Eindhoven University of Technology

MASTER

Behind the Political Frame

Exploring Visualization Methods to Increase Bias Awareness and Mitigate Framing Effects on Social Media

Ruyters, Wies J.H.

Award date:
2022

Link to publication

Disclaimer
This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

Department of Mathematics and Computer Science

Behind the Political Frame: Exploring Visualization Methods to Increase Bias Awareness and Mitigate Framing Effects on Social Media

Master Thesis

Wies Ruyters

Supervisor:
Huub van de Wetering

Thursday 1st September, 2022

Eindhoven, September 2022

Abstract

Social media posts from politicians are biased, i.e. strongly in favor of one group or argument, and framed, i.e. selected, emphasised and presented. Although bias and framing do not imply falsehood, they do shape the receiver's understanding of information, as politicians typically choose the most salient perspective to communicate their arguments effectively. These biases and frames obfuscate the versatility of policy issues, particularly as they are featured in homogeneous and personalized interfaces of social media users, known as filter bubbles. Literature indicated that information visualization can potentially increase political bias awareness, and mitigate political framing effects [1, 44, 46, 57, 59]. However, studies to date have mainly focused on news articles or long text documents, for which reason their findings are hard to generalize to the social media domain, where information needs to be fitted in shorter pieces of text, and often include other types of multimedia (e.g., hyperlinks).

The current study aimed to bridge this knowledge gap, by validating the impact of information visualization in a simulated social media context, based on Twitter. The data consisted of microblogs (known as tweets) from politicians and news outlets. Three types of visualization methods, aimed at increasing political bias awareness and mitigation of framing effects, were designed: 1) a static, context-dependent visualization on tweet-level, 2) frame opposition between political and news tweets, and 3) an interactive dashboard visualization that displayed aggregate data over time. The visualizations were tested in a user study (N=21). The results on the impact of visualization type 1) showed that participants' political bias awareness had increased. However, neither a mitigated framing effect was visible in the results on the impact of visualization type 2), nor the added value of visualization type 3). This report discusses the process and results of the current study, and concludes with suggestions for future research.

Acknowledgements

It has been my dream to perform research to the impact of information visualization on people's perception of political social media content. I feel grateful for the opportunity I got at Eindhoven University of Technology to realize this dream with my master thesis. First, my sincere thanks and appreciation go to my supervisor Huub, for his valuable guidance throughout the process. His critical questions and constructive feedback were of outstanding help, and our weekly meetings always provided me with interesting and important insights. Second, I would like to thank Martijn for sharing his expertise on experimental psychology when thinking along with me about the design of the experiment and the construction of the questionnaires. Third, I would like to thank my friend Tar for his help on crucial parts of the technical realization of the experiment interface. Fourth, my gratitude goes to all 21 participants that took part in the experiment. Fifth, I would like to thank my loved ones for their support and patience. Sixth and last, I feel deeply grateful for all the inspiring conversations I was able to have about the topic - either with family, friends, acquaintances or strangers. These conversations have made me believe in my research, and empowered me to make the most out of this project.

Contents

Chapter 1

Introduction

Although misinformation, propaganda and (political) media bias are nothing ‘new’, the number of possible channels through which they are communicated has grown enormously with the birth and development of social media. Social media facilitate politicians with a communication channel through which they communicate directly with their audience. Next to direct political communication, social media allow them to disclose personal information which serves several campaign purposes [1, 2]. In doing so, politicians bypass traditional media, which some of them claim to be ‘corrupt’, ’non-representative’ or ‘fake’ [2, 3, 4]. The users of these platforms, referred to as *users*, receive politician’s messages in a digital, personalized interface, i.e. the *feed*. Such a feed features content in the form of messages, microblogs, pictures, videos or other media as *posts*. Next to posts from politicians, a user’s feed is typically complemented with posts from friends, family, celebrities and news outlets. Users can *engage* by means of *liking* and *sharing* - without the source having to show any proof of correctness. Contrary to traditional media channels, e.g. newspapers, social media platforms are not liable for the content created by their users [5]. Not surprisingly, it was found that precisely these social media channels enable the spread of fake news, misinformation and disinformation at a rate that is faster, deeper and more broadly[1] than truthful content in all categories of information [6]. With a growing number of people relying on online platforms and social media for accessing news and information, the spread of false and misleading content imposes a critical problem - namely, not all users have the necessary skill to distinguish true from false [7].

1.1 The political discourse on social media

1.1.1 Categorization and context

Misinformation, disinformation, and fake news

Evaluated upon the intention with which false information is created and disseminated, three types can be defined: misinformation, disinformation, and fake news. Misinformation is generally referred to as false information that is mistaken for truthful. Disinformation is a type of misinformation, yet created and spread with the intention to deceive and cause harm. Fake news is in turn defined as a type of misinformation, yet specifically in the form of news articles, however, published and propagated through media regardless of the means or motives behind it [8, 9]. The intersection of politics, misinformation and social media has convoluted the public debate and political landscape on a global level. The impact of misinformation on democracy has become more apparent as it played a key role in recent events [10].

[1] The study by Vosoughi et al. researched Twitter data on misinformation dissemination. In Twitter, a post is called a *tweet* and a share of a post is a *retweet* it. The researchers defined *depth* as the number of separate retweets an original tweet got, *size* as the number of involved users over time, and *breadth* as the maximum number of users involved at any depth.

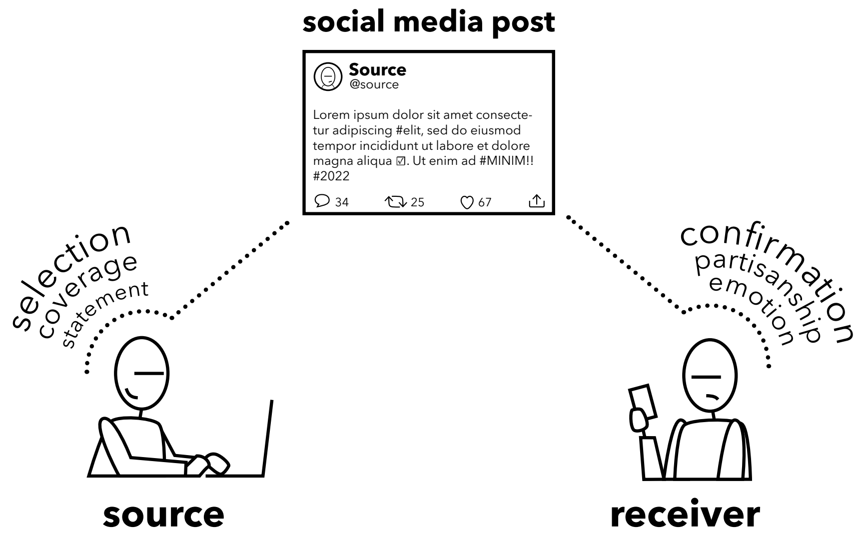


Figure 1.1: The different biases that occur on the source's side when sharing information, and on the receiver's side when interpreting information.

Bias

The Oxford Dictionary defines bias as a strong feeling in favor of or against one group of people or one side in an argument, often not based on fair judgement [11]. Bias, therefore, does not imply falseness. In the context of social media, a so-called *media bias* is present in the content itself (coming from the *source*) and the user's interpretation thereof (the *receiver*). The two main biases from the source's perspective are the choice of actually reporting something, i.e. selection bias, the completeness of information, i.e. coverage bias, and the way information is formulated, i.e. statement bias [12]. The biases that influence the receiver are cognitive biases. In studying user susceptibility to (political) misinformation and media bias, the most widely researched cognitive biases are 1) confirmation bias and bandwagon effect, 2) motivated reasoning and partisan bias, and 3) emotional bias and personality factors. First, theory on confirmation bias suggests that people prefer to be exposed to information that supports their existing views. This bias is often mentioned in line with selective exposure [13]. The bandwagon effect is a theory derived hereupon that explains user's inclination to seek proximity with others that uphold similar worldviews [14]. The concept of confirmation bias roots in cognitive dissonance theory, where cognitive dissonance describes the discomfort of holding two conflicting thoughts, caused by a misalignment of rationale and behavior [15, 16]. Second, in judging whether information is true or false, the reliance on personal political attitudes and/or partisanship is known as motivated reasoning or partisan bias [17, 18]. Third, emotional bias is described as the role of emotional processing in (mis)belief [19]. Studies on personality factors attempt to find if, and if so, what personality factors in news consumption relate to one's truth discernment capabilities and bias susceptibility [20]. As one user may be subject to multiple biases simultaneously, where some biases are more present than others, individual differences do occur. Figure 1.1 exemplifies the described source and receiver biases that come with social media posts.

Emphasis and value framing

In reporting truthful information, the selection, emphasis and presentation of information shape the receiver's understanding of a situation. This is known as *framing* [21]. In the political domain, a *framing effect* occurs when presenting the same political issue or common problem in a different way to alter citizen's attitudes, emotions or behavior [22]. Both the exposure of users to incorrect information, as well as framed content, can impact user's perception on reality, and even their emotional response [23]. In politics, *emphasis framing* is the idea that issues can be presented from multiple perspectives, and depending on which of those perspectives is highlighted in the frame, citizens tend to follow the more salient issue perspective when interpreting the issue [24]. In case this emphasis frame highlights well-known and culturally shared political values, it can be considered a *value frame* [24]. As framing does not require more than selecting a convincing aspect of an issue to make it more salient, it can be considered an effective political communication tool. With the shift of the political discourse to social media platforms, where short, rather than long, pieces of information have set the norm, the popularity of emphasis and value framing amongst politicians is logically explained.

1.1.2 Online platform dynamics

As the main interface from where a user consumes content on social media is the personal feed, the construction of these feeds plays a central role in shaping the user's sense of reality. Social networks typically enable one-to-one and one-to-many connections, where the former is mutual and the latter allows users to follow any other user, without that other user necessarily having to follow back. Therefore, users are to a large extent able to choose which connections to include in their network. The included content on a user's feed, however, does not only depend on these (often personal) connections. Platforms curate content algorithmically with the help of affective feedback loops that fetch a user's interest and suggest desired content in return [25]. Nonetheless, users are more likely to engage with content that is in line with their preexisting beliefs [26]. Users are continuously encouraged by social media platforms to engage with content coming from other users by means of liking, sharing and replying. Those forms of engagement reflect users' emotions about what they see. Since a platform's profit typically depends on how much time users spend there, social media are incentivized to curate feeds that are as engaging as possible - with content uniformity as a result. This phenomenon is known in literature as the *filter bubble*, i.e. the potential of online personalization to isolate people from diverse content and views [27]. The resulting reinforcement of a shared narrative, opinion or belief because of the interactions with like-minded users, is described as the *echo chamber* [28]. In political context, these notions entail situations where user's beliefs are strengthened, regardless of them being truthful [27].

1.1.3 Post-truth

The involvement of emotion in politics that is witnessed today is not novel, as emotion has proven to play a key role in many historical events - crises in particular. It is yet the shift in awareness and recognition of emotion as a determining political factor that distinguishes the "Post-truth" society seen as of this moment [29]. *Post-truth* concerns the societal situation in which emotion and personal beliefs form the public opinion rather than objective facts. It is partially driven by *truthiness*, which refers to a personal feeling of something to be true, even when it is not necessarily true. Politicians are able to anticipate on this phenomenon with *computational propaganda*, defined as the assemblage of social media platforms, autonomous agents, and big data tasked with the manipulation of public opinion [30, 29]. The embedding of affective feedback loops is to a certain extent able to amplify these developments, as the underlying recommender systems rely on emotional responses coming from users that result from the interactions that they have on social media platforms.

1.1.4 Today's debates on reality

A primary example of the impact of political misinformation is the 2016 U.S. presidential election, where a consultancy company called Cambridge Analytica was involved in the campaign of republican candidate Trump. Personal data of millions of Facebook users was used to microtarget swaying voters on the platform with fake and malicious content [31]. This form of *computational propaganda* is claimed to have determined Trump's election at the time, and also played a crucial role in the British Brexit referendum [30, 29, 10]. A second example is the spread of misinformation and conspiracy theories about the coronavirus during 2020's pandemic, therefore often referred to as "infodemic" [32]. Claims about the origin and severity of the coronavirus and the possible measures against it spread rapidly, where those claims lacked any proof. This led a non-negligible minority to distrust governments, not live up to imposed social distancing measures, and refuse getting vaccinated [33]. A third example is the Capitol Hill riot by supporters of former U.S. president Trump, whom himself claimed that the victory of his opponent Biden was stolen rather than won, as he repeatedly stated on his media channels. During his presidency, Trump's Twitter was famous for dominating the political agenda. However, his reluctant relationship with facts and posed provocations relating to the election results caused the platform to take measures. As of January 2021, Trump was banned permanently [34]. A last example is the military invasion of Ukraine by the Russian army in 2022. The invasion unleashed a secondary "war on information" which effects reached beyond the borders of these two nations; it polarized the geopolitical debate. Back and forth, those at war accuse each other of spreading misleading and incorrect claims, making it difficult for outsiders to know what is actually happening on the battleground [35].

1.1.5 Detection and correction of misinformation

Prebunking versus debunking

With any social media user being able to create and disseminate content, and any other user being able to consume and share that content, social media platforms became notorious for their *freedom of speech* and *freedom of reach*. The interplay of biases at the source and the receiver's side make that content is not always interpreted the way it was intended. Whereas platforms have no legal liability over the content created or shared on their service, several institutions have requested them to take more responsibility in warning for and taking down misinformation, misleading content and hate speech. Their reply to this request became apparent during the COVID-19 pandemic, where platforms as Facebook, Twitter and Spotify added warnings to content concerning the coronavirus or the available vaccinations against it [36, 37, 38]. Warnings are forms of *prebunking*, i.e. a preventive approach that seeks to help people recognize and resist misinformation they might encounter in the future [33]. The intervention of these platforms showed that limits to freedom of speech and freedom of reach do exist. The pressure on social media platforms to act has created a demand for automated bias and misinformation detection techniques. The detection and removal of malicious content is a form of *debunking* - a reactive intervention that corrects for or deletes misinformation after it was published [33]. In research thereto, four components were found to characterize online fake news; the creator, target victim, news content, and social context [39]. Detection approaches are focused on one or a combination of these characteristics, where the content-based approach includes the algorithmic analysis of semantic features of text. Emotion and subjectivity play an important part in those textual analyses [9, 40, 41].

