

Table 1. Census data per country

	BE		CH		DE		UK		FR		US	
	n	%	n	%	n	%	n	%	n	%	n	%
Gender												
Male	514	50%	569	50%	507	51%	602	50%	611	49%	549	49%
Female	530	50%	579	50%	487	49%	792	50%	656	51%	500	51%
Age (years)												
18 - 29	249	22%	259	21%	207	21%	270	24%	242	21%	246	26%
30 - 39	181	20%	253	21%	179	18%	262	20%	210	19%	188	19%
40 - 49	188	20%	244	21%	204	21%	256	20%	245	20%	255	20%
50 - 59	225	21%	233	22%	231	23%	322	20%	264	20%	213	20%
60 - 69	200	17%	159	16%	173	17%	284	16%	306	19%	147	15%
Education												
Low	249	30%	88	18%	283	29%	430	21%	297	27%	447	42%
Medium	429	38%	652	48%	346	34%	520	43%	596	44%	280	29%
High	366	33%	408	34%	365	36%	44	37%	374	30%	317	29%

Note: In the survey we used a string variable for age. Certain age group frequencies might be slightly higher or lower than the quota since respondents had to be excluded that indicated their age incorrectly.

Appendix 2

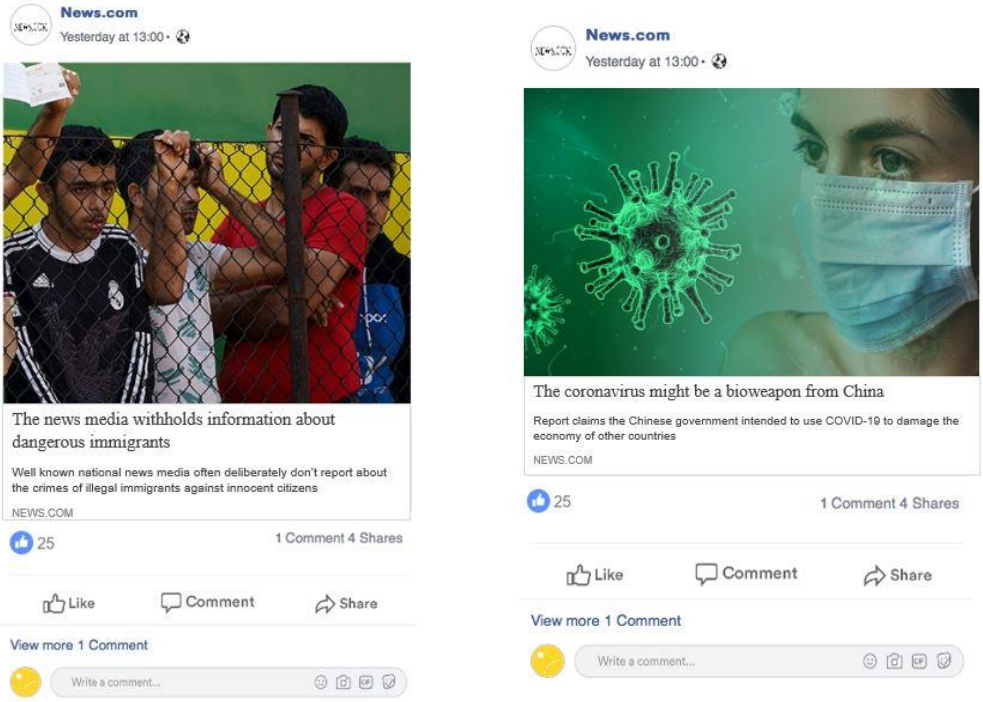
Table 1. Census data per country

	BE		CH		DE		UK		FR		US	
	n	%	n	%	n	%	n	%	n	%	n	%
Gender												
Male	514	50%	569	50%	507	51%	602	50%	611	49%	549	49%
Female	530	50%	579	50%	487	49%	792	50%	656	51%	500	51%
Age (years)												
18 - 29	249	22%	259	21%	207	21%	270	24%	242	21%	246	26%
30 - 39	181	20%	253	21%	179	18%	262	20%	210	19%	188	19%
40 - 49	188	20%	244	21%	204	21%	256	20%	245	20%	255	20%
50 - 59	225	21%	233	22%	231	23%	322	20%	264	20%	213	20%
60 - 69	200	17%	159	16%	173	17%	284	16%	306	19%	147	15%
Education												
Low	249	30%	88	18%	283	29%	430	21%	297	27%	447	42%
Medium	429	38%	652	48%	346	34%	520	43%	596	44%	280	29%
High	366	33%	408	34%	365	36%	44	37%	374	30%	317	29%

Note: In the survey we used a string variable for age. Certain age group frequencies might be slightly higher or lower than the quota since respondents had to be excluded that indicated their age incorrectly.

Appendix 2.1

Figure I: Examples for online conspiracy theories vignettes



Appendix 2.2

Liking

			Immigration				
BE	5.43	1.77	4.00	2.21	2.81	2.11	160
CH	5.78	1.44	4.03	2.17	2.62	2.11	110
DE	5.50	1.70	3.97	2.24	2.78	2.09	144
UK	5.37	1.62	4.23	2.05	3.07	2.20	155
FR	5.50	1.71	4.87	1.87	2.48	1.95	167
US	5.52	1.57	4.56	1.99	3.46	2.25	208
Total	5.50	1.64	4.31	2.10	2.91	2.15	944
			COVID-19				
BE	5.04	1.75	4.13	2.11	3.07	2.14	164
CH	5.09	1.82	3.66	2.10	2.87	2.18	129
DE	4.92	2.01	3.74	2.15	3.14	2.19	138
UK	5.39	1.64	3.94	2.16	3.15	2.18	205
FR	5.17	1.92	4.83	1.89	2.45	1.96	219
US	5.30	1.72	4.42	2.08	3.42	2.22	265
Total	5.18	1.80	4.20	2.11	3.03	2.16	1120

Sharing

			Immigration				
BE	5.07	2.01	4.12	2.12	2.61	2.01	178
CH	5.34	1.82	4.22	2.08	2.81	2.20	119
DE	5.11	2.14	4.19	2.14	2.87	2.19	159
UK	5.37	1.64	4.41	2.06	2.88	2.05	169
FR	5.07	1.84	5.13	1.76	2.39	2.06	189
US	5.40	1.78	4.70	1.97	3.30	2.26	204
Total	5.22	1.88	4.49	2.04	2.82	2.14	1018
			COVID-19				
BE	4.88	1.92	4.28	2.07	2.85	2.15	192
CH	4.81	2.06	3.72	2.08	2.91	2.20	139
DE	4.78	2.07	4.11	2.15	3.04	2.26	160
UK	5.16	1.72	4.19	2.08	3.19	2.21	231
FR	5.04	1.90	5.06	1.84	2.44	1.94	271
US	5.06	1.91	4.53	2.01	3.32	2.23	276
Total	4.98	1.92	4.40	2.06	2.96	2.17	1269

Commenting

			Immigration				
BE	4.59	2.26	4.06	2.11	2.77	2.16	173
CH	4.36	2.39	3.94	2.26	2.59	2.11	114
DE	4.24	2.38	4.03	2.22	2.56	2.06	173
UK	4.83	2.11	4.26	2.10	2.99	2.18	167
FR	4.55	2.26	5.16	1.79	2.42	2.03	167
US	5.14	1.91	4.69	2.01	3.36	2.22	208
Total	4.65	2.22	4.39	2.11	2.81	2.15	1002
			COVID-19				
BE	4.20	2.24	4.49	2.10	2.75	2.18	178
CH	3.94	2.38	3.50	2.22	2.63	2.19	120
DE	3.93	2.39	3.94	2.23	3.04	2.22	167
UK	4.33	2.81	4.18	2.12	2.85	2.14	223
FR	4.24	2.41	4.98	1.94	2.35	1.96	245
US	4.73	2.29	4.57	2.11	3.11	2.24	257
Total	4.29	2.34	4.37	2.15	2.79	2.16	1190

Appendix 2.3

Items Dark Triad of personality traits

Narcissism

To what extent do you agree with the following statements?

I tend to want others to admire me.

I tend to want others to pay attention to me.

I tend to seek prestige or status.

I like to get acquainted with important people.

Psychopathy

To what extent do you agree with the following statements?

I tend to lack remorse.

I tend to be callous or insensitive.

People who mess with me always regret it.

Machiavellianism

To what extent do you agree with the following statements?

I have used flattery to get my way.

It's wise to keep track of information that you can use against people later.

Make sure your plans benefit you, not others.

Most people can be manipulated.

Appendix 2.4

Likert items attitudinal congruence

Immigration

Now we would like to know more about your opinion on immigration. Please indicate to what extent you agree with the following statements.

Immigration to the U.S. should be restricted.

Immigrants contribute to the welfare of the U.S.

The social services in the U.S. are burdened by immigrants.

Immigrants are entitled to social support.