In-time: labelling and visualization

The limitation of both prebunking and debunking is their misalignment in time with respect to the user's content consumption. Prebunking lacks tailoring to the information in question, and debunking comes after the malicious content has already spread. As a successive step to detection, platforms and researchers have started to explore in-time information visualization to help users become more resilient to misinformation and bias. Fact-check labels provide a first example. It was found that the strength and specificity of these labels determine their impact - the stronger the claim of an article to be false, and the more detailed the justification of that claim, the better able users were to discern truth from fake headlines after they were exposed to the label [42]. In turn, information visualization approaches can focus on the spreader of misinformation (i.e., the source), the content itself, or a combination of both. With respect to source visualization, graph theory is a well-established field of research that enables the visualization of network interactivity, in which echo chambers can be detected [43]. With respect to content visualization, several approaches have been proposed to make implicit bias and falseness in information more explicit. A first example is highlighting biased language in text, which was found to increase bias awareness compared to a control group [44]. A second example is the widely studied WordCloud visualization. WordClouds extract the most common words and depict them in different sizes and angles that are proportional to their popularity in the original text, thereby disrupting the text's linearity. They were proven to enhance people's understanding when presented in parallel with the original text [21]. In the political context, other examples can be found in visualizations of televised debates, where expressions of politicians are put in context and fact-checked to enhance a user's understanding [45, 46].

1.2 Problem Statement

At the foundation of democracy is its assumption of a shared sense of reality [47]. The disagreements between politicians - representatives of a democracy - on what is factual and fictional, along with an increased reliance on emotion in the interpretation of information, described in literature as *post-truth*, obfuscates this reality [29]. Politicians have anticipated the popularity of social media platforms as a way to communicate directly with their audience, yet are not legally accountable for disseminating false or misleading content. Fact is that any reported form of information, regardless of who may be the source, is biased; it was selected, covered and framed. However, the emphasis and value frames that politicians themselves incorporate in their communication, only state one aspect of an issue, and can therefore be misleading. Figure 1.2 provides with a characterization of the different forms of political discourse on social media. First, a distinction can be made between blatant *falsehoods* in the form of misinformation, disinformation and fake news, and more subtle *biased truths*, in the form of news frames (i.e., reported information) by news media. Value and emphasis frames are seen as a type of content that falls within these two categories, as they are not false by definition, yet, significantly deform information. Together with the homogenization of content as a result of social media platform dynamics and the increased reliance on emotion in the post-truth society, politicians contribute to shape users' *own* reality, rather than a commonly and culturally shared reality by including these frames. This constitutes the following problem statement:

By incorporating emphasis and value frames in their communication on social media, politicians obfuscate the versatility of issues and shape users' interpretation of reality in favor of the perspectives they choose to include in their frames.

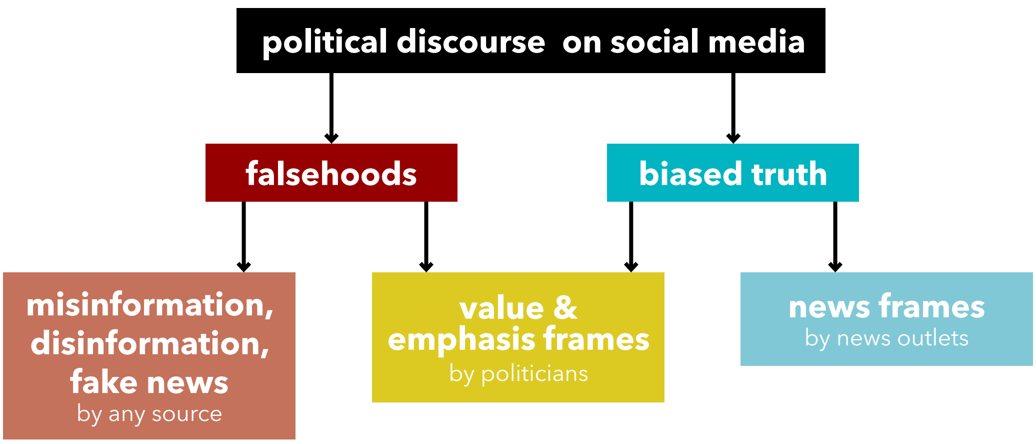


Figure 1.2: The different types of political discourse on social media.

Information visualization has been explored to warn and inform users about false and misleading content. Solutions explored to date range from very basic, e.g. labels, to rather sophisticated, e.g. text visualization. However, when concentrating on the mitigation of framing effects on *social media* with information visualization, little empirical evidence can be found. This has led to the formulation of the following research question:

How can information visualization increase political bias awareness and help mitigate the political framing effect on social media?

In light of this research question, the *political framing effect* is defined as the presentation of the same political issue or common problem in a different way to alter citizen's attitudes, emotions or behavior [22]. A *mitigated* political framing effect is seen as an increased sense of awareness at the user about the versatility of policy issues and the bias of the source as result of an intervention. Here, the intervention is considered to be the exposure to information visualization.

Research approach The approach of the current study is threefold. First, a literature study was done in three research domains, them being 1) psychological drivers promoting belief in misinformation, 2) misinformation identification, and 3) current state of the art in visualization methods for increased bias awareness and framing effect mitigation. Second, a data and task abstraction were executed to arrive at requirements and designs for the visualization that is to be tested. Third, the impact of these visualizations was tested in a qualitative, mixed-method user study (N=21). The social media setting for the current study was based on Twitter. Twitter is a social media platform that is to-date popular amongst politicians - and their audience - worldwide, and with its 280 character limit on posts, referred to as a *tweets*, users are forced to frame their messages wisely. Twitter therefore seemed a logical choice for a representative social media platform on which political framing takes place.

Chapter 2

Related Work

This chapter summarizes the main findings that followed from the literature review that was set up for the current study. This literature review had three focus domains; 1) understand what psychological drivers promote belief in misinformation, 2) identify characteristics of misinformation and misleading content online, and 3) map the current state of the art in information visualization aimed at creating more bias awareness or at mitigating political framing effects. The findings obtained in the first domain should make clear what psychological factors the information visualization, which is to be tested in the user study, needs account for in attempt to influence participants' interpretation of political tweets for the better. A literature review within the second domain should provide with an overview of what textual or contextual features are contained in misinformative or misleading social media posts, and hence, which features - that are currently invisible and implicit on social media - could be visualized with information visualization in order to contribute to a participant's misinformation, bias and framing resilience. The literature study within the third domain should provide with insights on which information visualization methods have already been developed, which results they have achieved, and which suggestions were done by their researchers that can be taken into account for the current study.

2.1 Psychological factors promoting belief in misinformation

This section highlights four psychological factors that were found to influence one's interpretation and understanding of information; 1) cognitive biases, 2) analytical and critical thinking, 3) emotion and personality factors, and 4) feed diversification. Every subsection is concluded with a statement on how the user study, that is part of the current study, may account for the concerning factor.

2.1.1 Cognitive biases

The way presented information is processed is subject to cognitive biases. In the context of online misinformation, two forms of cognitive biases stand out in terms of research coverage; 1) repeated exposure, and 2) motivated reasoning.

Repeated exposure

First, theory on repeated exposure suggests that the more a claim is repeated, the more familiar it becomes, hence the easier it is to process. This "Illusory truth" effect states that repeated claims are more likely to be judged as true than novel claims [48]. This was also found to apply in the misinformation domain [49]. As fake news spreads deeper, faster and more broadly than the truth in all categories of information, a corrective measure, such as debunking, is therefore often ineffective: the falsehood outweighs the correction in cognitive fluency [6, 48]. This effect is not moderated by knowledge; studies have shown that even when topic knowledge is available, cognitive fluency is still more often relied on [50]. Translating this to the current study, repeated exposure is a factor impossible to test for, as it would require a longitudinal experiment set-up.

Motivated reasoning

Second, theory on motivated reasoning and confirmation bias suggests that one tends to overrate the accuracy of information that is consistent with one's preexisting beliefs or political attitudes. This happens as a result of naturally reasoning in the direction of those preexisting beliefs as cognitive goals [17, 48]. In other words, this theory argues that one believes what one wants to believe. Mixed evidence was found on whether this cognitive bias influences one's ability to discern fake from real news. Nonetheless, it was found that specific demographic groups may be more susceptible to be subject to this bias than others [49, 48]. Contrary to repeated exposure, motivated reasoning is not impossible to test for in the current study; including a self-reported question on a participant's political preference would allow to test for motivated reasoning on article-level. Given the contradictory findings in literature, and the small number of participants included in the current study, motivated reasoning will however not be of primary focus. However, a self-reported question will be included in the experiment, such that the experiment results can be associated to the participants' political preference. This allows to compare results that may be obtained in future work, in which participants identify with a different political preference.

2.1.2 Analytical and critical thinking

Social media have transformed information consumption behavior. When comparing a typical Twitter feed to a traditional newspaper, two things stand out in particular. First, the diversity of information is larger since a Twitter feed comprises tweets from not only news outlets, but also social contacts, politicians and possibly many other sources. Second, the information density of a tweet is significantly smaller than a journalistic article. As a result, the information consumption behavior people have adopted has become increasingly volatile. In what is known in literature as the "classical reasoning perspective", the way people consume information on social media does no longer activate analytical reasoning, but intuitive reasoning instead [48][1]. However, it is one's analytic reasoning capabilities[2] that were found to be negatively correlated with the perceived accuracy of fake news, and positively correlated with the ability to discern truth from falsehood [18]. On top of this, deliberation was found to reduce false accuracy ratings of untruthful headlines [17]. Interestingly, these findings were unrelated to a participant's partisanship - hence, dispute the motivated reasoning theory. In line with the classical reasoning perspective, it was found that the addition of cues that shift attention to accuracy or scrutiny in judging information can improve truth discernment capabilities - again, regardless of one's partisanship with respect to the article under inspection [49]. The current study allows to test what information visualizations may trigger one's analytical or critical thinking, within the volatile information consumption environment of a social network. Hence, this psychological factor is included in the research goals of the current study.

2.1.3 Emotion and personality factors

Beyond findings that relate to cognition, emotion and personality factors also appear to correlate with misinformation susceptibility. Reliance on emotions and, in turn, a stronger experience of these emotions, were found a predictor of greater belief in fake news [10]. However, this appears

[1]In cognition literature, intuitive reasoning and analytical reasoning are described as System 1 and System 2 respectively.

[2]In the study by Pennycook and Rand (2019), in line with several other studies in this field, one's analytic reasoning capabilities were measured with a Cognitive Reflection Test (CRT). This method poses mathematical questions where the intuitive answer is usually wrong.

to be specific to fake news only, as no similar evidence was found for greater belief in truthful information [10]. In contrast to the classical reasoning perspective, where no consistent evidence is found on the influence of partisanship, research in the field of emotion has found that the experiences of the two independent emotions of anger and fear do seem to correlate with political preference. In this case, anger encouraged motivated reasoning, which led to mistakenly believing misinformation that was in line with political party preference. Fear had an opposite effect in respect to to anger, as this emotion triggered to take the whole information environment into account when discerning truth from false, instead of only focusing on partisan beliefs [51]. When focusing on specific personality traits, it was found that agreeableness, conscientiousness and open-mindedness were positively related to political news discernment[3], contrary to extraversion, political conservatism and the number of hours participants consume political news, which were negatively related [20]. Although research to the visualization of emotions has made promising progress, little evidence was found on how it could be used for information-sparse content, such as tweets. Moreover, given the context that was set for the current study, a sophisticated model to measure emotions in tweets would have to be developed from scratch as no such model exists for Dutch language yet. Hence, the psychological factor of emotion was not included in the research goals for the current study.

2.1.4 Feed diversification

An individual user's feed composition relies on choices that were made in the dynamic context of a social media platform, and is therefore considered personalized. This personalization is twofold. On one side, it consists of self-selected personalization as a result of users choosing to encounter like-minded content, referred to as *selective exposure*. On the other hand, it is a consequence of pre-selected personalization, referred to as the *filter bubble*, which is driven by other actors on a platform - not necessarily with explicit consent of the user [52]. However, feed *homogenization* can be counterbalanced by feed *diversification*. This diversification can happen on two grounds, namely on topic diversity and multiperspectivity. Research to recommender systems revealed that participants were willing to consume viewpoint diverse news recommendations, for which it may be concluded that diverse news recommendations do not imply reading aversion as a result of cognitive bias [53]. However, when suggested diversified content within a filter bubble, i.e. one's usual feed, only one third of the participants opted the viewpoint-divergent article, compared to two thirds opting to read a viewpoint-concordant article [54]. Building upon these findings in diversified news recommendations, positioning effects were studied as moderators of one's inclination to choose diversified over homogenized content with respect to one's personal preferences. It was found that the positioning in content that follows a reading pattern plays a significant role in participants' article selection [55]. This finding was taken into account when designing the layout of the experiment interface included in the current study.

2.2 Misinformation identification

Although the psychological factors, as described in section 2.1, influence one's interpretation and understanding of a social media post, it is the post itself that contains the potential to deceive. Three primary characteristics of misinformation have been identified; the source, the content and its generated engagement [8]. Although defined for online misinformation in general, these characteristics are of particular interest in the domain of the current study. First, as the current study investigates the framing effect for tweets that come from politicians directly, source credibility relates to the politician's personal Twitter account. Studies have shown that participants use the source to evaluate the accuracy of a news headline, and are inclined to judge ideologically congruent sources to be less politically biased than ideologically incongruent sources [49, 56, 57]. With respect to content, fake news articles were found to have longer titles, more capitalized words, use

[3] Here, agreeableness and conscientiousness were related to lower perceived accuracy of misinformation. Open-mindedness, however, was related to less belief in misinformation, and greater perceived accuracy of true news.

less stop words, and are typically shorter in body text. Moreover, deceptive stories were shown to be lower in cognitive complexity, and an analysis on textual level features of fake news revealed that they include more self-referencing, negation statements, complaints, generalizing terms, more social and more temporal words [8]. Last, with respect to engagement, true news was found to have more neutral engagement (e.g., neutral replies to an article) compared to fake news, which was found to have more negative engagement (e.g., dislikes) [8].

2.3 Current state of the art in information visualization

In the research domain of information visualization, also known as InfoVis, several methods relating to social media dynamics, bias awareness and framing effect mitigation have been studied. However, most of these methods apply either to big data streams or information-dense documents, whereas the focus of the current study is to individual tweets, which contain no more than 280 characters. Therefore, this chapter discusses some information visualization methods which, despite of their different input types and applications, may still provide with valuable insights for the visualizations to be included in the current study.

2.3.1 Creating bias awareness

Information visualizations aimed at creating more media bias awareness have been studied on three levels: 1) general, independent warnings, 2) article specific, and 3) text specific. On the first level, research has shown that general forewarning messages increased media bias awareness [57]. However, more specific warnings proved to be more effective than general reminders [58]. On the second level, preliminary results of a qualitative study that tested users' experience with a bias aware news recommendation system, of which a snippet is depicted in figure 2.1, showed that users thought the bias indications were useful when browsing [59]. However, in studying the effects of political bias visualization, research found evidence that visualizing political classification of articles had no effect on the user's perceived bias in an article [57]. On the third level, research found that participants became more aware of bias when presented with bias-aware visualizations on text level, which, in this case, concerned textual annotations as shown in figure 2.2 [44].

2.3.2 Framing effect mitigation

Research to framing effect mitigation differs from research to bias awareness, since the framing effect is typically "quantified" in the sense that it is not dependent on self-reports, but based on the comparison of individual or group responses to certain conditions. A first example comes from an exploratory study to the influence that text visualization has in the impact of framing on the perception of political issues. The results indicate that exposure to text visualization can mitigate political framing effects, and moreover showed a transfer learning effect, which implied that the experiment participants were left uninfluenced by framing in subsequent events [21]. The featured text visualization in the study mentioned was a WordCloud, a visualization approach to disrupt a text linearity, as shown in figure 2.3. Second, similar to bias awareness research, warnings have also been explored in the framing effect mitigation domain. Research found that the impact warnings may have on participants, depends on the participants' involvement with the issue - framing effects disappeared when high involvement participants were exposed to weak or strong warning messages, yet less involvement participants were more susceptible to framing effects, and only strong warning messages were able to eliminate those effects [60]. Involvement, as moderator of framing effects, was also studied in political context, where it is referred to as *issue importance*; on both individual and group level, framing effects occur less among high importance issues, whereas large effects are observed for low-importance issues [61]. Third and last, several studies found frame opposition able to mitigate framing effects when presented next to the original frame [62]. However, in attempt to apply research to framing effects on the social media context (Twitter), frame opposition was not always able to influence participants' perceptions [63]. Different from

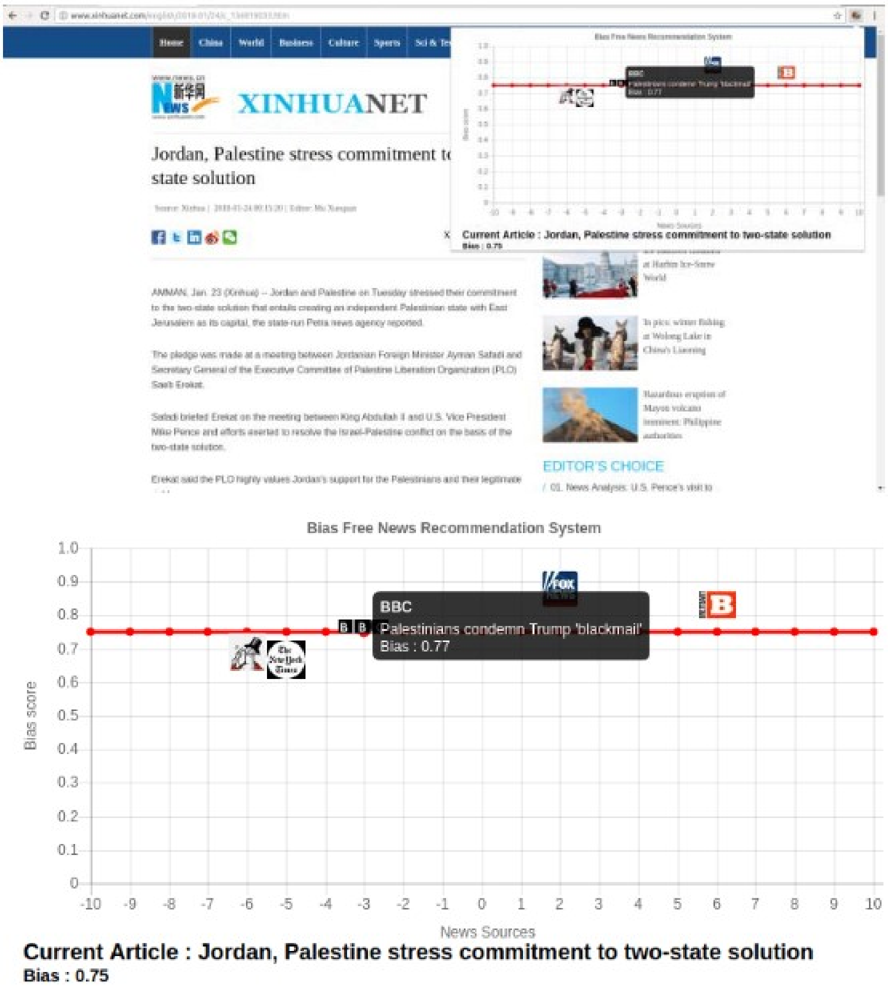


Figure 2.1: Snippet from the bias aware news recommendation system developed by Patankar et al. (2019). The researchers developed their own algorithm to measure political bias in articles from newspapers. This allowed them to create a scale to which they could map newspapers according to their bias, which is seen in the top right of the figure and highlighted in the bottom part. This way, the potential user would be able to compare on what side of the spectrum the current article is positioned with respect to the bias of other newspapers. [59]

Figure 2.2: Highlighted biased language example visualization adopted from Spinde (2020) [44]

Figure 2.3: The WordCloud visualization used in the study by Baumer (2018), which was found to mitigate political text framing effects

controlling what frames experiment participants are exposed to, the study mentioned outsourced the experiment setting conditions to a true discussion on Twitter itself, hence had no control over the content of (opposing) frames. Although this set-up couldn't have been more close to a real-life social media environment as it was in fact the Twitter environment itself, the study concluded that more in-depth analysis of the different dynamics in a social media environment need to be taken into account in future research before taking experiments to real-life social media environments.

Chapter 3

Current Study

The findings summarized in chapter 2 provide with several insights into the problem statement, of which the stimulation of analytic reasoning seems to be most promising in aiding one's truth discernment capabilities. However, evidence on how analytical reasoning can be triggered with information visualization on social media is lacking. This chapter highlights the approach of the current study, with which it aims to connect the research domains of psychology and social media with research in information visualization.

3.1 Research focus

Information visualization was found to be able to create more bias awareness and mitigate political framing effects under certain conditions. The current study will build upon prior research, yet with different visualization methods that are tailored to a Twitter-inspired social media setting. This approach bridges the gap between highly controlled, news article-focused experiment conditions - which are most common in literature - and a real-world social media setting, in which actual social media content is under inspection. Based on the summarized findings in chapter 2, the focus of the current study was set to trigger analytic reasoning with information visualization, as this was found to most evidently improve participant's truth discernment capabilities. Within this focus, two research goals were set: 1) creating more bias awareness and 2) mitigating political framing effects.

3.2 Setting

A scope was set on the general social media domain, by choosing 1) a platform, and 2) a political landscape.

Twitter Section 1.2 argued the choice for Twitter as a representative social media platform to facilitate the current study. Moreover, with its accessible Application Programming Interface (API), and plenty of available online resources, Twitter allowed for an inclusive data collection. This richness of data is of considerable value for the quality of the experiment.

The Netherlands Many studies to post-truth political discourse, fake news, and (social) media bias were done in a U.S. politics context. However, the bi-party system that is employed in the U.S. is significantly different from most democracies in Europe. Therefore, in attempt to get more insights in a different political setting, the focus of this study was set on the Dutch political landscape. The Dutch government is a multi-party system that works with a coalition and an opposition. A delegation of politicians coming from coalition parties constitute the ministry members of the cabinet, and both coalition and opposition party representatives take seat in the parliament. The parliament, in turn, discusses, approves and disapproves policy proposals from the cabinet.

3.3 Visualization methods

Given the two research goals set for the current study, three visualization methods were chosen to be featured in the experiments. This section argues those choices, as well as it explains the most important definitions.

3.3.1 Article-level visualizations

First, article-level visualizations are featured in order to test for political bias awareness. In the current study, an article is a political tweet, which will be referred to as the *main tweet*. Bearing in mind the characteristics of misinformation online, the most important contextual aspects of a main tweet are its context, its content, and its source (the *author*; a politician in this case). With respect to content, a main tweet's textual *subjectivity* and *polarity* (which is also referred to as *sentiment*) were chosen to be visually encoded. Although highlighting subjective or sentimental words in a main tweet's body was considered, this was excluded from the current study as the chance that Twitter would hypothetically adopt such an intrusive measure was assumed to be zero. Hence, the visual encodings for subjectivity and sentiment concerned standalone graphics. With respect to the main tweet's source, the political leaning of the author was chosen to be visualized. Although this was not proven to influence a participants' bias awareness when tested on news articles, it was still chosen to be included in the current study - howbeit, political tweets and news articles do differ greatly [57]. These article-level visualizations are *static*, and will be referred to as *context indicators*.

3.3.2 Frame opposition

Second, aimed at validating the influence of visualization on mitigation of political framing effects, the current study implements *frame opposition*. The opposing frames come from both politicians and news outlets, with the requirement that they share different arguments or information on the same issue as the main tweet. However, the types of frames that come from politicians, and the types of frames that come from news outlets, differ significantly. Politicians' tweets usually concern a political statement, and often contain counterarguments to oppose other politicians, for which their tweets can be considered *opposing tweets*. Tweets from news outlets are rather informative of nature, and can therefore be considered *supportive tweets* as they provide background knowledge for the arguments that politicians share. Including both allows participants to see a political frame in context; tweets by news outlets provide participants with information that may support their understanding of the opposing tweets by politicians.

3.3.3 Interactive visualization

Lastly, targeted at both research goal 1) and 2), an *interactive* visualization was developed with which participants could explore general, aggregate data about Twitter behavior of politicians, referred to as the *dashboard*. This dashboard was designed to take into account the four crucial requirements in critical information visualization; disclosure, plurality, contingency and empowerment[1] [64].

[1] In the article by Dörk (2013), disclosure concerns a creator's openness about what decisions were made in order to arrive at a visualization, plurality concerns the inclusion of multiple perspectives in the visualizations, contingency concerns the flexibility of visualizations, and empowerment refers to the users' potential to take a critical perspective on what they see.

3.4 User study hypotheses

The current study will validate the influence of the proposed information visualizations with a user study. The user study consists of independent experiments with participants (N=21). A participant is exposed to a total of nine main tweets under three conditions, yielding three tweets in each condition. The first condition is always the control condition, the second condition features the above mentioned article-level visualizations, i.e. the context indicators and the frame opposition, and the third condition features the same visualizations as in the second condition, yet with the addition of the dashboard. This sets the stage to test the following hypotheses:

H1 Visual context indicators increase political bias awareness.

H2 Frame opposition, applied to political tweets, mitigates possible political framing effects.

H3 The interactive visualization of aggregate Twitter data by means of a dashboard, allows participants to better understand the context of political tweets.

Chapter 4

Data

As the current study aims to validate the impact of information visualization in a setting that resembles a real-life social media platform as closely as possible, the data was collected from Twitter itself. With the current multi-party, indirect democracy, all politicians in the Netherlands represent the political party they are member of. Hence, the assumption was made that politicians' tweets on policy issues coming from their personal Twitter accounts represent their political party perspectives on those issues. First, eight Dutch political parties were selected to be included in the study[1]. Second, two politicians per political party were chosen. From this selection onward, the data collection, processing and enrichment was done in Python[2]. Figure 4.1 globally visualizes the steps that were taken in order to arrive at the final dataset. All depicted steps will be explained in this chapter, with visual support of two example tweets; one from a left-conservative and male politician (Geert Wilders, PVV), and one from a left-progressive and female politician (Corinne Ellemeet, GL).

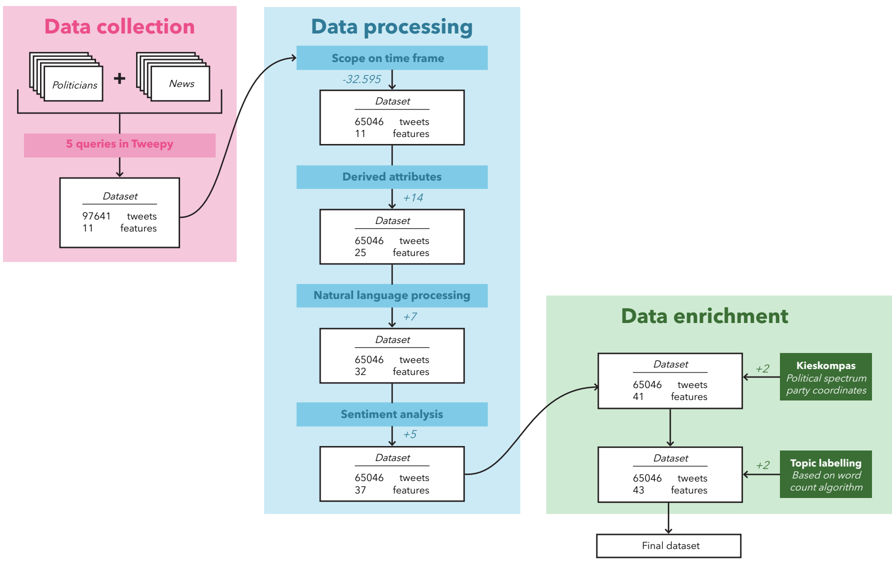


Figure 4.1: The process to arrive at the final dataset.

[1]Including all the 20 possible parties that take seat in the parliament was not expected to add much value to the process of finding the answer to the research question. Moreover, including over eight parties would have complicated the analysis of the user study.

[2]Python is a well-known programming language that is often used in data processing and analysis applications. The choice for the current study to use Python was based on the personal preference of the researcher, and most resources about working with Twitter data being available in Python.

4.1 Description of the dataset

This section discusses the selection of politicians and news outlets for the dataset.

4.1.1 Selection of politicians

The general elections of 2021 distributed the available 150 seats of the Dutch parliament among the politicians that take seat today. The eight largest political fractions according to this distribution were chosen to be included in the study[3]. Three out of the eight selected parties are part of the coalition, i.e. they formulated the coalition agreement and a subset of their representatives take part in the cabinet as ministers and state secretaries. The remaining five parties take part in the opposition, i.e. they take seat in the parliament and are, among others, involved in the approval and disapproval of policy proposals. To induce a gender-neutral starting point for the data collection process, one male and one female representative of each party were included. For every party, the politician that received the largest number of votes from the electorate was picked, as well as the second-most popular politician of a different gender[4]. Some politicians holding a position in the ministry employ two Twitter accounts; one in name of their personal identity, and one in name of their function. The latter type of Twitter account was excluded from this study, as the content shared is edited by a team of ministry employees rather than by the politicians themselves. This measure was appointed to comply with the scope of this study, which focuses on the influence of individual politicians on social media. All together, 16 Twitter accounts of politicians were included in the dataset, of which a brief description can be found in table 4.1.