Appendix 2.5

Means per country of all independent variables

Table 1. Means per country for personality traits

Narcissism

Country	Mean	Std. Deviation	N
BE	2.97	1.54	1063
CH	3.18	1.42	1250
DE	3.04	1.52	1019
UK	2.86	1.59	1380
FR	2.54	1.45	1253
US	3.25	1.68	1038
Total	2.96	1.55	7003
Psychopathy			
BE	3.16	1.35	1062
CH	2.86	1.20	1250
DE	2.87	1.30	1019
UK	2.45	1.36	1380
FR	3.34	1.39	1253
US	2.69	1.53	1038
Total	2.89	1.39	7002

Machiavellianism			
BE	3.61	1.21	1062
CH	4.09	1.19	1249
DE	4.08	1.23	1019
UK	3.26	1.35	1380
FR	3.88	1.23	1253
US	3.36	1.46	1038
Total	3.71	1.32	7001
Need for Drama			
BE	2.97	1.36	1063
CH	2.89	1.16	1250
DE	2.85	1.25	1019
UK	2.93	1.50	1380
FR	3.42	1.40	1253
US	3.27	1.50	1037
Total	3.05	1.39	7002

Table 2. Means per country for political orientation

Country	Mean	Std. Deviation	N
BE	6.38	2.56	1062
CH	5.82	2.43	1248
DE	5.65	2.15	1018
UK	6.02	2.15	1380
FR	6.02	2.67	1252
US	6.39	2.71	1038
Total	6.04	2.46	6998

Table 3. Means per country for political interest

Country	Mean	Std. Deviation	N
BE	3.93	1.89	1064
CH	4.26	1.78	1251
DE	4.71	1.70	1019
UK	4.20	1.88	1381
FR	3.72	1.93	1254
US	4.52	1.84	1038
Total	4.21	1.87	7007

Table 4. Means per country for social media use

Country	Mean	Std. Deviation	N
BE	2.79	1.50	1065
CH	2.40	1.37	1251
DE	2.72	1.54	1019
UK	2.75	1.61	1381
FR	2.63	1.43	1255
US	3.18	1.77	1038
Total	2.73	1.55	7009

Table 5. Means per country for social media activity

Country	Mean	Std. Deviation	N
BE	3.35	1.45	1065
CH	3.06	1.51	1251
DE	3.25	1.67	1019
UK	3.46	1.60	1381
FR	3.44	1.40	1255
US	3.72	1.65	1038
Total	3.38	1.56	7009

Table 6. Means per country and issue for trust

Trust in News			
Country	Mean	Std. Deviation	N
BE	4.59	1.40	1065
CH	4.30	1.45	1251
DE	4.50	1.56	1019
UK	4.20	1.52	1381
FR	3.80	1.43	1255
US	4.07	1.68	1037
Total	4.23	1.53	7008
Trust in Social media News			
BE	3.28	1.56	1065
CH	3.21	1.48	1251
DE	3.37	1.60	1019
UK	2.99	1.59	1381
FR	3.14	1.48	1255
US	3.30	1.73	1037
Total	3.20	1.57	7008

Trust in Government			
BE	3.24	1.63	1065
CH	4.07	1.54	1251
DE	3.79	1.71	1019
UK	3.46	1.64	1381
FR	2.72	1.63	1255
US	3.12	1.61	1038
Total	3.40	1.69	7009

Appendix 2.6

Pairwise correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
(1) like immigration	1.000																			
(2) share immigration	0.627	1.000																		
	(0.000)																			
(3) comment immigration	0.514	0.595	1.000																	
	(0.000)	(0.000)																		
(4) like covid	0.450	0.347	0.315	1.000																
	(0.000)	(0.000)	(0.000)																	
(5) share covid	0.347	0.460	0.344	0.599	1.000															
	(0.000)	(0.000)	(0.000)	(0.000)																
(6) comment covid	0.310	0.367	0.488	0.458	0.539	1.000														
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)															
(7) attitudes immigration	0.161	0.165	0.064	0.088	0.100	0.046	1.000													
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)														
(8) attitudes covid	0.021	0.051	0.045	0.019	0.042	0.042	0.065	1.000												
	(0.08	(0.00	(0.00	(0.11	(0.00	(0.00	(0.00													

	1)	0)	0)	1)	0)	0)	0)							
(9) political orientation	0.230	0.211	0.138	0.146	0.152	0.115	0.416	-0.041	1.000					
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)						
(10) political interest	0.157	0.135	0.181	0.072	0.068	0.142	-0.130	-0.090	0.077	1.000				
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
(11) Narcissism	0.198	0.182	0.193	0.186	0.172	0.169	-0.122	-0.099	0.080	0.169	1.000			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				
(12) Psychopathy	0.229	0.221	0.218	0.196	0.179	0.168	0.109	0.031	0.169	0.040	0.367	1.000		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.010)	(0.000)	(0.001)	(0.000)			
(13) Machiavellianism	0.188	0.182	0.165	0.150	0.143	0.124	0.106	-0.044	0.136	0.102	0.412	0.513	1.000	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
(14) Need for drama	0.233	0.239	0.254	0.225	0.217	0.219	0.033	0.061	0.133	0.114	0.383	0.515	0.506	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
(15) sex	0.097	0.101	0.110	0.054	0.051	0.084	0.027	-0.013	0.117	0.270	0.094	0.162	0.132	0.170	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.026)	(0.268)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
(16) age	0.001	0.013	0.014	-0.024	-0.004	0.025	0.215	0.018	0.068	0.136	-0.269	-0.165	-0.149	-0.101	0.124	1.000

	(0.922)	(0.288)	(0.241)	(0.044)	(0.732)	(0.039)	(0.000)	(0.124)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				
(17) education	-0.057	-0.071	-0.046	0.100	0.105	0.069	0.186	0.066	0.043	0.178	0.121	0.009	0.061	0.011	0.056	-0.179	1.000		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.456)	(0.000)	(0.378)	(0.000)	(0.000)			
(18) Social media activity	0.244	0.267	0.271	0.276	0.290	0.265	-0.084	0.055	0.053	0.112	0.329	0.221	0.166	0.242	0.019	-0.262	0.016	1.000	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.107)	(0.000)	(0.182)		
(19) Trust in news	0.035	0.029	0.045	0.070	0.080	0.078	-0.105	-0.241	-0.018	0.125	0.154	0.005	0.053	-0.005	0.023	0.069	-0.016	0.078	1.000
	(0.004)	(0.014)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.123)	(0.000)	(0.000)	(0.705)	(0.000)	(0.668)	(0.050)	(0.000)	(0.187)	(0.000)	
(20) Trust in social media	0.180	0.205	0.192	0.217	0.240	0.191	-0.003	-0.063	0.069	0.020	0.258	0.166	0.127	0.154	0.016	-0.070	-0.093	0.396	0.465	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.775)	(0.000)	(0.000)	(0.090)	(0.000)	(0.000)	(0.000)	(0.000)	(0.174)	(0.000)	(0.000)	(0.000)	(0.000)	

Appendix 2.7

Note: This is an example of the data structure with multiple respondents for a few variables. In this case we have three respondents. Respondent 1 indicated to like and share the immigration story, respondent 2 indicated to like, share and comment on both stories. Respondent 3 likes the immigration and COVID-19 story.

ID	Respondent ID	Issue	Type	Motivation conviction (DV1)	Motivation entertainment (DV2)	Motivation provoke (DV3)	Attitudinal congruence	Psychopathy	…
1	1	Immigration (0)	Like (0)	6	2	3	6	3.25	..
2	1	Immigration (0)	Share (1)	5	3	6	6	3.25	..
3	2	Immigration (0)	Like (0)	1	2	5	3.25	5.5	..
4	2	Immigration (0)	Share (1)	2	3	6	3.25	5.5	..
5	2	Immigration (0)	Comment (2)	3	4	5	3.25	5.5	..
6	2	Covid (1)	Like (0)	1	3	6	4.25	5.5	..
7	2	Covid (1)	Share (1)	2	3	5	4.25	5.5	..
8	2	Covid (1)	Comment (2)	3	4	4	4.25	5.5	..
9	3	Immigration (0)	Like (0)	4	5	2	4	4.75	..
10	3	Covid (1)	Like (0)	5	4	3	5	4.75	..
11	..	..	..	..	..	..	..	..	..