Seats	Political party	Orientation	Position	Male politician	Female politician
34	VVD	Right-conservative	Coalition	Mark Rutte	Dilan Yesilgöz
24	D66	Left-progressive	Coalition	Rob Jetten	Sigrid Kaag
17	PVV	Left-conservative	Opposition	Geert Wilders	Fleur Agema
15	CDA	Right-conservative	Coalition	Wopke Hoekstra	Anne Kuik
9	SP	Left-progressive	Opposition	Mahir Alkaya	Lilian Marijnissen
9	PVDA	Left-progressive	Opposition	Henk Nijboer	Lilianne Ploumen
8	GL	Left-progressive	Opposition	Jesse Klaver	Corinne Ellemeet
8	FVD	Right-conservative	Opposition	Thierry Baudet	Simone Kerseboom

Table 4.1: The eight political parties included in the dataset

4.1.2 Selection of news outlets

With the declining trust in political institutions and the post-truth context of today, news outlets were chosen to be included in the frame opposition visualization as their tweets relate to political topics, yet are informative rather than argumentative of nature. Withal, some misinformation, disinformation or fake news may be present in the tweets that were collected from the selected politicians. The algorithmic detection of falsehoods was out of scope for this study, and as inspecting the correctness of information in individual tweets would take too much time, no fact-checking was done. Therefore, the data may contain some misinformation. However, as news outlets to-date enforce the journalistic code of reporting the truth, the possibly false content that would occur in the political tweets could potentially be nuanced by including tweets from news outlets in the experiment. Hence, in line with the national context of the current study, seven nationally

[3]By the time the current study was performed, one of the parties included in the selection (FVD) had split up shortly after the elections - i.e., some members of this party kept their seat in the parliament, but under the umbrella of a different, newly established political party. This implies that the listing of the eight largest fractions in the study is slightly different from the present listing of the eight largest parties in the Dutch parliament.

[4]One of the politicians (Lilianne Ploumen, PVDA) left the parliament shortly after the data collection had started. However, in order to keep the data consistent, she remained as one of the included politicians.

operating news outlets were selected - some traditional outlets for which printed press is their main business, as well as some hybrid outlets, and one online-only outlet. A description of the selected news outlet Twitter accounts is found in table 4.2.

News outlet	Classification
De Telegraaf	Newspaper & online
Het Algemeen Dagblad (AD)	Newspaper & online
De Volkskrant	Newspaper & online
NRC Handelsblad	Newspaper & online
Trouw	Newspaper & online
NOS	Television & online
NU.nl	Online

Table 4.2: News outlets included in the dataset

4.2 Data collection

This section lists the operational steps that were required to facilitate the data collection.

4.2.1 Twitter for Academia

In order to facilitate researchers with Twitter data via the Twitter API, i.e. a set of programmatic endpoints that can be used to understand or build the Twitter dialogue, Twitter offers researchers the possibility to request an academic developer account authorization, referred to as *academic authorization* [34]. The academic authorization has several advantages over the non-academic authorization, of which the most important ones for this study were 1) the possibility to query up to the 3200 most recent tweets of any public Twitter account based on a username, and 2) an eased data collection cap of ten million tweets per month instead of 500 thousand. The academic authorization request for the current study was approved in early March 2022, after which the data collection started.

4.2.2 Queries

The queries were supported by Tweepy, which is a Python library for accessing the Twitter API [65]. Within a Tweepy query, up to 21 attributes of any tweet can be asked for, ranging from the language of the tweet to the number of retweets it has gotten. The attributes that were included in the queries are mentioned in table 4.3. The query that was designed for the current study featured the request of the 3200 most recent tweets of 16 politicians and seven news outlets, i.e. a maximum of 73600 tweets per query. This query was done five times in total, on the following dates: 1) March 26, 2022, 2) April 15, 2022, 3) April 30, 4) June 22[5], and 5) July 2, 2022. The retrieved tweets for any of the politician and news accounts were stored locally as comma separated files. When establishing the complete dataset, the separate files were merged and duplicate tweets were removed. This step is visualized in the pink box in figure 4.1. In order to better understand the type of data that was included in the query, figure 4.2 lists these attributes for the two example tweets.

[5]The data collection for this study was decided to end after the third query. However, this decision changed after the presentation of the government's proposed policies to reduce the amount of nitrogen in the Netherlands. The policy proposal triggered several farmer protests across the country, as well as it caused heated debates in the parliament. Since politicians took this debate to Twitter too, two new queries were planned. Given the controversial topic, these tweets were expected to be valuable for the experiment. However, this decision caused a tweet-gap of approximately ten days for one of the most actively tweeting news outlets, as their 3200 most recent tweets would go back less far in time than the time between the third and fourth query.

Attribute	Description
status_id	Unique identifier of the tweet
created_at	Moment in time the tweet was published
user_id	Unique identifier of the tweet's author
screen_name	The author's username on Twitter
name	The author's full name
lang	Tweet language
full_text	The tweet's complete text content
in_reply_to_screen_name	The username of the account that the tweet's author is directly replying to
is_quote_status	Boolean indicating whether the author has quoted another tweet in the current tweet
retweet_count	The number of retweets the tweet has gotten
favorite_count	The number of likes the tweet has gotten

Table 4.3: Attributes that were included in the Tweepy query

Data collection
Queried attributes

status_id	1535270709689303041	1516307131624050688
created_at	2022-06-15 09:37:46+00:00	2022-04-19 06:46:06+00:00
user_id	41778159	1164004764
screen_name	geertwilderspvv	CorinneEllemeet
name	Geert Wilders	Corinne Ellemeet
lang	nl	nl
full_text	Rutte en Kaag moeten zich kapot schamen dat ze miljoenen Nederlanders die het financieel niet meer redden in de steek laten. De PVV zet Nederlanders wél op 1! #voorjaarsnota #koopkracht #PVV #Wilders https://t.co/dyNLf036ld	Dit is waar en het is een hele treurige en zorgelijke constatering. Voor veel Kamerleden is het geen onwil: Hoe hard we ook werken, het is niet allemaal (goed) te doen. Kamerleden staan niet in de rij om deel te nemen aan belangrijk commissiewerk https://t.co/kw6EKvwn39
in_reply_to_screen_name	-	-
is_quote_status	False	False
retweet_count	739	0
favorite_count	3450	12

Figure 4.2: The attributes contained in the Tweepy query (table 4.3), highlighted for the two example tweets. The black boxes on top are screenshots from the original tweets on Twitter. Multimedia (e.g., video clips) were removed from the screenshots.

4.3 Data processing

The collected data already contained several important attributes, as briefly explained in table 4.3 and examplified in figure 4.2. However, the data required a set of processing steps in order to find an appropriate bound on date to scope the time frame, derive relevant attributes to visualize, and induce natural language processing in order to eventually enable the sentiment analyses and topic labelling. These steps are shown in the blue box of figure 4.1. This section discusses the processing steps in more detail.

4.3.1 Scope on time frame

Due to Twitter's constraint of 3200 tweets per username per query, the time period that was covered per politician and per news outlet varied widely. Moreover, news outlets tweet significantly more than politicians. Three arguments led to the decision to scope the dataset with a lower and upper bound on tweet publishing date. First, the tweets of the least frequently tweeting politician led back to the year 2011. The topics that might have occurred in those tweets are outdated, and it would be impossible to find opposing frames for those tweets since other politicians were neither active on Twitter, nor could the query fetch tweets from that time as it would bypass the tweet-cap of 3200. Moreover, in order to ensure that - when it comes to tweeting about policy issues - politicians would potentially tweet about similar events and arguments, alignment in time is necessary. Second, including very old tweets in the experiment might confuse participants, as the political landscape looked differently before the elections of March 2021. Therefore, the lower bound was set for the first of December 2021, as this was the earliest month that was covered for the most frequently tweeting politician. June 30, 2022 was chosen as upper bound, such that only full months would be taken into account.

4.3.2 Derived attributes

Based on the *full text* attribute that was featured in the queried data (table 4.3 and figure 4.2), several other attributes were derived that could benefit the visualizations[6]. These attributes were retrieved with regular expressions in Python, and are listed in table 4.4. Figure 4.3 shows these attributes for the two example tweets. Note that some attributes seem the overlap between the queried attributes in table 4.3 and derived attributes in table 4.4; 1) *lang* with *nl/en*, and 2) *is quote status* with *is qt*. These derived attributes were added in order to make the dataset more consistent. First, the language attribute that was included in the Tweepy query was found to be unreliable - many tweets that contained non-normative language, or consisted of only hashtags or only emojis, were labelled arbitrary languages (e.g., Finnish). Manual inspection of the data had led to the conclusion to map all tweets with the *lang* label of "en" to English ("en"), and tweets with the *lang* label that was not "en" to Dutch ("nl"). This mapping is contained in the *nl/en* attribute in table 4.4. Second, manual inspection of the data resulted in the finding that the *is quote status* attribute showed "True" values for both *quote statuses* and *retweets*. However, these two types of statuses are different in nature; a quote status is the author's reflection on the tweet coming from a different user, whereas a retweet is the direct copy of a tweet coming from a different user, without the added reflection of the author. Therefore, two independent boolean attributes were added to distinguish the two types of tweet: *is rt* and *is qt*, as shown in table 4.4. All derived attributes are also highlighted for the example tweets in figure 4.3

4.3.3 Natural language processing

In order to prepare the data for natural language processing, the full text of a tweet was cleaned by stripping links, emoticons and hashtags[7]. The actual natural language processing was done with

[6]Note that not all of the derived attributes were ultimately featured in the visualizations.

[7]Hashtags are used on Twitter to emphasize certain words or claims, as well as they operate as topic identifiers. Twitter keeps track of all hashtags posted by their users, and are therefore able to easily identify which hashtags

Derived attribute	Description
date	Simplified publishing table featuring only the date
rt_of	The person that was being retweeted by the tweet's author
retweet	The retweet content
hashtags	All hashtags that occurred in the tweet
hashtags_not_polit	Hashtags that did not feature political party names or political self-promotion
hashtags_polit	Hashtags containing only political party names or political self-promotion
hashtags_count	Total number of hashtags occurring in the tweet
uppercase_words	Full uppercase words occurring in the tweet
exclamation_marks_count	Total number of exclamation marks occurring in the tweet
question_marks_count	Total number of question marks occurring in the tweet
emoji	The emoji(s) that the tweet contained
nl/en	Boolean value indicating whether the tweet was Dutch or English
is_rt	Boolean value indicating whether the tweet was retweeted or originally written by the politician or news outlet
is_qt	Boolean value indicating whether the tweet has quoted another tweet

Table 4.4: Derived attributes

help of the natural language toolkit (NLTK) [66]. In order to prepare the data for exploratory topic mining, the cleaned data was tokenized, lemmatized and stemmed. In topic mining, the recurrence of word combinations determines what the algorithm picks up as clusters. Therefore, the algorithm is helped if derivations of words are brought back to their root, i.e. their stem[8]. Lemmatization was induced in advance of stemming to group together forms of words such that they can be considered a single item, i.e. in dictionary form (the word's lemma)[9] [67]. These steps are highlighted for the two example tweets in figure 4.4.

4.3.4 Sentiment analysis: pattern-nl

Sentiment analysis is typically done with either machine learning-based methods or rule-based methods. Machine learning methods are trained on a labelled input, i.e. textual data items that each have a label that indicates the sentiment. The idea is that the trained algorithm can later tell the sentiment for new data items, based on the features it has learned from training. In turn, rule-based methods calculate the sentiment of data items based on scores for individual words or word combinations that it retrieves from a lexicon. This section argues the choice for a rule-based method for the current study.

Machine learning based methods

As manually labelling tweets on their sentiment would require too much time and increase the inflicted bias from the researcher, an off-the shelf method was searched for. RobBERT is a Dutch BERT-based language model that is pre-trained on a dataset of book reviews. Its evaluation on sentiment analysis tasks is promising, however, the model operates as a black-box [68]. This implies that retrieving the sentiment assessments on word-level would require the application of a different algorithm, such as SHAP, yielding results that are rather complex to understand for

are popular at different times and at different places around the world. The most popular hashtags are considered *trending topics*.

[8]The stemmed version of the words "likes", "likely", "liking", "liked" is "like".

[9]The word "better" has "good" as its lemma.

Data processing *Derived attributes*		
date	15/06/2022	19/04/2022
rt_of	-	-
retweet	-	-
hashtags	['voorjaarsnota', 'koopkracht', 'pvv', 'wilders']	-
hashtags_not_polit	['voorjaarsnota', 'koopkracht', 'wilders']	-
hashtags_polit	['pvv']	-
hashtags_count	4	-
uppercase_words	['PVV', 'PVV']	-
exclamation_marks_count	1	-
question_marks_count	0	-
emoji	-	-
nl/en	nl	nl
is_rt	Original	Original
is_qt	Original	Original

Figure 4.3: The derived attributes, highlighted for the two example tweets.

laypersons, i.e. experiment participants in this case [69]. Bearing in mind that algorithms are biased in nature, the difficulty to retrieve transparency about the sentiment assessments was the main argument to exclude this model from performing sentiment analysis on the current dataset, as the disclosure of information is a key component in reliable information visualization [64].

Rule-based methods

Pattern-nl is the Dutch submodule of Pattern, an open-source Python package for natural language processing. It contains a rule-based sentiment analyzer, built on a lexicon of about 4000 Dutch lemmata. This method assigns a subjectivity and a polarity score to each word in a sentence, and takes the average of those values to arrive at a total sentiment score of the input text [70]. The package allows to directly retrieve the sentimental words on which the total sentiment score was based, including the subjectivity and polarity scores of each of those words. Here, subjectivity is a continuous score mapped on a domain of [0,1] where 0 represents "objective" and 1 represents "subjective", and polarity is a continuous score mapped on a domain of [-1,1] where -1 represents "negative" and 1 represents "positive". Although achieving an accuracy of up to 80% on book reviews, the total score calculation lacks a proper anticipation on context-specific cues. This causes the algorithm to perform worse in settings where the lexicon has little coverage [70]. However, as the ease and transparency to obtain the sentiment assessments was prioritized over accuracy for the current study, pattern-nl was the preferred option for performing sentiment analysis. After execution, the subjectivity and polarity score for each tweet were added to the data. Additionally, the words that were picked up by the algorithm - i.e., the tweet's words that were included in pattern-nl's lexicon - were added in a separate column, together with the subjectivity and polarity scores for each word. These are contained in the *sentiment assessments* attribute, as also visible in figure 4.5.

Bucket mapping of sentiment values

Summary statistics were added to the sentiment scores in order to ease participants' interpretation. For the sentiment analysis done in pattern-nl, the two numerically continuous scores within a domain of $[0, 1]$ for subjectivity (table 4.5) and $[-1, 1]$ for polarity (table 4.6) were mapped into *buckets* that correspond with specific domain ranges. It is noteworthy that a "neutral" bucket is missing for the subjectivity domain - this is the case since the bucket for "objective" is in fact the value that most closely corresponds with one's perception of neutral. The sentiment scores and bucket mappings are highlighted for the two example tweets in figure 4.5.

Domain	Subjectivity
[0,0.25)	Objective
[0.25,0.5)	Considerably objective
[0.5,0.75)	Considerably subjective
(0.75,1]	Subjective

Table 4.5: Subjectivity domain buckets

Domain	Polarity
[-1,-0.5)	Negative
[-0.5,0)	Fairly negative
[0,0]	Neutral
(0,0.5]	Fairly positive
(0.5,1]	Positive

Table 4.6: Polarity domain buckets

Data processing

Natural language processing

full_text	Rutte en Kaag moeten zich kapot schamen dat ze miljoenen Nederlanders die het financieel niet meer redden in de steek laten. De PVV zet Nederlanders wél op 1! #voorjaarsnota #koopkracht #PVV #Wilders https://t.co/dyNLf036ld	Dit is waar en het is een hele treurige en zorgelijke constatering. Voor veel Kamerleden is het geen onwil: Hoe hard we ook werken, het is niet allemaal (goed) te doen. Kamerleden staan niet in de rij om deel te nemen aan belangrijk commissiewerk https://t.co/kw6EKvwn39
lowercase	rutte en kaag moeten zich kapot schamen dat ze miljoenen nederlanders die het financieel niet meer redden in de steek laten. de pvv zet nederlanders wél op 1! #voorjaarsnota #koopkracht #pvv #wilders https://t.co/dynlf036ld	dit is waar en het is een hele treurige en zorgelijke constatering. voor veel kamerleden is het geen onwil: hoe hard we ook werken, het is niet allemaal (goed) te doen. kamerleden staan niet in de rij om deel te nemen aan belangrijk commissiewerk https://t.co/kw6ekvwn39
clean_text	rutte en kaag moeten zich kapot schamen dat ze miljoenen nederlanders die het financieel niet meer redden in de steek laten de pvv zet nederlanders wél op 1 voorjaarsnota koopkracht pvv wilders	dit is waar en het is een hele treurige en zorgelijke constatering voor veel kamerleden is het geen onwil hoe hard we ook werken het is niet allemaal goed te doen kamerleden staan niet in de rij om deel te nemen aan belangrijk commissiewerk
clean_text_no_stopwords	rutte kaag moeten kapot schamen miljoenen nederlanders financieel niet redden steek laten pvv zet nederlanders wél 1 voorjaarsnota koopkracht pvv wilders	waar hele treurige zorgelijke constatering kamerleden geen onwil hard werken niet allemaal goed kamerleden staan niet rij deel nemen belangrijk commissiewerk
tokenized	[‘rutte’, ‘en’, ‘kaag’, ‘moeten’, ‘zich’, ‘kapot’, ‘schamen’, ‘dat’, ‘ze’, ‘miljoenen’, ‘nederlanders’, ‘die’, ‘het’, ‘financieel’, ‘niet’, ‘meer’, ‘redden’, ‘in’, ‘de’, ‘steek’, ‘laten’, ‘de’, ‘pvv’, ‘zet’, ‘nederlanders’, ‘wél’, ‘op’, ‘1’, ‘voorjaarsnota’, ‘koopkracht’, ‘pvv’, ‘wilders’]	[‘dit’, ‘is’, ‘waar’, ‘en’, ‘het’, ‘is’, ‘een’, ‘hele’, ‘treurige’, ‘en’, ‘zorgelijke’, ‘constatering’, ‘voor’, ‘veel’, ‘kamerleden’, ‘is’, ‘het’, ‘geen’, ‘onwil’, ‘hoe’, ‘hard’, ‘we’, ‘ook’, ‘werken’, ‘het’, ‘is’, ‘niet’, ‘allemaal’, ‘goed’, ‘te’, ‘doen’, ‘kamerleden’, ‘staan’, ‘niet’, ‘in’, ‘de’, ‘rij’, ‘om’, ‘deel’, ‘te’, ‘nemen’, ‘aan’, ‘belangrijk’, ‘commissiewerk’]
lemma	rutte kaag moeten kapot schamen miljoen nederlanders financieel niet redden steken laten pvv zetten nederlanders wél 1 voorjaarsnota koopkracht pvv wilders	waar heel treurig zorgelijk constatering kamerlid geen onwil hard werken niet allemaal goed kamerlid staan niet rij deel nemen belangrijk commissiewerk
stemmed	rutte kaag moeten kapot schamen miljoen nederlanders financieel niet redden steken laten pvv zetten nederlanders wel 1 voorjaarsnota koopkracht pvv wilder	waar heel treurig zorgelijk constatering kamerlid geen onwil hard werken niet allemaal goed kamerlid staan niet rij deel nemen belangrijk commissiewerk

Figure 4.4: Natural language processing steps, highlighted for the two example tweets.

Data processing *Sentiment analysis*		
polarity	-0.175	0.14
subjectivity	0.35	0.79
sentiment_assessments	[(['kapot'], -0.35, 0.7, None), (['financieel'], 0.0, 0.0, None)]	[(['waar'], 0.2, 0.5, None), (['hele', 'treurige', 'zorgelijke'], -0.4, 0.9, None), (['hard'], 0.0, 0.8, None), (['goed'], 0.55, 0.9, None), (['belangrijk'], 0.35, 0.8500000000000001, None)]
Bucket classification		
bucket_polarity	redelijk negatief	redelijk positief
bucket_subjectivity	overwegend objectief	subjectief

Figure 4.5: The sentiment analysis values and bucket mappings, highlighted for the two example tweets.

4.4 Data enrichment

This section explains the steps taken to add externally sourced attributes to the dataset.

4.4.1 Political spectrum classification

One of the key features that the information visualization could show in attempt to enhance a users' understanding of the political context of a tweet, is the political leaning of the author. Several ways exist to classify political parties, of which the political spectrum is the most common one to date. The political spectrum consists of two axes; the x-axis, distinguishing left-wing from right-wing, and the y-axis, distinguishing conservative from progressive. Kieskompas is an independent, scientific institute in the Netherlands that develops political spectra for every election based on the analysis of political parties' election programs [71]. Their spectrum to map the Dutch political landscape ranges form [-2,2] on both the x- and the y-axis, yielding a unique coordinate for all political parties. The coordinates of the eight political parties included in the current study were used based on KiesKompas' 2021 general election spectrum. These values are showcased for the two example tweets in figure 4.6

Figure 4.6: The Kieskompas coordinated for the two example tweets of Wilders and Ellemeet, based on their political parties, PVV and GL, respectively.

4.4.2 Topic labelling

To realize frame opposition in the experiment, tweets on the same policy issue yet yielding different arguments from different politicians needed to be found. In order to ease this search, exploratory topic mining was applied in advance to identify topics that were shared among politicians in the dataset.

Exploratory topic mining with Latent Dirichlet Allocation

Latent Dirichlet Allocation (LDA) is a probablistic Bayesian model that considers a text pattern as a distribution of words, called topics. Its possible applications are widely studied and used, yet mainly on documents that are typically richer in text than tweets, which are messy in language and contain less than 280 characters. Hence, several methods were inspected to improve the algorithm's applicability for Twitter data, such as tweet pooling by author, which was used for the current study[10] [72]. As the topic modeling part of this study was exclusively for exploratory reasons, LDA was used on the dataset to uncover broad topics, which would be further inspected at later

[10] In the study by Mehrota (2013), tweet pooling by hashtag revealed better results than tweet pooling by author. Tweet pooling by hashtag was however impossible for the current study, as the way included politicians use hashtags is highly inconsistent.

stage, namely in the manual tweet selection process. Since the main focus in the experiment is the opposition of political frames, with news media frames of supportive nature, the tweets of news outlets were excluded from this exploratory topic modelling step in order to reduce complexity. Moreover, was assumed that if different politicians tweet about a topic, news outlets would also tweet about this topic.

Wordlist-based algorithm

The application of LDA modelling to the dataset with news media tweets excluded, revealed three main topics: 1) the war in Ukrain, 2) the coronavirus, and 3) general policy criticism towards the Dutch cabinet. However, this topic labelling was done after the first three queries, hence excluded the data contained in query four and five. Therefore, a fourth topic was added by hand; i.e. the political crisis concerning nitrogen policy in the Netherlands. Based on the algorithm's output in terms of most frequently occurring words for each of the first three topics, complemented with related topic words that were found on the internet[11], wordlists were constructed. Wordlists could be of the type *main topic* or *subtopic*, where the first type of wordlist is generally descriptive of a topic, whereas the latter is more specific for an aspect of that topic. By means of example, "Violence" was found a subtopic of the main topic "War in Ukraine". Next, an algorithm was written in Python to check whether a tweet would classify as a main topic or subtopic, based on the wordlist-counts for words occurring in the tweet. This process is explained in Appendix A. The correctness of the labelling algorithm was checked by hand. Soon as random checks showed satisfactory results, the algorithm was run on the whole dataset (including tweets from news outlets), after which the labels were added. Figure 4.7 shows the labels that were granted to the two example tweets.

Figure 4.7: The externally sourced data additions relating to the political background of a tweet.

4.5 Main, opposing and supportive tweet selection

The quality of the experiment is dependent on the content of the tweets that participants would be exposed to. Next to ensuring that each political party would be featured as a main tweet at least once, an additional set of five strict requirements was enforced in the selection of main tweets, opposing tweets, and supportive tweets. This section describes those requirements in further detail, as well as it contains a description of the manual selection process.

[11]The wordlist of the fourth topic only features topic words that were found online, as the LDA model was only applied to data generated with the first three queries.

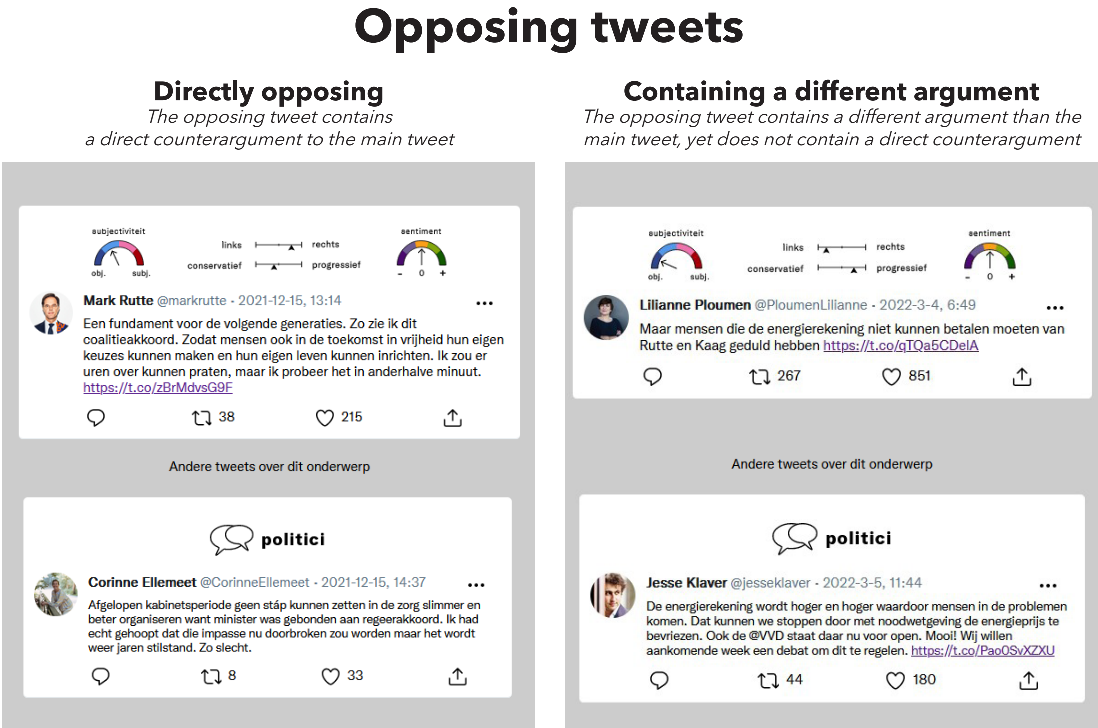


Figure 4.8: The manual selection matched two types of opposing tweets to any main tweet. On the left, frame opposition is realized with directly opposing tweets that contain counterarguments on the same issue. On the right, the frame opposition is realized with tweets that contain two different arguments on the same issue, yet are no direct counterarguments. All main tweets were accompanied with a combination of both types of opposing tweets.

4.5.1 Requirements

First, in order to resemble Dutch politics as close as possible, the language of the tweet needed to be Dutch. Second, retweets and direct replies were excluded from the selection in order to ensure that all political tweets were written by politicians themselves[12] and were standalone statements[13]. Third, the tweets must cover a controversial issue that is topical in the Netherlands, either at times of the experiment, or had enjoyed allover news coverage around the time of publication. Fourth, when a main tweet was chosen, the opposing tweets by politicians and supportive tweets by news media, to be featured in experiment conditions 2) and 3), must have been published within a total period of ten days. Fifth and last, the opposing tweets by politicians must be authored by at least two different politicians and communicate different arguments. The opposing tweets were therefore not always directly opposing, yet were made sure to share different arguments than the arguments presented in the main tweet. Figure 4.8 highlights this distinction. The supportive tweets by news media often overlapped, i.e. different news media reported about the same events in similar ways. When making a selection, the main goal was to include tweets from different news media such that the overlap between tweets was minimal.

[12]On Twitter, any user can “retweet” a tweet that was originally written by someone else. This means that if user A retweets user B, all users that follow user A will then see the original tweet from user B, however with the note that user A retweeted that tweet. A “direct reply” concerns the situation where user A replies to a tweet from user B as means of a conversation starter. The conversation that follows is known as a “Twitter thread”.

[13]In this case, a standalone statement meant a statement that did not need additional information in order to be understood.

4.5.2 Manual selection

At the start of the selection process, the dataset was manually explored. Later, more specific tweets were picked and matching tweets from other politicians and news media were looked for within the time constraint. The final selection of the nine tweets was relatively balanced, with five left-wing and four right-wing tweet authors, of which three authors take part in the coalition and the remaining six take part in the opposition. The selection of main tweets is highlighted in Appendix B. The possible bias of the researcher was tried to be minimized as much as possible, yet, an inherent bias may be present.

Chapter 5

Visualization

In the process of deciding what visualizations to include, the current study followed the framework described by Munzner (2014) [73]. This framework builds upon four nested levels of visualization design, which are a *domain situation* mapping, a *data* and *task abstraction*, a choice of *visual encoding* and *interaction idiom*, and the *technical implementation* respectively. In line with previous studies to framing effects, experiment participants were questioned about their interpretation of one central frame at a time. As the current study builds upon real-world Twitter data from politicians, this central frame was a politician's tweet and is therefore referred to as the *main tweet*. The impact that the visualizations may or may not have, will hence always be relative to this main tweet. In case of frame opposition, both tweets from other politicians as well as news outlets are shown. These are referred to as *opposing tweets* and *supportive tweets* respectively.

5.1 Domain situation

The domain situation mapping was done in order to better understand what motives drive potential users, referred to as *users*, and in which setting they find themselves.

User As the information visualization is designed for experiment purposes only, the user is a participant of the experiment. This participant is assumed to be a representative social media user. No evidence was found about typical demographics of the population of people that follow and engage with Dutch politicians on Twitter, hence accounting for demographics was kept out of scope. Therefore, the user of the visualization was defined to be any social media user, without any further assumptions about political dedication, party preference, age, or education level. Concretely, the set of visualizations should meet different types of social media behavior (e.g. in terms of scrolling, browsing, engaging), which can range from superficial to heavily invested.