Appendix 3

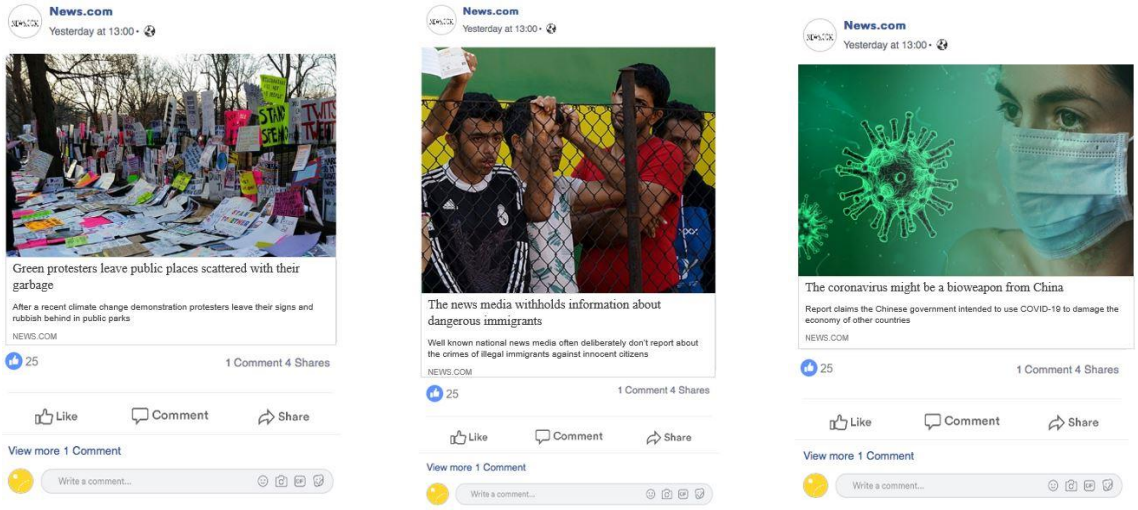
Table 1. Census data per country

	BE		CH		DE		UK		FR		US	
	n	%	n	%	n	%	n	%	n	%	n	%
Gender												
Male	514	50%	569	50%	507	51%	602	50%	611	49%	549	49%
Female	530	50%	579	50%	487	49%	792	50%	656	51%	500	51%
Age (years)												
18 - 29	249	22%	259	21%	207	21%	270	24%	242	21%	246	26%
30 - 39	181	20%	253	21%	179	18%	262	20%	210	19%	188	19%
40 - 49	188	20%	244	21%	204	21%	256	20%	245	20%	255	20%
50 - 59	225	21%	233	22%	231	23%	322	20%	264	20%	213	20%
60 - 69	200	17%	159	16%	173	17%	284	16%	306	19%	147	15%
Education												
Low	249	30%	88	18%	283	29%	430	21%	297	27%	447	42%
Medium	429	38%	652	48%	346	34%	520	43%	596	44%	280	29%
High	366	33%	408	34%	365	36%	44	37%	374	30%	317	29%

Note: In the survey we used a string variable for age. Certain age group frequencies might be slightly higher or lower than the quota since respondents had to be excluded that indicated their age incorrectly.

Appendix 3.1

Figure I: Examples for online misinformation vignette



Appendix 3.2

Table 2. Distributions of people who indicated they would interact with the social media post in percentages

	Climate Change	Respondents (*n*)	Immigration	Respondents (*n*)	COVID-19	Respondents (*n*)
BE	21.4%	186	14.3%	133	17.0%	157
CH	15.5%	171	8.3%	96	10.5%	119
DE	13.2%	111	14.9%	134	13.1%	115
UK	19.4%	235	11.1%	141	15.7%	191
FR	26.1%	266	15.7%	182	21.9%	239
US	27.6%	232	19.3%	173	24.9%	211
N		1201		859		1032

Note: Percentages in the table represent people who answered 5-7 on a seven point scale on how likely it is that they would like, share or comment on the post.

References

AFP, CORRECTIV, Pagella Politica/Facta, Full Fact, & Maldita.es. (2020). *INFODEMIC COVID-19 IN EUROPE: A VISUAL ANALYSIS OF DISINFORMATION*. 1–17.

Ahadzadeh, A. S., Ong, F. S., & Wu, S. L. (2021). Social media skepticism and belief in conspiracy theories about COVID-19: The moderating role of the dark triad. *Current Psychology*. https://doi.org/10.1007/s12144-021-02198-1

Ahmed, S. (2021). Who inadvertently shares deepfakes? Analyzing the role of political interest, cognitive ability, and social network size. *Telematics and Informatics*, *57*, 101508. https://doi.org/10.1016/j.tele.2020.101508

Albahar, M. (2021). A hybrid model for fake news detection: Leveraging news content and user comments in fake news. *IET Information Security*, *15*(2), 169–177. https://doi.org/10.1049/ise2.12021

Alemanno, A. (2018). How to Counter Fake News? A Taxonomy of Anti-fake News Approaches. *European Journal of Risk Regulation*, *9*(1), 1–5. https://doi.org/10.1017/err.2018.12

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The "Nasty Effect:" Online Incivility and Risk Perceptions of Emerging Technologies: Crude comments and concern. *Journal of Computer-Mediated Communication*, *19*(3), 373–387. https://doi.org/10.1111/jcc4.12009

Anspach, N. M., & Carlson, T. N. (2020). What to Believe? Social Media Commentary and Belief in Misinformation. *Political Behavior*, *42*(3), 697–718. https://doi.org/10.1007/s11109-018-9515-z

Apuke, O. D., & Omar, B. (2020). User motivation in fake news sharing during the COVID-19 pandemic: An application of the uses and gratification theory. *Online Information Review*, *45*(1), 220–239. https://doi.org/10.1108/OIR-03-2020-0116

Arendt, H. (1967, February 25). Truth & Politics. *The New Yorker*.

Arif, A., Robinson, J. J., Stanek, S. A., Fichet, E. S., Townsend, P., Worku, Z., & Starbird, K. (2017). A Closer Look at the Self-Correcting Crowd: Examining Corrections in Online Rumors. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 155–168. https://doi.org/10.1145/2998181.2998294

Ashworth-Hayes, S. (2022, April 11). Nadine Dorries didn't say 'food doesn't grow on trees.' *Full Fact*. https://fullfact.org/online/nadine-dorries-fake-tweet/

Auxier, B., & Anderson, M. (2021). A majority of Americans say they use YouTube and Facebook, while use of Instagram, Snapchat and TikTok is especially common among adults under 30. *Pew Research Center*, 8.

Balakrishnan, V., Ng, K. S., & Rahim, H. A. (2021). To share or not to share – The underlying motives of sharing fake news amidst the COVID-19 pandemic in Malaysia. *Technology in Society*, *66*, 101676. https://doi.org/10.1016/j.techsoc.2021.101676

Barkun, M. (2017). President Trump and the "Fringe." *Terrorism and Political Violence*, *29*(3), 437–443. https://doi.org/10.1080/09546553.2017.1313649

Bartlett, J., & Miller, C. (2010). *The Power of Unreason: Conspiracy Theories, Extremism and Counter-Terrorism*. Demos.

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, *33*(2), 122–139. https://doi.org/10.1177/0267323118760317

Bergmann, E. (2021). The Eurabia conspiracy theory. In A. Önnerfors & A. Krouwel (Eds.), *Europe: Continent of Conspiracies* (1st ed., pp. 36–53). Routledge. https://doi.org/10.4324/9781003048640-3

Berinsky, A. J. (2017). Rumors and Health Care Reform: Experiments in Political Misinformation. *British Journal of Political Science*, *47*(2), 241–262. https://doi.org/10.1017/S0007123415000186

Bessi, A., Coletto, M., Davidescu, G. A., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015). Science vs Conspiracy: Collective Narratives in the Age of Misinformation. *PLOS ONE*, *10*(2), e0118093. https://doi.org/10.1371/journal.pone.0118093

Bode, L. (2020). User correction as tool in the battle against social media misinformation. *Georgetown Law Technology Review*, *4*(2), 367–378.

Bode, L., & Vraga, E. (2021). Value for Correction: Documenting Perceptions about Peer Correction of Misinformation on Social Media in the Context of COVID-19. *Journal of Quantitative Description: Digital Media*, *1*. https://doi.org/10.51685/jqd.2021.016

Bode, L., & Vraga, E. K. (2018). See Something, Say Something: Correction of Global Health Misinformation on Social Media. *Health Communication*, *33*(9), 1131–1140. https://doi.org/10.1080/10410236.2017.1331312

Borah, P., Irom, B., & Hsu, Y. C. (2021). 'It infuriates me': Examining young adults' reactions to and recommendations to fight misinformation about COVID-19. *Journal of Youth Studies*, 1–21. https://doi.org/10.1080/13676261.2021.1965108

Bradshaw, S., & Howard, P. N. (2017). Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation. *Oxford, England: Project on Computational Propaganda.*, 37.

Bruder, M., & Kunert, L. (2021). The conspiracy hoax? Testing key hypotheses about the correlates of generic beliefs in conspiracy theories during the COVID -19 pandemic. *International Journal of Psychology*, ijop.12769. https://doi.org/10.1002/ijop.12769

Brüggemann, M., Engesser, S., Büchel, F., Humprecht, E., & Castro, L. (2014). Hallin and Mancini Revisited: Four Empirical Types of Western Media Systems: Hallin and Mancini Revisited. *Journal of Communication*, *64*(6), 1037–1065. https://doi.org/10.1111/jcom.12127

Buchanan, T. (2020a). Why Do People Share Disinformation On Social Media? Policy Brief. *CREST*. https://crestresearch.ac.uk/resources/disinformation-on-social-media/

Buchanan, T. (2020b). Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation. *PLOS ONE*, 33.

Buchanan, T., & Benson, V. (2019). Spreading Disinformation on Facebook: Do Trust in Message Source, Risk Propensity, or Personality Affect the Organic Reach of "Fake News"? *Social Media + Society*, *5*(4), 205630511988865. https://doi.org/10.1177/2056305119888654

Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences*, *67*, 97–102. https://doi.org/10.1016/j.paid.2014.01.016

Burke, M., & Kraut, R. E. (2016). The Relationship Between Facebook Use and Well-Being Depends on Communication Type and Tie Strength: FACEBOOK AND WELL-BEING. *Journal of Computer-Mediated Communication*, *21*(4), 265–281. https://doi.org/10.1111/jcc4.12162

Calvillo, D. P., Ross, B. J., Garcia, R. J. B., Smelter, T. J., & Rutchick, A. M. (2020). Political Ideology Predicts Perceptions of the Threat of COVID-19 (and Susceptibility to Fake News About It). *Social Psychological and Personality Science*, *11*(8), 1119–1128. https://doi.org/10.1177/1948550620940539

Camargo, C. Q., & Simon, F. M. (2022). Mis- and disinformation studies are too big to fail: Six suggestions for the field's future. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-106

Cassauwers, T. (2019). Can artificial intelligence help end fake news? *Horizon*. https://ec.europa.eu/research-and-innovation/en/horizon-magazine/can-artificial-intelligence-help-end-fake-news

Chadwick, A., & Vaccari, C. (2019). News sharing on UK social media: Misinformation, disinformation, and correction. *Loughborough: Online Civic Culture Centre*.

Chadwick, A., Vaccari, C., & O'Loughlin, B. (2018). Do tabloids poison the well of social media? Explaining democratically dysfunctional news sharing. *New Media & Society*, *20*(11), 4255–4274. https://doi.org/10.1177/1461444818769689

Chan, M. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A Meta-Analysis of the Psychological Efficacy of Messages Countering Misinformation. *Psychological Science*, *28*(11), 1531–1546. https://doi.org/10.1177/0956797617714579

Chen, X. (2016). The Influences of Personality and Motivation on the Sharing of Misinformation on Social Media. *IConference 2016 Proceedings*. iConference 2016: Partnership with Society, Philadelphia, USA. https://doi.org/10.9776/16145

Chen, X., & Sin, S.-C. J. (2013). 'Misinformation? What of it?' Motivations and individual differences in misinformation sharing on social media: 'Misinformation? What of it?' Motivations and Individual Differences in Misinformation Sharing on Social Media. *Proceedings of the American Society for Information Science and Technology*, *50*(1), 1–4. https://doi.org/10.1002/meet.14505001102

Chen, X., Sin, S.-C. J., Theng, Y.-L., & Lee, C. S. (2015). Why Students Share Misinformation on Social Media: Motivation, Gender, and Study-level Differences. *The Journal of Academic Librarianship*, 10.

Clarke, S. (2002). Conspiracy Theories and Conspiracy Theorizing. *Philosophy of the Social Sciences*, *32*(2), 131–150.

Clemm von Hohenberg, B. (2019). An Ocean of Possible Truth: Biased Processing of News on Social Media. *SSRN*, 47. http://dx.doi.org/10.2139/ssrn.3281038

Coe, K., Kenski, K., & Rains, S. A. (2014). Online and Uncivil? Patterns and Determinants of Incivility in Newspaper Website Comments. *Journal of Communication*, *64*(4), 658–679. https://doi.org/10.1111/jcom.12104

Colliander, J. (2019). “This is fake news”: Investigating the role of conformity to other users’ views when commenting on and spreading disinformation in social media. *Computers in Human Behavior*, *97*, 202–215. https://doi.org/10.1016/j.chb.2019.03.032

Coluccia, A. (2020). On the probabilistic modeling of fake news (hoax) persistency in online social networks and the role of debunking and filtering. *Internet Technology Letters*, *3*(5). https://doi.org/10.1002/itl2.204

Cosentino, G. (2020). From Pizzagate to the great replacement: The globalization of conspiracy theories. In *Social media and the post-truth world order* (pp. 59–86). Palgrave Pivot.

Council of Europe. (n.d.). *Dealing with propaganda, misinformation and fake news*. Retrieved October 25, 2022, from https://www.coe.int/en/web/campaign-free-to-speak-safe-to-learn/dealing-with-propaganda-misinformation-and-fake-news

Craker, N., & March, E. (2016). The dark side of Facebook®: The Dark Tetrad, negative social potency, and trolling behaviours. *Personality and Individual Differences*, *102*, 79–84. https://doi.org/10.1016/j.paid.2016.06.043

Daniel, L., & Harper, D. J. (2020). The Social Construction of Conspiracy Beliefs: A Q-Methodology Study of How Ordinary People Define Them and Judge Their Plausibility. *Journal of Constructivist Psychology*, 1–22. https://doi.org/10.1080/10720537.2020.1837695

Duffy, A., Tandoc, E., & Ling, R. (2019). Too good to be true, too good not to share: The social utility of fake news. *Information, Communication & Society*, 1–15. https://doi.org/10.1080/1369118X.2019.1623904

Eberl, J.-M., Tolochko, P., Jost, P., Heidenreich, T., & Boomgaarden, H. G. (2020). What’s in a post? How sentiment and issue salience affect users’ emotional reactions on Facebook. *Journal of Information Technology & Politics*, *17*(1), 48–65. https://doi.org/10.1080/19331681.2019.1710318

Ecker, U. K. H., Lewandowsky, S., & Chadwick, M. (2020). Can corrections spread misinformation to new audiences? Testing for the elusive familiarity backfire effect. *Cognitive Research: Principles and Implications*, *5*(1), 41. https://doi.org/10.1186/s41235-020-00241-6

Egelhofer, J. L., & Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: A framework and research agenda. *Annals of the International Communication Association*, *43*(2), 97–116. https://doi.org/10.1080/23808985.2019.1602782

Enders, A. M., Uscinski, J. E., Seelig, M. I., Klofstad, C. A., Wuchty, S., Funchion, J. R., Murthi, M. N., Premaratne, K., & Stoler, J. (2021). The Relationship Between Social Media Use and Beliefs in Conspiracy Theories and Misinformation. *Political Behavior*. https://doi.org/10.1007/s11109-021-09734-6

Epstein, Z., Lin, H., Pennycook, G., & Rand, D. (2022). *How many others have shared this? Experimentally investigating the effects of social cues on engagement, misinformation, and unpredictability on social media* (arXiv:2207.07562). arXiv. http://arxiv.org/abs/2207.07562

Esau, K. (2021). Impoliteness (Hate Speech/Incivility). *DOCA - Database of Variables for Content Analysis*. https://doi.org/10.34778/5b

Esser, F., & Vliegenthart, R. (2017). Comparative Research Methods. In J. Matthes, C. S. Davis, & R. F. Potter (Eds.), *The International Encyclopedia of Communication Research Methods* (pp. 1–22). Wiley. https://doi.org/10.1002/9781118901731

Europa Publications. (2018). *The Europa World Year Book 2018*. Routledge.

European Commission. (2022). European Centre for Algorithmic Transparency. *European Commission*. https://digital-strategy.ec.europa.eu/en/node/11157/printable/pdf

EUvs.Disinfo. (2021). *Hannah Arendt and the Fragility of Facts*. https://euvsdisinfo.eu/hannah-arendt-and-the-fragility-of-facts/

Ferret Journalists. (2021a, February 18). Nigel Farage claim of 'EU Human Rights Act' is False. *The Ferret*. https://theferret.scot/nigel-farage-eu-human-rights-act-is-false/

Ferret Journalists. (2021b, November 24). *Claim about MPs' debate attendance is Mostly False*. https://theferret.scot/claim-about-mps-debate-attendance-is-mostly-false/

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

Flynn, D. J., Nyhan, B., & Reifler, J. (2017). The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs About Politics: Nature and Origins of Misperceptions. *Political Psychology*, *38*, 127–150. https://doi.org/10.1111/pops.12394

Frankowski, S., Lupo, A. K., Smith, B. A., Dane'El, M., Ramos, C., & Morera, O. F. (2016). Developing and Testing a Scale to Measure Need for Drama. *Personality and Individual Differences*, *89*, 192–201. https://doi.org/10.1016/j.paid.2015.10.009

Freiling, I., Krause, N. M., Scheufele, D. A., & Brossard, D. (2021). Believing and sharing misinformation, fact-checks, and accurate information on social media: The role of anxiety during COVID-19. *New Media & Society*, 146144482110114. https://doi.org/10.1177/14614448211011451

Full Fact. (2020). *Viral post about someone's uncle's coronavirus advice is not all it's cracked up to be*. https://fullfact.org/online/coronavirus-claims-symptoms-viral/

Galpin, C., & Trenz, H.-J. (2019). Participatory Populism: Online Discussion Forums on Mainstream News Sites During the 2014 European Parliament Election. *Journalism Practice*, *13*(7), 781–798. https://doi.org/10.1080/17512786.2019.1577164

Garrett, R. K., & Poulsen, S. (2019). Flagging Facebook Falsehoods: Self-Identified Humor Warnings Outperform Fact Checker and Peer Warnings. *Journal of Computer-Mediated Communication*, *24*(5), 240–258. https://doi.org/10.1093/jcmc/zmz012

Garrett, R. K., & Weeks, B. E. (2013). The promise and peril of real-time corrections to political misperceptions. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work - CSCW '13*, 1047. https://doi.org/10.1145/2441776.2441895

Gearhart, S., Moe, A., & Zhang, B. (2020). Hostile media bias on social media: Testing the effect of user comments on perceptions of news bias and credibility. *Human Behavior and Emerging Technologies*, *2*(2), 140–148. https://doi.org/10.1002/hbe2.185

Goreis, A., & Voracek, M. (2019). A Systematic Review and Meta-Analysis of Psychological Research on Conspiracy Beliefs: Field Characteristics, Measurement Instruments, and Associations With Personality Traits. *Frontiers in Psychology*, *10*, 205. https://doi.org/10.3389/fpsyg.2019.00205

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, *363*(6425), 374–378. https://doi.org/10.1126/science.aau2706

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, *5*(1), eaau4586. https://doi.org/10.1126/sciadv.aau4586

Guo, C., Cao, J., Zhang, X., Shu, K., & Yu, M. (2019). Exploiting Emotions for Fake News Detection on Social Media. *Preprint ArXiv*, 8.