Setting The setting of the domain is simulated, as participants would not truly find themselves on Twitter, but rather in a Twitter-inspired yet controlled environment. However, participants were asked to adopt reading behaviour as if they would be on Twitter itself. As the problem statement is independent of a specific social network, Twitter served as a representative example of a social media platform.

5.2 Data and task abstraction

Although visualizations are usually tailored to a domain, an abstraction of the data and possible user tasks supports their design process. Here, a data abstraction mainly provides with better insights on what types of data are at disposal, and a task abstraction helps to approach the construction of goal-oriented visualizations in a more fundamental and structured manner. This section lists both the data and task abstraction done for the current study.

5.2.1 Data abstraction

The data that was queried from Twitter contained several noteworthy tweet attributes, such as the publishing time and *engagement*, i.e. the number of *likes* and *retweets* a tweet had gotten. However, some of these attributes are already visible to users when viewing the tweet on Twitter, and would therefore not add much visible value. Bearing in mind the characteristics of fake news and misinformation, other data attributes were derived or calculated to enrich the dataset further. Precisely these *unraveled* tweet attributes that are not directly visible to the user, may prove the added value of information visualization. A full overview of the data attributes and their types is found in table 5.1. The attributes that resulted from natural language processing, e.g. the lemmatized *full text* attribute, were omitted from the data abstraction, as they were removed from the dataset after the sentiment analysis and topic labelling were realized.

5.2.2 Task abstraction

The methodology described in Munzner (2014) derives the task abstraction from action-target pairs [73]. *Actions* describe active steps a user can take, and is therefore formulated in verb-form. *Targets* are stated in return to those actions, as they describe data features a user may be interested in [73]. A classical example of an action-target pair is to *lookup* a *word* in the dictionairy, where *lookup* is the action and *word* is the target. The main focus of the study was formulated in chapter 3, i.e. to trigger one's analytic reasoning, within which two research goals were defined: 1) increase political bias awareness, and 2) mitigate political framing effects. The task abstraction for the current study was based on these research goals. An overview of action-target pairs as a result of the task abstraction is presented in table 5.2, where the last column indicates within which visualization a task was realized. This could be one of the three visualizations included in the user study: 1) *static* visualization, referred to as *context indicators*, 2) *frame opposition* visualization, or 3) *interactive* visualization, referred to as *dashboard*.

Context indicators

This static visualization is aimed at informing the users about specific features of the main tweet, which were otherwise not explicitly visible. Phrased differently, the context indicators serve as a subtle attention message upon which the user may decide what to think of it - hence, the action-target pair would be to *enjoy features* according to Munzner's framework [73]. These context indicators are aimed at increasing a participant's bias awareness, which is one of the research goals of the current study. The context indicators were introduced in the second experiment condition (section 3.4), and will be visually explained in section 5.3.1.

Frame opposition

Research showed that frame opposition is able to mitigate framing effects when shown next to the original frame, i.e. the *main tweet* in the current study. In the experiment set-up of the current study, frame opposition was achieved by exposing participants to *opposing tweets* from other politicians and *supportive tweets* from news outlets. As real Twitter data was used, no control was imposed over the exact content of these frames - hence, frames were not always directly opposing, since politicians do not always reply directly to each other, nor do they always talk about the same issues within a same time frame. More specifically, inspection of the data proved that some issues were thoroughly discussed by opposition politicians, yet were not granted any coverage by coalition politicians. In these cases, tweets by news outlets oftentimes communicated more background information about certain policy issues from a more coalition-like perspective. In any case, the task for the user is to *compare frames* on their content and *identify (dis)similarities* between arguments. Like the context indicators, the frame opposition visualization was introduced in the second experiment condition (section 3.4), and will be visually explained in section 5.3.1.

Data	Attribute	Data type	Note
Queried	*status_id*	Ordinal	
	created_at	Date/time	
	user_id	Ordinal	
	screen_name	Categorical	
	name	Categorical	
	lang	Categorical	
	full_text	String	
	$in\_reply\_to\_screen_name$	Categorical	
	is_quote_status	Boolean	
	retweet_count	Quantitative	
	favorite_count	Quantitative	
Derived	*date*	Date	
	rt_of	Categorical	
	retweet	String	
	hashtags	List	List of strings (hastags occurring in the tweet)
	hashtags_not_polit	List	List of strings (non-political hastags occurring in the tweet)
	hashtags_polit	List	List of strings (political hastags occurring in the tweet)
	hashtags_count	Quantitative	
	uppercase_words	List	List of uppercase words that occurred in the tweet
	exclamation_marks_count	Quantitative	
	question_marks_count	Quantitative	
	emoji	Categorical	
	nl/en	Categorical	Valued 'en' or 'nl'
	is_rt	Categorical	Valued 'Original' or 'Retweet'
	is_qt	Categorical	Valued 'Original' or 'Quote status'
Sentiment analysis	*subjectivity*	Quantitative	
	polarity	Quantitative	
	bucket_polarity	Categorical	
	bucket_subjectivity	Categorical	
	sentiment_assessments	List	List of tuples, where a tuple is of the form (word, polarity value, subjectivity value)
Political spectrum	*political_party*	Categorical	
	leftwing_rightwing	Quantitative	
	progressive_conservative	Quantitative	
Topic labelling	*label*	Categorical	
	sublabel	Categorical	

Table 5.1: The data abstraction for all processing steps

Level	Action	Target	Task description (participant)	Realization
Analyze	Enjoy	Features	Be exposed to the context indicators, in order to become aware of the tweet's bias	Static visualization
	Discover	Trends, correlations	Use the dashboard to explore new or confirm existing hypotheses about politicians' Twitter behavior	Interactive visualization
Search	Browse	Trends, correlations	Find out whether the author of a tweet is an outlier, or shows behavioral correlations between any of the following: tweet frequency, generated engagement, or calculated bias	
	Locate	Trends, correlations	Find out which politicians show trends or correlations between any combination of the following: tweet frequency, generated engagement, or calculated bias	
Query	Compare	Trends, correlations	Compare political parties or individual politicians on tweet frequency, generated engagement, or calculated bias	
		Features	Compare tweets in-depth on their bias, by comparing the biased language of opposing frames	
			Compare arguments from different politicians on the same issue	Frame opposition
	Identify	(Dis)similarities	Identify whether there are (dis)similarities between arguments from different politicians	

Table 5.2: The task abstraction made for the current study, based on the methodology described in Munzner (2014) [73].

Dashboard

Whereas the frame opposition and context indicators were focused on the main tweet, this interactive visualization is aimed at showing what goes beyond this tweet. The dataset allowed to place the main tweets in context, e.g. by comparing the main tweet to the average Twitter behavior of its author, or comparing it to the average Twitter behavior of other authors (i.e., other politicians). Hence, a first set of tasks is to *compare trends*, *compare correlations*, and *compare features*. Second, a user might be interested in searching for outliers and correlations, such as those between author tweet frequency, author engagement and author bias. This might differ between situations where the user is already knowing of the main tweet's author, or contrarily, has never seen or heard about this politician before. In other words, the user either *locates* whether the author of the main tweet is an outlier (i.e., the location is unknown, but the target is known) or *browses* who may be the outlier if not knowing yet who that outlier may be (i.e. the location is known, but the target is unknown). The same applies to the search for correlations. Lastly, a user might want to *discover trends* in tweet frequency, engagement, and bias by the author, as it enables to understand whether the main tweet follows or differs from that trend. The dashboard was introduced in the third experiment condition (section 3.4), and will be visually explained in section 5.3.1.

5.3 Visual encoding and interaction idiom

The visual encoding and interaction idiom concern the more visual part of the visualization design process. Building upon the task abstraction, this section argues what visual encodings were chosen, and with which interaction idioms some of them were complemented.

5.3.1 Visual encoding

As the main theme of the study - politics - is a highly delicate subject, the visual encodings were kept as neutral as possible[1]. This starts with the design of the main interface that the experiment participants would find themselves in. The main interface, as shown in figure 5.1 and figure 5.2, contained a blue navigation bar where the participants could track their progress in the experiment once started, and a light grey background covering the whole page underneath. The placeholder for the main tweet, the opposing tweets and the supportive tweets, as well as the placeholder for the questions, all had white backgrounds. Serving as reading area's with the featuring of black text, this was assumed to favor general readability. The layout of all tweets exposed on the interface was copied from Twitter's website, such that it would resemble a real-life tweet as closely as possible[2].

Context indicators

Since Twitter includes the depiction of different forms of engagement in a tweet's footer, it seemed a logical choice to position the context indicators in a tweet's header. The contextual tweet features that were chosen to be shown in this header were article level-based and consisted of 1) author-specific features, being political spectrum orientation, and 2) tweet-specific features, being textual subjectivity and polarity. The context indicators are shown in figure 5.3.

Frame opposition

In order to ease the distinction between the opposing tweets and the supportive tweets, two separate placeholders were designed with clear labels - one for *politicians* and one for *news outlets*. These are depicted in figure 5.4. As comparing more than two tweets of each type at the same time was considered to be an overload of information, the containers only showed one tweet at a time. The participant could scroll horizontally within every placeholder. The placeholder for politicians was positioned directly below the main tweet, to implicitly depict them as of similar author type (i.e., political). The placeholder for news media was positioned directly below the placeholder for politicians. Moreover, positioning other politicians more central in the screen was expected to unconsciously draw more attention from the participants, compliant with the findings on positioning effects in news recommendation research [55].

Political spectrum orientation For this author-specific feature, two quantitative values needed to be visualized - one on a scale from left-wing to right-wing, and one on a scale from progressive to conservative, both according to the political spectrum. The position of the author's coordinate relative to the conceptual center in this spectrum was assumed to be the most important in a

[1] An implicit bias of the researcher as designer of these visualizations will however be present, yet was attempted to be kept minimal.

[2] There is, however, an exception in layout for tweets that include an external link. Twitter itself shows a preview of articles below the tweet at display, in which the headline and intro can be read directly. This implies that a user might not have to click on the external link in order to read an article's headline. This preview was omitted in the current study.

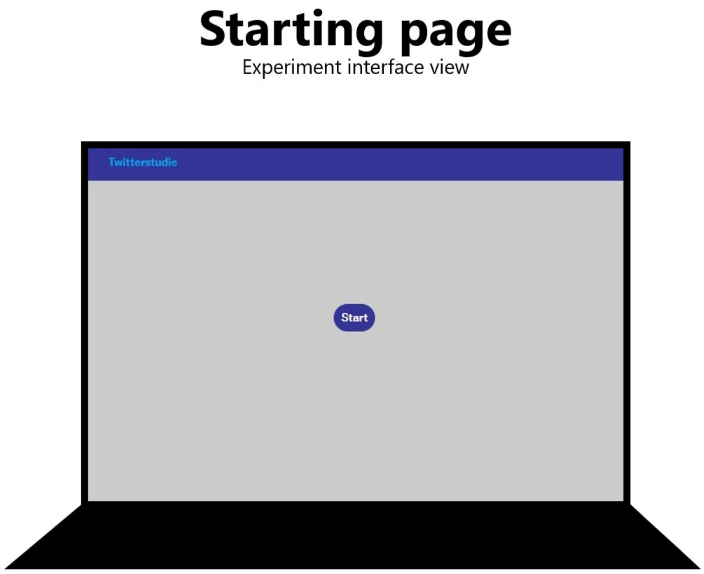


Figure 5.1: The experiment interface view for the starting page.

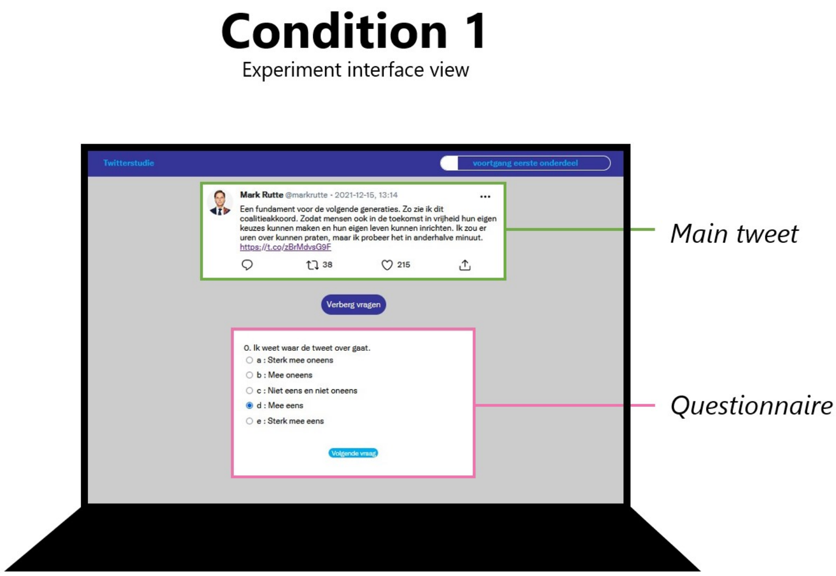


Figure 5.2: The experiment interface view for a main tweet in condition 1.

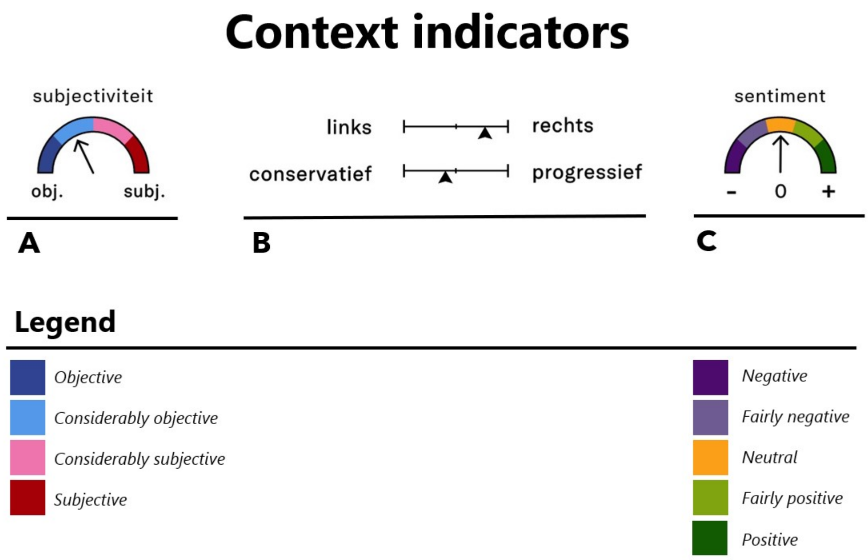


Figure 5.3: The context indicators, including the corresponding legend

participant's interpretation. Hence, the two coordinate scores were mapped separately onto two simple axes, as shown in figure 5.3 (B).