Guo, H., Cao, J., Zhang, Y., Guo, J., & Li, J. (2018). Rumor Detection with Hierarchical Social Attention Network. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 943–951. https://doi.org/10.1145/3269206.3271709

Hall Jamieson, K., & Albarracín, D. (2020). The Relation between Media Consumption and Misinformation at the Outset of the SARS-CoV-2 Pandemic in the US. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-012

Halpern, D., Valenzuela, S., Katz, J., & Miranda, J. P. (2019). From Belief in Conspiracy Theories to Trust in Others: Which Factors Influence Exposure, Believing and Sharing Fake News. In G. Meiselwitz (Ed.), *Social Computing and Social Media. Design, Human Behavior and Analytics* (Vol. 11578, pp. 217–232). Springer International Publishing. https://doi.org/10.1007/978-3-030-21902-4_16

Hameleers, M. (2019). To Like Is to Support? The Effects and Mechanisms of Selective Exposure to Online Populist Communication on Voting Preferences. *International Journal of Communication*, *13*, 2417–2436.

Hameleers, M., Humprecht, E., Möller, J., & Lühring, J. (2021). Degrees of deception: The effects of different types of COVID-19 misinformation and the effectiveness of corrective information in crisis times. *Information, Communication & Society*, 1–17. https://doi.org/10.1080/1369118X.2021.2021270

Hameleers, M., Powell, T. E., Van Der Meer, T. G. L. A., & Bos, L. (2020). A Picture Paints a Thousand Lies? The Effects and Mechanisms of Multimodal Disinformation and Rebuttals Disseminated via Social Media. *Political Communication*, *37*(2), 281–301. https://doi.org/10.1080/10584609.2019.1674979

Hameleers, M., & van der Meer, T. G. L. A. (2020). Misinformation and Polarization in a High-Choice Media Environment: How Effective Are Political Fact-Checkers? *Communication Research*, *47*(2), 227–250. https://doi.org/10.1177/0093650218819671

Hameleers, M., van der Meer, T., & Vliegenthart, R. (2021). Civilized truths, hateful lies? Incivility and hate speech in false information – evidence from fact-checked statements in the US. *Information, Communication & Society*, 1–18. https://doi.org/10.1080/1369118X.2021.1874038

Han, J., Cha, M., & Lee, W. (2020). Anger contributes to the spread of COVID-19 misinformation. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-39

Herrero-Diz, P., Conde-Jiménez, J., & Reyes de Cózar, S. (2020). Teens' Motivations to Spread Fake News on WhatsApp. *Social Media + Society*, 1–14.

Hopp, T., Ferrucci, P., & Vargo, C. J. (2020). Why Do People Share Ideologically Extreme, False, and Misleading Content on Social Media? A Self-Report and Trace Data–Based Analysis of Countermedia Content Dissemination on Facebook and Twitter. *Human Communication Research*. https://doi.org/10.1093/hcr/hqz022

Houston, J. B., Hansen, G. J., & Nisbett, G. S. (2011). Influence of User Comments on Perceptions of Media Bias and Third-Person Effect in Online News. *Electronic News*, *5*(2), 79–92. https://doi.org/10.1177/1931243111407618

Hughes, S., & Machan, L. (2021). It's a conspiracy: Covid-19 conspiracies link to psychopathy, Machiavellianism and collective narcissism. *Personality and Individual Differences*, *171*, 110559. https://doi.org/10.1016/j.paid.2020.110559

Humprecht, E. (2019). Where 'fake news' flourishes: A comparison across four Western democracies. *Information, Communication & Society*, *22*(13), 1973–1988. https://doi.org/10.1080/1369118X.2018.1474241

Humprecht, E., Esser, F., & Van Aelst, P. (2020). Resilience to Online Disinformation: A Framework for Cross-National Comparative Research. *The International Journal of Press/Politics*, *25*(3), 493–516. https://doi.org/10.1177/1940161219900126

Hunt, K., Wang, B., & Zhuang, J. (2020). Misinformation debunking and cross-platform information sharing through Twitter during Hurricanes Harvey and Irma: A case study on shelters and ID checks. *Natural Hazards*, *103*(1), 861–883. https://doi.org/10.1007/s11069-020-04016-6

Islam, A. K. M. N., Laato, S., Talukder, S., & Sutinen, E. (2020). Misinformation sharing and social media fatigue during COVID-19: An affordance and cognitive load perspective. *Technological Forecasting and Social Change*, *159*, 120201. https://doi.org/10.1016/j.techfore.2020.120201

Iyengar, S., & Hahn, K. S. (2009). Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use. *Journal of Communication*, *59*(1), 19–39. https://doi.org/10.1111/j.1460-2466.2008.01402.x

Jensen, E. A., Pfleger, A., Herbig, L., Wagoner, B., Lorenz, L., & Watzlawik, M. (2021). What Drives Belief in Vaccination Conspiracy Theories in Germany? *Frontiers in Communication*, *6*, 678335. https://doi.org/10.3389/fcomm.2021.678335

Jiang, S., & Wilson, C. (2018). Linguistic Signals under Misinformation and Fact-Checking: Evidence from User Comments on Social Media. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 1–23. https://doi.org/10.1145/3274351

Johnson, B. K., Neo, R. L., Heijnen, M. E. M., Smits, L., & van Veen, C. (2020). Issues, involvement, and influence: Effects of selective exposure and sharing on polarization and participation. *Computers in Human Behavior*, *104*, 106155. https://doi.org/10.1016/j.chb.2019.09.031

Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment*, *22*(2), 420–432. https://doi.org/10.1037/a0019265

Jones, D. N., & Paulhus, D. L. (2014). Introducing the Short Dark Triad (SD3): A Brief Measure of Dark Personality Traits. *Assessment*, *21*(1), 28–41. https://doi.org/10.1177/1073191113514105

Jones-Jang, S. M., Kim, D. H., & Kenski, K. (2021). Perceptions of mis- or disinformation exposure predict political cynicism: Evidence from a two-wave survey during the 2018 US midterm elections. *New Media & Society*, *23*(10), 3105–3125. https://doi.org/10.1177/1461444820943878

Kalsnes, B. (2018). Fake News. In B. Kalsnes, *Oxford Research Encyclopedia of Communication*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190228613.013.809

Kapantai, E., Christopoulou, A., Berberidis, C., & Peristeras, V. (2020). A systematic literature review on disinformation: Toward a unified taxonomical framework. *New Media & Society*, 26.

Katz, E., Blumler, J. G., & Gurevitch, M. (1974). Uses and Gratifications Research. *The Public Opinion Quarterly*, *37*(4), 509–523.

Katz, E., Haas, H., & Gurevitch, M. (1973). On the Use of the Mass Media for Important Things. *American Sociological Review*, *38*(2), 164. https://doi.org/10.2307/2094393

Keeley, B. L. (1999). Of Conspiracy Theories. *The Journal of Philosophy*, *96*(3), 109–126.

Kim, C., & Yang, S.-U. (2017). Like, comment, and share on Facebook: How each behavior differs from the other. *Public Relations Review*, *43*(2), 441–449. https://doi.org/10.1016/j.pubrev.2017.02.006

Kim, J. W., & Masullo Chen, G. (2021). Exploring the Influence of Comment Tone and Content in Response to Misinformation in Social Media News. *Journalism Practice*, *15*(4), 456–470. https://doi.org/10.1080/17512786.2020.1739550

Kim, J. won. (2014). Scan and click: The uses and gratifications of social recommendation systems. *Computers in Human Behavior*, *33*, 184–191. https://doi.org/10.1016/j.chb.2014.01.028

Kim, Y., & Kim, Y. (2019). Incivility on Facebook and political polarization: The mediating role of seeking further comments and negative emotion. *Computers in Human Behavior*, *99*, 219–227. https://doi.org/10.1016/j.chb.2019.05.022

Kirchner, J., & Reuter, C. (2020). Countering Fake News: A Comparison of Possible Solutions Regarding User Acceptance and Effectiveness. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW2), 1–27. https://doi.org/10.1145/3415211

Klar, S., Leeper, T., & Robison, J. (2020). Studying Identities with Experiments: Weighing the Risk of Posttreatment Bias Against Priming Effects. *Journal of Experimental Political Science*, *7*, 56–60.

Klimiuk, K., Czoska, A., Biernacka, K., & Balwicki, Ł. (2021). Vaccine misinformation on social media – topic-based content and sentiment analysis of Polish vaccine-deniers' comments on Facebook. *Human Vaccines & Immunotherapeutics*, *17*(7), 2026–2035. https://doi.org/10.1080/21645515.2020.1850072

Koo, A. Z.-X., Su, M.-H., Lee, S., Ahn, S.-Y., & Rojas, H. (2021). What Motivates People to Correct Misinformation? Examining the Effects of Third-person Perceptions and Perceived Norms. *Journal of Broadcasting & Electronic Media*, *65*(1), 111–134. https://doi.org/10.1080/08838151.2021.1903896

Kumar, A. (2020). Trump says blame China. His supporters are listening. *Politico*. https://www.politico.com/news/2020/05/03/trump-supporters-china-226309

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R.,

Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Lee, C. S., & Ma, L. (2012). News sharing in social media: The effect of gratifications and prior experience. *Computers in Human Behavior*, *28*(2), 331–339. https://doi.org/10.1016/j.chb.2011.10.002

Lee, E.-J., & Tandoc, E. C. (2017). When News Meets the Audience: How Audience Feedback Online Affects News Production and Consumption: When News Meets the Audience. *Human Communication Research*, *43*(4), 436–449. https://doi.org/10.1111/hcre.12123

Leiner, D. J., Kobilke, L., Rueß, C., & Brosius, H.-B. (2018). Functional domains of social media platforms: Structuring the uses of Facebook to better understand its gratifications. *Computers in Human Behavior*, *83*, 194–203. https://doi.org/10.1016/j.chb.2018.01.042

Lerma, M., Marquez, C., Sandoval, K., & Cooper, T. V. (2021). Psychosocial Correlates of Excessive Social Media Use in a Hispanic College Sample. *Cyberpsychology, Behavior, and Social Networking*, cyber.2020.0498. https://doi.org/10.1089/cyber.2020.0498

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and Its Correction: Continued Influence and Successful Debiasing. *Psychological Science in the Public Interest*, *13*(3), 106–131. https://doi.org/10.1177/1529100612451018

Lewandowsky, S., & van der Linden, S. (2021). Countering Misinformation and Fake News Through Inoculation and Prebunking. *European Review of Social Psychology*, *32*(2), 348–384. https://doi.org/10.1080/10463283.2021.1876983

Lijphart, A. (2012). *Patterns of Democracy. Government Forms and Performance in Thirty-Six Countries.* Yale University Press.

Lopes, B., & Yu, H. (2017). Who do you troll and Why: An investigation into the relationship between the Dark Triad Personalities and online trolling behaviours towards popular and less popular Facebook profiles. *Computers in Human Behavior*, *77*, 69–76. https://doi.org/10.1016/j.chb.2017.08.036

Mahl, D., Zeng, J., & Schäfer, M. S. (2021). From “Nasa Lies” to “Reptilian Eyes”: Mapping Communication About 10 Conspiracy Theories, Their Communities, and Main Propagators on Twitter. *Social Media + Society*, *7*(2), 205630512110174. https://doi.org/10.1177/20563051211017482

Majid, G. M., & Pal, A. (2020). Conspiracy and Rumor Correction: Analysis of Social Media Users’ Comments. *2020 3rd International Conference on Information and Computer Technologies (ICICT)*, 331–335. https://doi.org/10.1109/ICICT50521.2020.00058

March, E., & Springer, J. (2019). Belief in conspiracy theories: The predictive role of schizotypy, Machiavellianism, and primary psychopathy. *PLOS ONE*, *14*(12), e0225964. https://doi.org/10.1371/journal.pone.0225964

Marchlewska, M., Cichocka, A., & Kossowska, M. (2018). Addicted to answers: Need for cognitive closure and the endorsement of conspiracy beliefs: Need for cognitive closure and conspiracy beliefs. *European Journal of Social Psychology*, *48*(2), 109–117. https://doi.org/10.1002/ejsp.2308

Mari, S., Gil de Zúñiga, H., Suerdem, A., Hanke, K., Brown, G., Vilar, R., Boer, D., & Bilewicz, M. (2021). Conspiracy Theories and Institutional Trust: Examining the Role of Uncertainty Avoidance and Active Social Media Use. *Political Psychology*, pops.12754. https://doi.org/10.1111/pops.12754

Martin, J. R., & White, P. R. R. (2005). *The Language of Evaluation: Appraisal in English.* Palgrave Macmillan.

Masullo, G. M., & Kim, J. (2021). Exploring “Angry” and “Like” Reactions on Uncivil Facebook Comments That Correct Misinformation in the News. *Digital Journalism*, *9*(8), 1103–1122. https://doi.org/10.1080/21670811.2020.1835512

H.R.3611—Algorithmic Justice and Online Platform Transparency Act, H. R. 3611, 117th Congress, 1st Session (2021).

McCombs, M. E., & Shaw, D. L. (1972). The Agenda-Setting Function of Mass Media. *Public Opinion Quarterly*, *36*, 176–187.

McGregor, S. C. (2020). "Taking the Temperature of the Room." *Public Opinion Quarterly*, *84*(S1), 236–256. https://doi.org/10.1093/poq/nfaa012

Meta. (2022). *Meta Reports Second Quarter 2022 Results*.

Metzger, M. J., Flanagin, A. J., Mena, P., Jiang, S., & Wilson, C. (2021). From Dark to Light: The Many Shades of Sharing Misinformation Online. *Media and Communication*, *9*(1), 134–143. https://doi.org/10.17645/mac.v9i1.3409

Miller, J. M., Saunders, K. L., & Farhart, C. E. (2016). Conspiracy Endorsement as Motivated Reasoning: The Moderating Roles of Political Knowledge and Trust: CONSPIRACY ENDORSEMENT AS MOTIVATED REASONING. *American Journal of Political Science*, *60*(4), 824–844. https://doi.org/10.1111/ajps.12234

Mimikama. (2019). *Fake: Falsche Bilder zur FridaysforFuture-Demo!* Fake: Falsche Bilder Zur FridaysforFuture-Demo! https://www.mimikama.at/aktuelles/fridaysforfuture-demo/

Min, S. J. (2021). Who Believes in Conspiracy Theories? Network Diversity, Political Discussion, and Conservative Conspiracy Theories on Social Media. *American Politics Research*, *49*(5), 415–427. https://doi.org/10.1177/1532673X211013526

Mitchell, A., & Oliphant, J. B. (2020). *Americans Immersed in COVID-19 News; Most Think Media Are Doing Fairly Well Covering It*. 9.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med*, *6*(7).

Mosinzova, V., Fabian, B., Ermakova, T., & Baumann, A. (2019). Fake News, Conspiracies and Myth Debunking in Social Media—A Literature Survey Across Disciplines. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3328022

Naab, T. K., Heinbach, D., Ziegele, M., & Grasberger, M.-T. (2020). Comments and Credibility: How Critical User Comments Decrease Perceived News Article Credibility. *Journalism Studies*, *21*(6), 783–801. https://doi.org/10.1080/1461670X.2020.1724181

Naab, T. K., & Küchler, C. (2021). Clarity (User Comments/Interactions). *DOCA - Database of Variables for Content Analysis*. https://doi.org/10.34778/5e

Nekmat, E., & Ismail, I. (2019). Issue-based micromobilization on social media: Mediated pathways linking issue involvement and self-network opinion congruity to expressive support. *Computers in Human Behavior*, *101*, 51–59. https://doi.org/10.1016/j.chb.2019.07.008

Newman, N., Fletcher, R., Nielsen, R. K., Kalogeropoulos, A., & Levy, D. A. L. (2017). *Reuters Institute Digital News Report 2017*. 136.

Newman, N., Fletcher, R., Robertson, C. T., Eddy, K., & Nielsen, R. K. (2022). *Reuters Institute Digital News Report 2022*. 164.

Newman, N., Fletcher, R., Schulz, A., Andı, S., & Nielsen, R. K. (2020). *Reuters Institute Digital News Report 2020*. 112.

Newman, N., Fletcher, R., Schulz, A., Andı, S., Robertson, C. T., & Nielsen, R. K. (2021). *Reuters Institute Digital News Report 2021*. 164.

Nielsen, R. K., & Graves, L. (2017). "News you don't believe": Audience perspectives on fake news. *Factsheet*, *Oxford*, 8.

Nyhan, B., & Reifler, J. (2010). When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior*, *32*(2), 303–330. https://doi.org/10.1007/s11109-010-9112-2

Obaidi, M., Kunst, J., Ozer, S., & Kimel, S. Y. (2021). The "Great Replacement" conspiracy: How the perceived ousting of Whites can evoke violent extremism and Islamophobia. *Group Processes*, 1–21.

Ognyanova, K., Lazer, D., Robertson, R. E., & Wilson, C. (2020). Misinformation in action: Fake news exposure is linked to lower trust in media, higher trust in government when your side is in

power. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-024

Oliphant, J., & Kahn, C. (2021). Half of Republicans believe false accounts of deadly U.S. Capitol riot-Reuters/Ipsos poll. *Reuters*. https://www.reuters.com/article/us-usa-politics-disinformation-idUSKBN2BS0RZ

Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2020). *Partisan polarization is the primary psychological motivation behind "fake news" sharing on Twitter* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/v45bk

Pantazi, M., Hale, S., & Klein, O. (2021). Social and Cognitive Aspects of the Vulnerability to Political Misinformation. *Political Psychology*, *42*(S1), 267–304. https://doi.org/10.1111/pops.12797

Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, *6*(2), 259–283. https://doi.org/10.1177/1461444804041444

Papakyriakopoulos, O., Medina Serrano, J. C., & Hegelich, S. (2020). The spread of COVID-19 conspiracy theories on social me-dia and the effect of content moderation. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-034

Pasquetto, I., Swire-Thompson, B., & Amazeen, M. A. (2020). Tackling misinformation: What researchers could do with social media data. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-49

Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality*, *36*(6), 556–563. https://doi.org/10.1016/S0092-6566(02)00505-6

Pena, M. M., Klemfuss, J. Z., Loftus, E. F., & Mindthoff, A. (2017). The effects of exposure to differing amounts of misinformation and source credibility perception on source monitoring and memory accuracy. *Psychology of Consciousness: Theory, Research, and Practice*, *4*(4), 337–347. https://doi.org/10.1037/cns0000137