Subjectivity and polarity The tweet-specific features of subjectivity and polarity were converted to categorical values in order to enhance a participant's intuitive understanding of these values (see table 4.6 and table 4.5). The two features at stake were mapped to divergent color scales. Since the literal Dutch translation of the word polarity is relatively unknown, the score was renamed to *sentiment*. In literature, negative to positive sentiment is often color-coded on a red to green color scale [74]. However, the current study deferred from this design, in order to not trigger an association with *bad* (red) and *good* (green). The subjectivity visualization is shown in figure 5.3 (A) and the sentiment visualization is shown in figure 5.3 (C), with the legend added at the bottom of the figure.

Dashboard

The interactive visualizations were contained in a dashboard. Its embedding in the experiment interface is depicted in figure 5.5. The dashboard contained two sheets, both with a different focus. The first sheet was aimed at giving the user a more in-depth perspective on the main tweet in comparison to the opposing and supportive tweets, facilitating the comparison tasks in particular, and is depicted in figure 5.6. The second sheet was aimed at giving the user an overall insight into Twitter behavior among politicians, i.e. on aggregate level rather than individual tweet level, and is depicted in figure 5.7. News media were excluded from the second sheet.

5.3.2 Interaction idiom

The dashboard was made interactive, since interactivity enables the user to truly make information their own. As the visualizations were linked, the user was able to further inspect data by selecting filters that are shared between visuals. These types of interactions differed between the dashboard's first sheet, i.e. tailored to the comparison of the main tweet respective to the opposing and supportive tweets, and the second sheet, aimed at providing a general overview of politician's Twitter behavior.

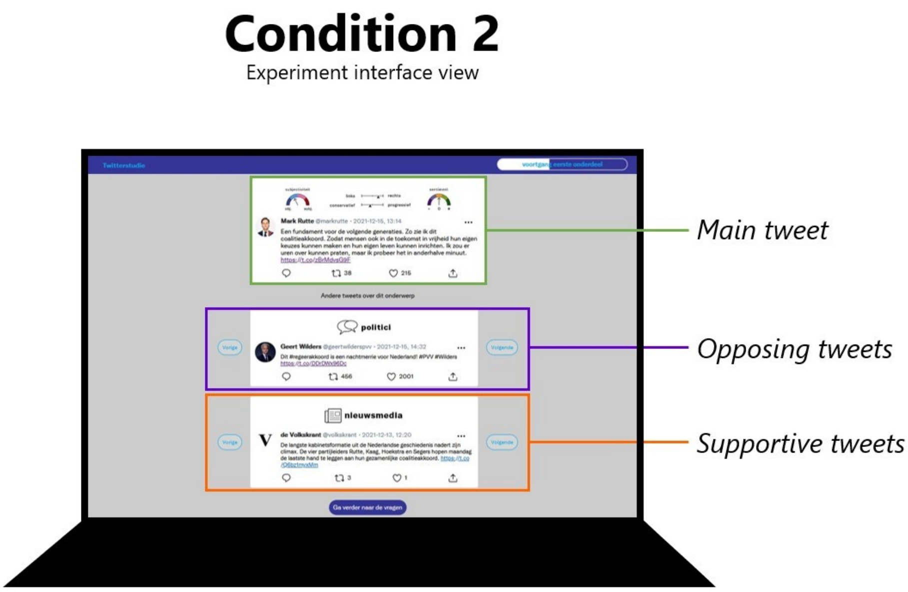


Figure 5.4: The experiment interface in condition 2

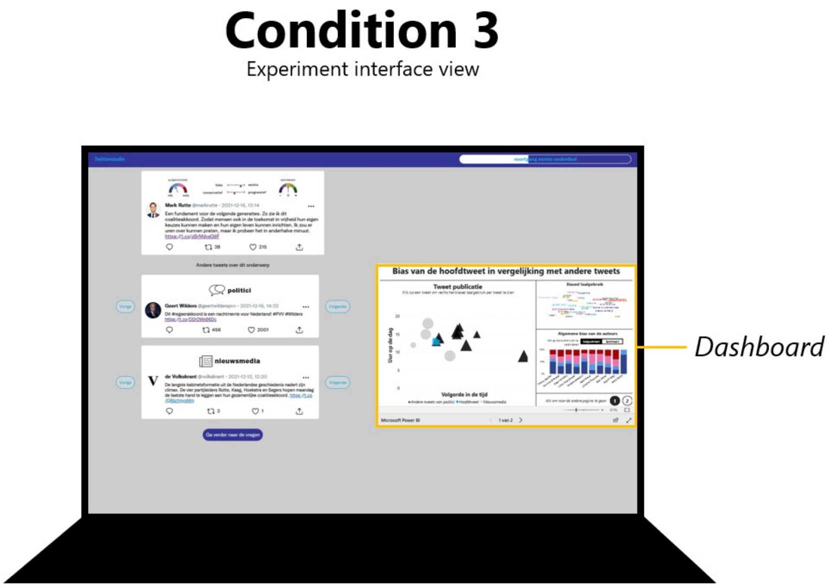


Figure 5.5: The experiment interface in condition 3

First sheet

In the first sheet, the user was able to see the exposed tweets - i.e. the main tweet, opposing tweets and supportive tweets - on a timeline, where hover effects were implemented to fetch a tweet's content and bias scores. By clicking on the main tweet, the user was able to see what words the algorithm had picked up as biased. Figure 5.8 depicts the dashboard view at inspection of, in this case, the main tweet. By deselecting a tweet, the user was brought back to the original WordCloud that featured all biased language contained in the exposed tweets all together (figure 5.6 (B)). The percentage bar charts on the right side of the dashboard (figure 5.6 (C)) allowed to

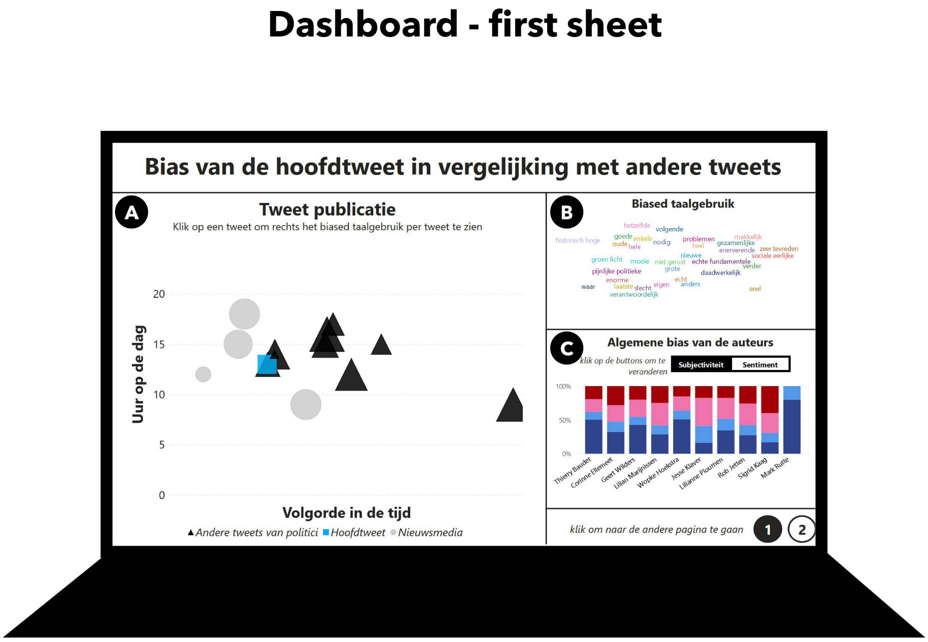


Figure 5.6: The first sheet of the dashboard visualization. (A) shows a timeline visualization in which the x-axis displays the chronological order, and the y-axis displays the hour at which a tweet is published. The main tweet is encoded as a blue square, opposing tweets are encoded as black triangles, and supportive tweets are encoded as grey circles. The marker sizes are proportional to the total calculated bias, which is the absolute sum of a tweet's *subjectivity* and *polarity* (in the experiment, polarity was renamed to *sentiment*). (B) depicts a WordCloud visualization, featuring all the biased words that were picked up by pattern-nl, i.e. the sentiment analysis algorithm used in the current study. The words shown in the WordCloud are specific for the tweets that are featured in (A). (C) contains two percentage stacked bar chart visualizations, one for *subjectivity* and one for *sentiment*. A participant is able to choose which one is displayed by selecting one of the buttons above the corresponding visual. The color codes were made consistent with the context indicators. Only politicians that authored an opposing tweet were featured in these bar charts; news outlets were excluded.

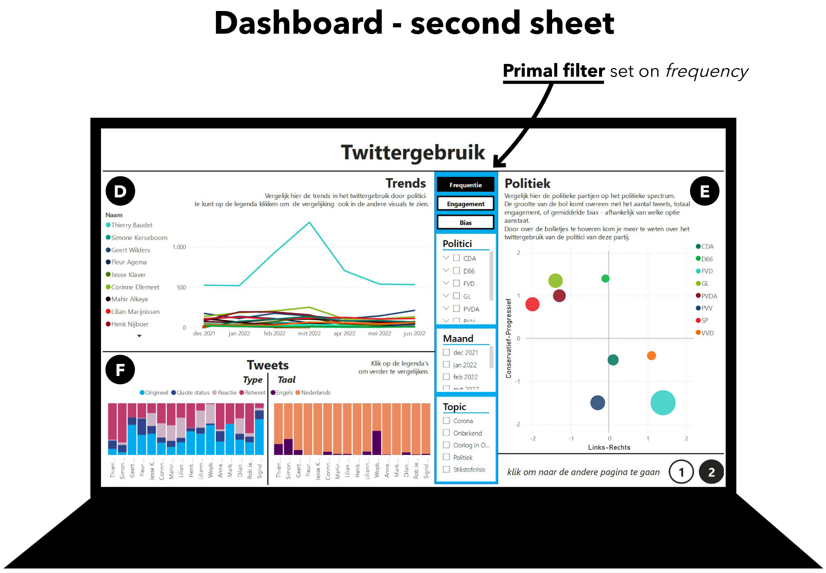


Figure 5.7: The second sheet of the dashboard visualization. Three primal filters were defined for this sheet, determining what type of data the user may compare; *frequency, engagement*, and *bias*. The central blue box contained custom filters that the user could select if interested. The data that is visualized in (D) and (E) is determined by this primal filter. In this figure, the primal filter is set on *frequency*, hence comparing how much politicians tweet. (D) displays a line-chart over time, specified for individual politicians. (E) shows the political spectrum, in which all eight political parties are mapped, with their size proportional to the total amount of tweets their politicians published within the time scope. For visualization (D) and (E), the colors at display were inspired by the political party logo's of the corresponding politicians. (F) is independent of the primal filter, and contains two percentage stacked bar charts; 1) displaying the distribution of tweet type, i.e. *original, retweet, quote* or *direct reply*, shown on the left, and 2) displaying the distribution of tweet language, i.e. English or Dutch, shown on the right.

compare author sentiment relatively, yet no trend over time was made visible. In order to gain more insight into the development of a politician's sentiment over time, an absolute line chart was added by means of a hover effect. This is depicted for an arbitrary politician in figure 5.9.

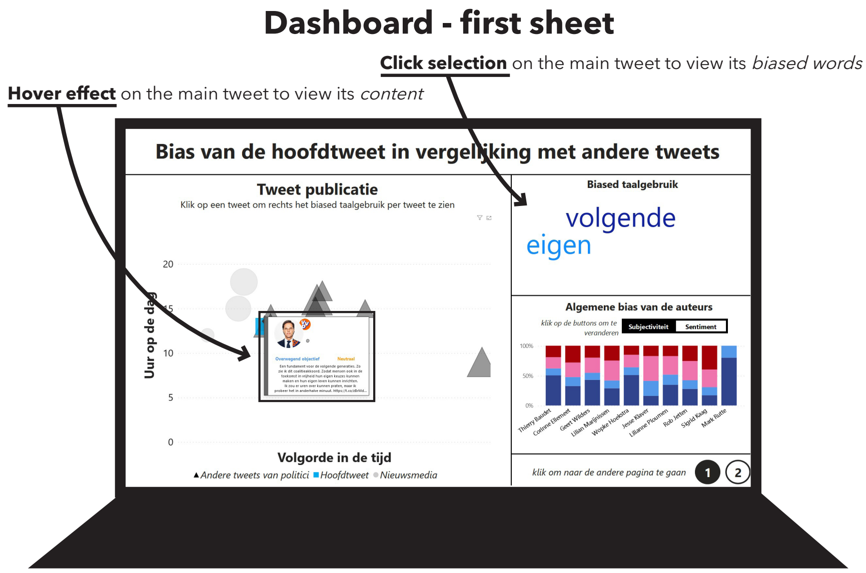


Figure 5.8: The first sheet of the dashboard visualization when inspecting the main tweet.

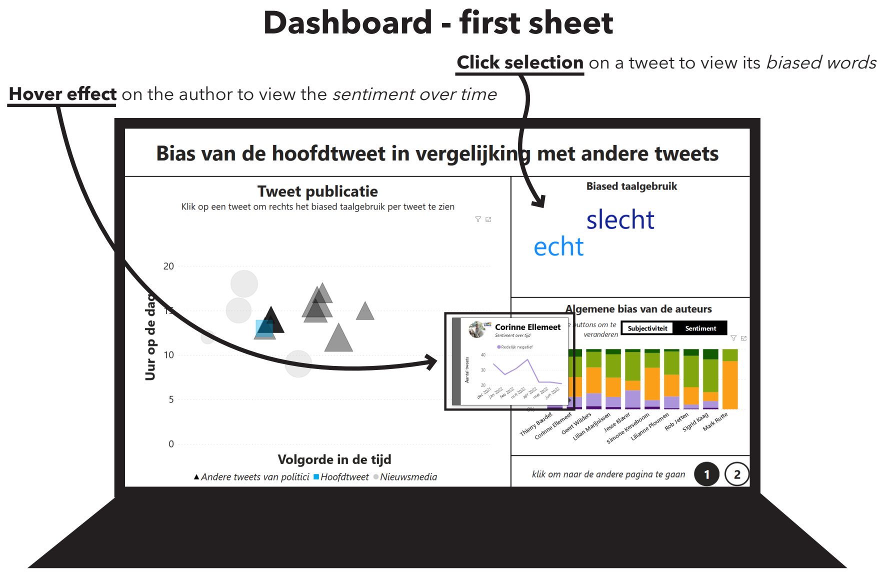


Figure 5.9: The first sheet of the dashboard visualization when inspecting the sentiment of a politician over time.