Pennycook, G., Bear, A., Collins, E. T., & Rand, D. G. (2020). The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines Without Warnings. *Management Science*, *66*(11), 4944–4957. https://doi.org/10.1287/mnsc.2019.3478

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*. https://doi.org/10.1038/s41586-021-03344-2

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

Pennycook, G., & Rand, D. G. (2021). The Psychology of Fake News. *Trends in Cognitive Sciences*, *25*(5), 388–402. https://doi.org/10.1016/j.tics.2021.02.007

Petersen, M. B., Osmundsen, M., & Arceneaux, K. (2018). *The "Need for Chaos" and Motivations to Share Hostile Political Rumors* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/6m4ts

Pew Research Center. (2021). *More Americans now say government should take steps to restrict false information online than in 2018*. Pew Research Center. https://www.pewresearch.org/fact-tank/2021/08/18/more-americans-now-say-government-should-take-steps-to-restrict-false-information-online-than-in-2018/

Pickles, K., Cvejic, E., Nickel, B., Copp, T., Bonner, C., Leask, J., Ayre, J., Batcup, C., Cornell, S., Dakin, T., Dodd, R., Isautier, J. M., & McCaffery, K. J. (2020). *COVID-19: Beliefs in misinformation in the Australian community* [Preprint]. Public and Global Health. https://doi.org/10.1101/2020.08.04.20168583

Poell, T., Nieborg, D., & van Dijck, J. (2019). Platformisation. *Internet Policy Review*, *8*(4). https://doi.org/10.14763/2019.4.1425

Politifact. (2022). *A video taken from a car driving through Bucha, Ukraine, shows a corpse "moving his arm," and then "in the rear view mirror the 'corpse' sits down."*

https://www.politifact.com/factchecks/2022/apr/04/russian-ministry-defense/russia-pushes-false-crisis-actor-claims-about-vide/

Pop, M.-I. (2019). Influence of the educational level on the spreading of Fake News regarding the energy field in the online environment. *Proceedings of the 13th International Conference on Business Excellence 2019*, *13*, 1108–1117.

Poynter. (2020). *Here's what we know about the source of the coronavirus pandemic*. Here's What We Know about the Source of the Coronavirus Pandemic. https://www.poynter.org/fact-checking/2020/heres-what-we-know-about-the-source-of-the-coronavirus-pandemic/

Prochazka, F., Weber, P., & Schweiger, W. (2018). Effects of Civility and Reasoning in User Comments on Perceived Journalistic Quality. *Journalism Studies*, *19*(1), 62–78. https://doi.org/10.1080/1461670X.2016.1161497

Rauthmann, J. F., & Kolar, G. P. (2013). The perceived attractiveness and traits of the Dark Triad: Narcissists are perceived as hot, Machiavellians and psychopaths not. *Personality and Individual Differences*, *54*(5), 582–586. https://doi.org/10.1016/j.paid.2012.11.005

Reilly, I. (2012). Satirical Fake News and/as American Political Discourse. *The Journal of American Culture*, *35*(3), 258–275. https://doi.org/10.1111/j.1542-734X.2012.00812.x

Ricard, J., & Medeiros, J. (2020). Using misinformation as a political weapon: COVID-19 and Bolsonaro in Brazil. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-013

Romer, D., & Jamieson, K. H. (2020). Conspiracy theories as barriers to controlling the spread of COVID-19 in the U.S. *Social Science & Medicine*, *263*, 113356. https://doi.org/10.1016/j.socscimed.2020.113356

Rooke, M. (2021). Alternative media framing of COVID-19 risks. *Current Sociology*, *69*(4), 584–602. https://doi.org/10.1177/00113921211006115

Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L. J., Recchia, G., van der Bles, A. M., & van der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science*, *7*(10), 201199. https://doi.org/10.1098/rsos.201199

Rösner, L., Winter, S., & Krämer, N. C. (2016). Dangerous minds? Effects of uncivil online comments on aggressive cognitions, emotions, and behavior. *Computers in Human Behavior*, *58*, 461–470. https://doi.org/10.1016/j.chb.2016.01.022

Rossini, P., Stromer-Galley, J., Baptista, E. A., & Veiga de Oliveira, V. (2020). Dysfunctional information sharing on WhatsApp and Facebook: The role of political talk, cross-cutting exposure and social corrections. *New Media & Society*, 146144482092805. https://doi.org/10.1177/1461444820928059

Rowe, I. (2015). Civility 2.0: A comparative analysis of incivility in online political discussion. *Information, Communication & Society*, *18*(2), 121–138. https://doi.org/10.1080/1369118X.2014.940365

Ruggiero, T. E. (2000). Uses and Gratifications Theory in the 21st Century. *Mass Communication and Society*, *3*(1), 3–37. https://doi.org/10.1207/S15327825MCS0301_02

Sängerlaub, A. (2017). Verzerrte Realitäten. Die Wahrnehmung von “Fake News” im Schatten der Bundestagswahl. *Politico*. https://www.politico.eu/wp-content/uploads/2017/10/Zusammenfassung-1.pdf

Schaeffer, K. (2020). A look at the Americans who believe there is some truth to the conspiracy theory that COVID-19 was planned. *Pew Research Center*, 5.

Scott, M. (2022). ‘Grotesque’ Russian disinfo campaign mimics Western news websites to sow dissent. *Politico*. https://www.politico.eu/article/russia-influence-ukraine-fake-news/

Seely, N. (2018). Virtual Vitriol: A Comparative Analysis of Incivility Within Political News Discussion Forums. *Electronic News*, *12*(1), 42–61. https://doi.org/10.1177/1931243117739060

Shin, J., & Thorson, K. (2017). Partisan Selective Sharing: The Biased Diffusion of Fact-Checking Messages on Social Media: Sharing Fact-Checking Messages on Social Media. *Journal of Communication*, *67*(2), 233–255. https://doi.org/10.1111/jcom.12284

Shu, K., Cui, L., Wang, S., Lee, D., & Liu, H. (2019). dEFEND: Explainable Fake News Detection. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 395–405. https://doi.org/10.1145/3292500.3330935

Starbird, K. (2017). Examining the Alternative Media Ecosystem through the Production of Alternative Narratives of Mass Shooting Events on Twitter. *Proceedings of the Eleventh International AAAI Conference on Web and Social Media*, 10.

Statista. (2022a). *Number of Instagram users worldwide from 2020 to 2025*. https://www.statista.com/statistics/183585/instagram-number-of-global-users/

Statista. (2022b). *Number of TikTok users worldwide from 2020 to 2025*. https://www.statista.com/statistics/1327116/number-of-global-tiktok-users/

Sternisko, A., Cichocka, A., Cislak, A., & Van Bavel, J. J. (2020). *Collective narcissism predicts the belief and dissemination of conspiracy theories during the COVID-19 pandemic* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/4c6av

Sterrett, D., Malato, D., Benz, J., Kantor, L., Tompson, T., Sonderman, J., & Loker, K. (2019). Who Shared It?: Deciding What News to Trust on Social Media. *Digital Journalism*, *7*(6), 783–801.

Stromer-Galley, J. (2007). Measuring Deliberation's Content: A Coding Scheme. *Journal of Deliberative Democracy*, *3*(1). https://doi.org/10.16997/jdd.50

Stroud, N. J., Duyn, E. V., & Peacock, C. (2016). News Commenters and News Comment Readers. *Engaging News Project*, 21.

Stubenvoll, M., Heiss, R., & Matthes, J. (2021). Media Trust Under Threat: Antecedents and Consequences of Misinformation Perceptions on Social Media. *International Journal of Communication*, *15*, 2765–2786.

Summers, J. S. (2017). Post hoc ergo propter hoc: Some benefits of rationalization. *Philosophical Explorations*, *20*, 21–36. https://doi.org/10.1080/13869795.2017.1287292

Sunstein, C. R., & Vermeule, A. (2008). Conspiracy Theories. *University of Chicago Public Law & Legal Theory Working Paper No. 199*, 30.

Sykora, M., Elayan, S., & Jackson, T. W. (2020). A qualitative analysis of sarcasm, irony and related #hashtags on Twitter. *Big Data & Society*, *7*(2), 205395172097273. https://doi.org/10.1177/2053951720972735

Talwar, S., Dhir, A., Kaur, P., Zafar, N., & Alrasheedy, M. (2019). Why do people share fake news? Associations between the dark side of social media use and fake news sharing behavior. *Journal of Retailing and Consumer Services*, *51*, 72–82. https://doi.org/10.1016/j.jretconser.2019.05.026

Tan, Y., & Weaver, D. H. (2007). Agenda-Setting Effects among the Media, the Public, and Congress, 1946–2004. *Journalism & Mass Communication Quarterly*, *84*(4), 729–744. https://doi.org/10.1177/107769900708400405

Tandoc, E. C., Lim, D., & Ling, R. (2020). Diffusion of disinformation: How social media users respond to fake news and why. *Journalism*, *21*(3), 18.

Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining "Fake News": A typology of scholarly definitions. *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Tang, W. Y., Reer, F., & Quandt, T. (2022). The interplay of the Dark Triad and social media use motives to social media disorder. *Personality and Individual Differences*, *187*, 111402. https://doi.org/10.1016/j.paid.2021.111402

Tee, W. J., & Murugesan, R. K. (2018). Trust Network, Blockchain and Evolution in Social Media to Build Trust and Prevent Fake News. *2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA)*, 1–6. https://doi.org/10.1109/ICACCAF.2018.8776822

Thompson, N., Wang, X., & Daya, P. (2019). Determinants of News Sharing Behavior on Social Media. *Journal of Computer Information Systems*, 1–9. https://doi.org/10.1080/08874417.2019.1566803

Thorson, K., Vraga, E., & Ekdale, B. (2010). Credibility in Context: How Uncivil Online Commentary Affects News Credibility. *Mass Communication and Society*, *13*(3), 289–313. https://doi.org/10.1080/15205430903225571

Tran, G. H., & Ngo, X. M. (2018). News comments on facebook – a systemic functional linguistic analysis of moves and appraisal language in reader-reader interaction. *Journal of World Languages*, *5*(1), 46–80. https://doi.org/10.1080/21698252.2018.1504856

Tran, T., Rad, P., Valecha, R., & Rao, H. R. (2020). Misinformation in Crises: A Conceptual Framework for Examining Human-Machine Interactions. *2020 IEEE / ITU International Conference on Artificial Intelligence for Good (AI4G)*, 46–50. https://doi.org/10.1109/AI4G50087.2020.9311010

Tucker, J., Guess, A., Barbera, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., & Nyhan, B. (2018). Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3144139

Twitter. (2022). *Twitter Announces Second Quarter 2022 Results*.

Uscinski, J. E., Klofstad, C., & Atkinson, M. D. (2016). What Drives Conspiratorial Beliefs? The Role of Informational Cues and Predispositions. *Political Research Quarterly*, *69*(1), 57–71. https://doi.org/10.1177/1065912915621621

Uscinski, J. E., & Olivella, S. (2017). The conditional effect of conspiracy thinking on attitudes toward climate change. *Research & Politics*, *4*(4), 205316801774310. https://doi.org/10.1177/2053168017743105

Van Aelst, P., Strömbäck, J., Aalberg, T., Esser, F., de Vreese, C., Matthes, J., Hopmann, D., Salgado, S., Hubé, N., Stępińska, A., Papathanassopoulos, S., Berganza, R., Legnante, G., Reinemann, C., Sheafer, T., & Stanyer, J. (2017). Political communication in a high-choice media

environment: A challenge for democracy? *Annals of the International Communication Association*, *41*(1), 3–27. https://doi.org/10.1080/23808985.2017.1288551

Van Bavel, J. J., Harris, E. A., Pärnamets, P., Rathje, S., Doell, K. C., & Tucker, J. A. (2021). Political Psychology in the Digital (mis)Information age: A Model of News Belief and Sharing. *Social Issues and Policy Review*, *15*(1), 84–113. https://doi.org/10.1111/sipr.12077

Van Bavel, J. J., & Pereira, A. (2018). The Partisan Brain: An Identity-Based Model of Political Belief. *Trends in Cognitive Sciences*, *22*(3), 213–224. https://doi.org/10.1016/j.tics.2018.01.004

van der Linden, S., Panagopoulos, C., Azevedo, F., & Jost, J. T. (2021). The Paranoid Style in American Politics Revisited: An Ideological Asymmetry in Conspiratorial Thinking. *Political Psychology*, *42*(1), 23–51. https://doi.org/10.1111/pops.12681

Van Duyn, E. V., Peacock, C., & Stroud, N. J. (2021). The Gender Gap in Online News Comment Sections. *Social Science Computer Review*, *39*(2), 181–196.

van Koningsbruggen, G. M., Hartmann, T., Eden, A., & Veling, H. (2017). Spontaneous Hedonic Reactions to Social Media Cues. *Cyberpsychology, Behavior, and Social Networking*, *20*(5), 334–340. https://doi.org/10.1089/cyber.2016.0530

van Mulukom, V., Pummerer, L., Alper, S., Bai, H., Cavojova, V., Farias, J. E. M., Kay, C. S., Lazarevic, L. B., Lobato, E. J. C., Marinthe, G., Pavela Banai, I., Šrol, J., & Zezelj, I. (2020). *Antecedents and consequences of COVID-19 conspiracy beliefs: A systematic review* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/u8yah

van Prooijen, J.-W., Krouwel, A. P. M., & Pollet, T. V. (2015). Political Extremism Predicts Belief in Conspiracy Theories. *Social Psychological and Personality Science*, *6*(5), 570–578. https://doi.org/10.1177/1948550614567356

van Prooijen, J.-W., Ligthart, J., Rosema, S., & Xu, Y. (2021). The entertainment value of conspiracy theories. *British Journal of Psychology*, bjop.12522. https://doi.org/10.1111/bjop.12522

Vogels, E. A., Auxier, B., & Anderson, M. (2021). Partisan differences in social media use show up for some platforms, but not Facebook. *Pew Research Center*, 7.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Vraga, E. K., & Bode, L. (2017). Using Expert Sources to Correct Health Misinformation in Social Media. *Science Communication*, *39*(5), 621–645. https://doi.org/10.1177/1075547017731776

Vraga, E. K., & Bode, L. (2018). I do not believe you: How providing a source corrects health misperceptions across social media platforms. *Information, Communication & Society*, *21*(10), 1337–1353. https://doi.org/10.1080/1369118X.2017.1313883

Vraga, E. K., & Bode, L. (2020). Defining Misinformation and Understanding its Bounded Nature: Using Expertise and Evidence for Describing Misinformation. *Political Communication*, *37*(1), 136–144. https://doi.org/10.1080/10584609.2020.1716500

Vraga, E. K., Bode, L., & Tully, M. (2020). Creating News Literacy Messages to Enhance Expert Corrections of Misinformation on Twitter. *Communication Research*, 009365021989809. https://doi.org/10.1177/0093650219898094

Waddell, T. F. (2018). What does the crowd think? How online comments and popularity metrics affect news credibility and issue importance. *New Media & Society*, *20*(8), 3068–3083. https://doi.org/10.1177/1461444817742905

Waddell, T. F. (2020). The Authentic (And Angry) Audience: How comment authenticity and sentiment impact news evaluation. *Digital Journalism*, *8*(2), 249–266. https://doi.org/10.1080/21670811.2018.1490656

Wanta, W., Golan, G., & Lee, C. (2004). Agenda Setting and International News: Media Influence on Public Perceptions of Foreign Nations. *J&MC Quarterly*, *81*(2), 364–377.

Wardle, C., & Derakhshan, H. (2018). Thinking about 'information disorder': Formats of misinformation, disinformation, and mal-information. In C. Ireton (Ed.), *Journalism, fake news & disinformation: Handbook for journalism education and training* (pp. 43–54). UNESCO.

Weeks, B. E., & Gil de Zúñiga, H. (2021). What's Next? Six Observations for the Future of Political Misinformation Research. *American Behavioral Scientist*, *65*(2), 277–289. https://doi.org/10.1177/0002764219878236

Young, J. C. (2021). Disinformation as the weaponization of cruel optimism: A critical intervention in misinformation studies. *Emotion, Space and Society*, *38*, 100757. https://doi.org/10.1016/j.emospa.2020.100757

Zhang, Y., Wang, L., Zhu, J. J. H., & Wang, X. (2021). Conspiracy vs science: A large-scale analysis of online discussion cascades. *World Wide Web*, *24*(2), 585–606. https://doi.org/10.1007/s11280-021-00862-x

Zivnuska, S., Carlson, J. R., Carlson, D. S., Harris, R. B., & Harris, K. J. (2019). Social media addiction and social media reactions: The implications for job performance. *The Journal of Social Psychology*, *159*(6), 746–760. https://doi.org/10.1080/00224545.2019.1578725

Zollo, F., Novak, P. K., Del Vicario, M., Bessi, A., Mozetič, I., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015). Emotional Dynamics in the Age of Misinformation. *PLOS ONE*, *10*(9), e0138740. https://doi.org/10.1371/journal.pone.0138740

Authors' contributions

Chapter Two: Identifying the Drivers Behind the Dissemination of Online Misinformation

Sophie Morosoli (first author): conception of the study (60%), data collection (100%), data analysis and interpretation (80%), writing of research paper (80%)

Peter Van Aelst (co-author): conception of the study (40%), data collection (100%), data analysis and interpretation (20%), writing of research paper (20%)

Edda Humprecht (co-author): conception of the study (40%), data collection (100%), data analysis and interpretation (20%), writing of research paper (20%)

Anna Staender (co-author): conception of the study (0%), data collection (100%), data analysis and interpretation (0%), writing of research paper (0%)

Frank Esser (co-author): conception of the study (0%), data collection (100%), data analysis and interpretation (0%), writing of research paper (0%)

Chapter Three: To Convince, to Provoke or to Entertain? A Study on Individual Motivations behind Engaging with Conspiracy Theories Online

Sophie Morosoli (first author): conception of the study (90%), data collection (100%), data analysis and interpretation (40%), writing of research paper (80%)

Peter Van Aelst (co-author): conception of the study (10%), data collection (100%), data analysis and interpretation (10%), writing of research paper (20%)

Patrick van Erkel (co-author): conception of the study (0%), data collection (0%), data analysis and interpretation (60%), writing of research paper (10%)

Chapter Four: "Come on people, this is clearly fake" How Social Media Users React to Political Online Misinformation by Commenting

Sophie Morosoli (single author)

Chapter Five: Who Approves, Disapproves or Ignores Online Misinformation? A Study on Individual Characteristics in Connection with Supporting and Renouncing Online Misinformation

Sophie Morosoli (single author)