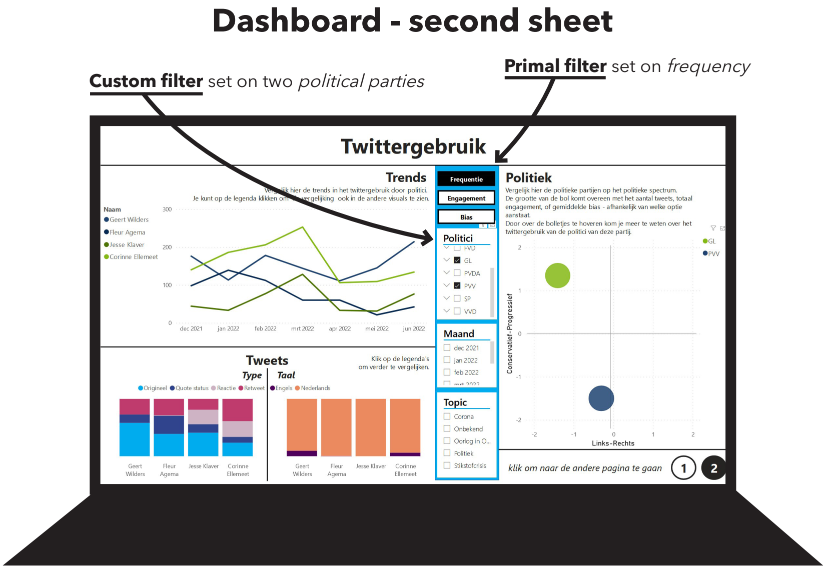


Figure 5.10: The second sheet of the dashboard visualization, with the primal filter set on *frequency* and the custom filter set on the *political parties* PVV and GL.

Second sheet

With an enormous amount of data to display, the interaction idiom in the second sheet was designed to be gradual. Users were first invited to look around and explore with hover effects, yet when interested, could use more custom filters to inspect the data in detail. The custom filters were contained in the central blue box, as is visible in figure 5.7 and figure 5.10. However, one primal filter was already set-up when opening the sheet, which determined the data at display - i.e. either *frequency, engagement* or average *bias*[3]. Figure 5.7 already showed the opening view of the second sheet with the primal filter set on *frequency*. Figure 5.11 shows the sheet with the primal filter set on *engagement*, and figure 5.12 shows the view of the second sheet with the primal filter set on average *bias*. The addition The two percentage stacked bar charts on the left (figure 5.7 (F)) were excluded from linked filtering and were not provided any hover effects, as this would increase the complexity and was not expected to add much additional insights. The only exception that could occur was at times a user would have a filter set on one or more specific politicians - in this case, the bar charts would highlight these selected politicians, as can be seen in figure 5.10.

[3]Bias was computed as the absolute sum of tweets' subjectivity and polarity values. The average bias is the average of this score over all tweets from a politician.

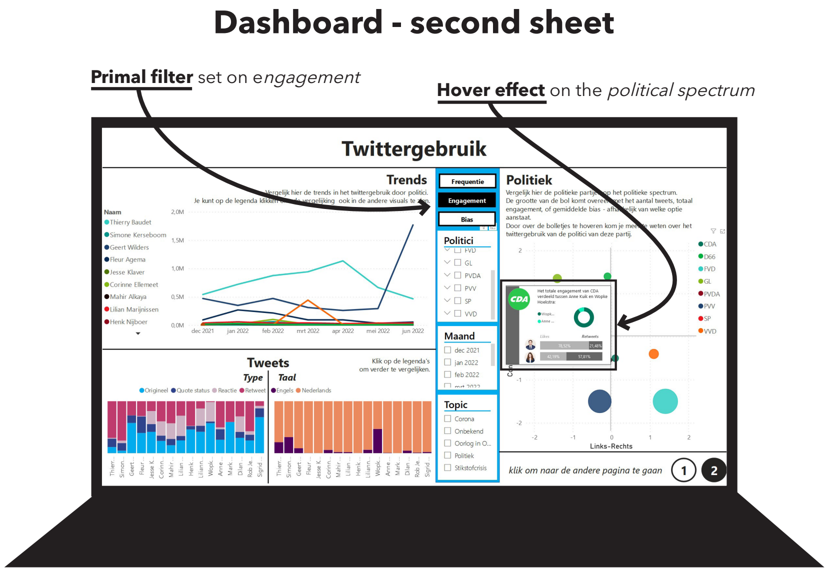


Figure 5.11: The second sheet of the dashboard visualization, with the primal filter set on *engagement* and the hover effect activated for a *political party* in the political spectrum: CDA. The hover effect shows the distribution of engagement for the two politicians that are part of CDA, namely Anne Kuik and Wopke Hoekstra.

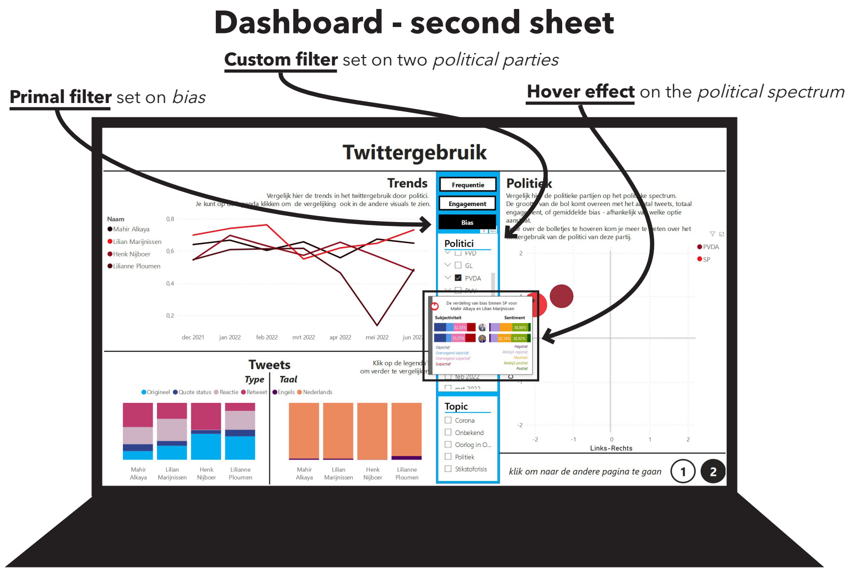


Figure 5.12: The second sheet of the dashboard visualization, with the primal filter set on *bias*, the custom filter set on the *political parties* PVDA and SP, and the hover effect activated for a political party in the *political spectrum*: SP. The hover effect shows the distribution of the bias (i.e., subjectivity and sentiment) for the two politicians that are part of SP, namely Lilian Marijnissen and Mahir Elkaya.

5.4 Technical implementation

This section features a brief technical explanation on how the visualizations, as described in section 5.3, were implemented in the experiment interface.

Frame opposition and context indicators The frame opposition and context indicators were programmed in HyperText Markup Language (HTML), styled by Cascading Style Sheets (CSS), and interactions were facilitated with JavaScript (JS). The static visualizations were designed in Adobe Illustrator AI and exported as Support Vector Graphics (SVG). The SVG items were parameterized, such that they were able to be made interactive with JS. By means of example: when a main tweet was loaded in its placeholder, its attribute value for *sentiment* would be called for in JS, and based on this value, e.g. "neutral", the corresponding "neutral" indicator would be made visible in the SVG object on the HTML page. The placeholders for the main tweet, opposing tweets and supportive tweets, retrieved their data from JavaScript Object Notation (JSON) objects.

Dashboard The dashboard was created with Microsoft PowerBI. For the current study, this tool was the most viable option because of two main reasons; 1) with its accessible entry level, linking graphs and sheets - a necessary property for empowering visualizations - was quickly established, and 2) it allowed for a straightforward HTML embedding, which was desired given the experiment interface setup.

Chapter 6

Experiment Design

Prior to participant recruitment, the experiment design and data analysis set-up were approved by the Ethical Review board of Eindhoven University of Technology. The number of participants, as well as the duration of the experiments, were constrained in order to be able to inspect all data in detail. A common experiment design in framing effects research is to expose participants to a frame, which can be an article, statement or pamphlet, and study participants' interpretation and response to that frame [61, 75, 76, 24]. Typically, the conditions (or *stimuli*) of these frame exposures can vary within or between experiments, and hence, results might tell what conditions influenced participants' interpretation and understanding. The current study was set up in a similar fashion. This chapter describes the set-up in further detail.

6.1 Recruitment of participants

The threshold for the number of participants was set at 21. The participants were recruited via the personal network of the researcher, as this was the most viable option within the scope of the study. This resulted in a sample that was demographically homogeneous, as most participants were university students with a left-progressive political preference. This implies that the results may not be generalized to other demographic groups, yet, do increase the applicability of the results for the demographic group that was included. Including many different backgrounds would have likely yielded a larger variance in the results, which would be extremely - if not impossible - to explain. In the recruitment message, the participants were told no more than that the experiment was going to be about the interpretation of Twitter data. The experiments, taking 45 to 100 minutes each, took place between July 30 and August 10. The participants were not told about possible rewards, but received a chocolate bar as a surprise at completion.

6.2 Experiment

This section discusses setting, flow, participant data analysis approach, and interface of the experiment.

6.2.1 Setting

Out of the 21 experiments planned, 17 took place on the TU/e campus in a private lecture room, 2 took place online via a video connection (Microsoft Teams), and 2 took place at the researcher's residence. The researcher was either physically or virtually present during the whole experiment in order to take notes of observed behavior from participants. The experiment interface was hosted on the researcher's laptop, and in case of physical experiments, participants would interact with this interface via a connected monitor and mouse. At the start of the experiment, the participants were instructed to read a printed text containing an explanation about the three experiment parts,

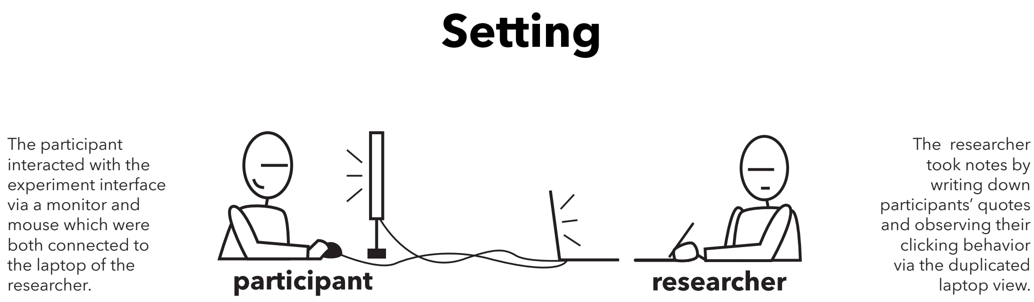


Figure 6.1: The setting of the physical experiments.

which is listed in Appendix C. Additionally, they were provided a list with general definitions about Twitter, as well as a legend to real the context indicators. After reading those instructions, the participants were verbally explained some last remarks about what to expect. Most importantly, the participants were asked to think out loud and enter the experiment as if they would enter a social media environment, e.g. Twitter, Facebook, or Instagram. When everything was clear to the participants, the experiment started. At completion, a small interview took place between the researcher and the participant in order to attain a few more insights where possible. A visualization of this setting is depicted in figure 6.1.

6.2.2 Flow

Every experiment consisted of three parts. In the first part, participants were exposed to a political frame, i.e. the *main tweet*, in three different conditions: 1) without any information visualization, 2) with the additional *context* indicators on top of the tweet as well as frame opposition with *opposing tweets* and *supportive tweets* respectively, and 3) the second condition with the additional interactive *dashboard*. The questions participants were tasked to answer, always referred to the main tweet, and were kept the same in all conditions[1]. Ecery condition features three main tweets, yielding nine main in total. In the second part, participants were tasked to answer a survey about their overall experience. In the third part, participants answered some final questions that served as control variables. This experiment flow is made visual in figure 6.2. The participants were not screened beforehand on issue importance [2].

6.2.3 Approach

The experiment followed a 3x3 within and between subject approach. In order to correct for possible learning effects and participant biases, the participants were divided in three groups of seven, where every group had a different order, referred to as *offset*, in which they were exposed to main tweets. The participants were randomly assigned to a group. This set-up allowed to test the differences in answers between conditions for one participant, as well as it allowed to analyze whether the participants' answers for main tweets differed between conditions in general. A visualization of this approach is depicted in figure 6.3, where (A) shows the main tweets, their authors and their subjects, and (B) shows the set-up approach as described.

[1]In many studies, questions to framing effects are often issue specific. The current study differed from this method to allow for better generalization between frames.

[2]In many studies to framing effects, participants are screened beforehand on perceived issue importance, and later matched to different experiment conditions on those grounds. The current study was designed differently, as the issues depended on the available issues that were covered in the dataset. The selected main tweets therefore did not influence the experiment design, nor the selection of participants. However, issue importance was included in the questions that participants asked after each tweet (i.e. in the first part of the study).

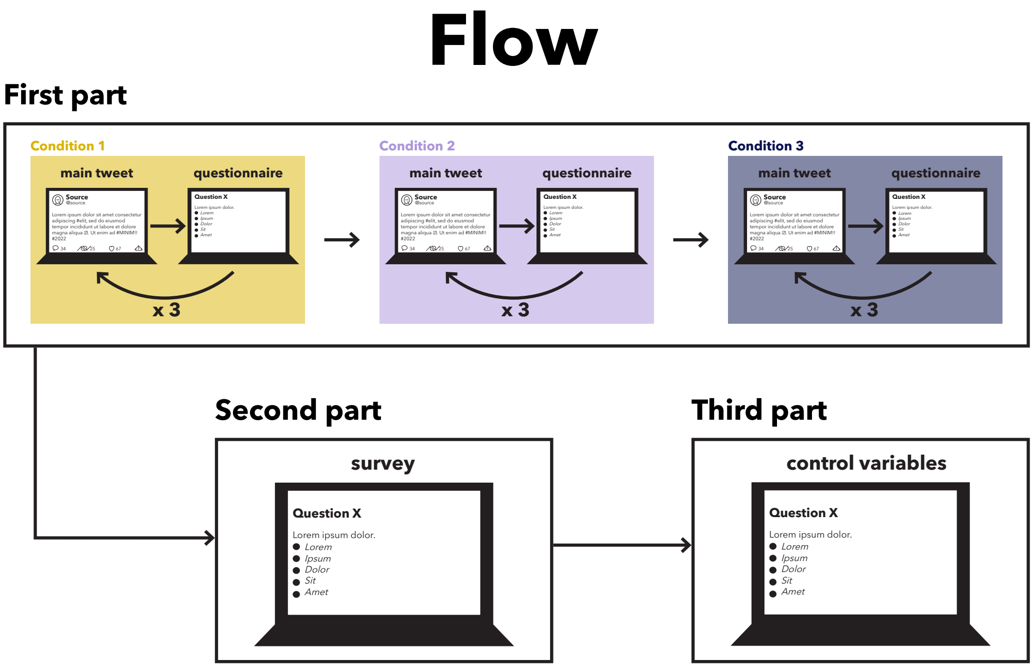


Figure 6.2: The experiment flow that participants followed when taking part in the experiment.

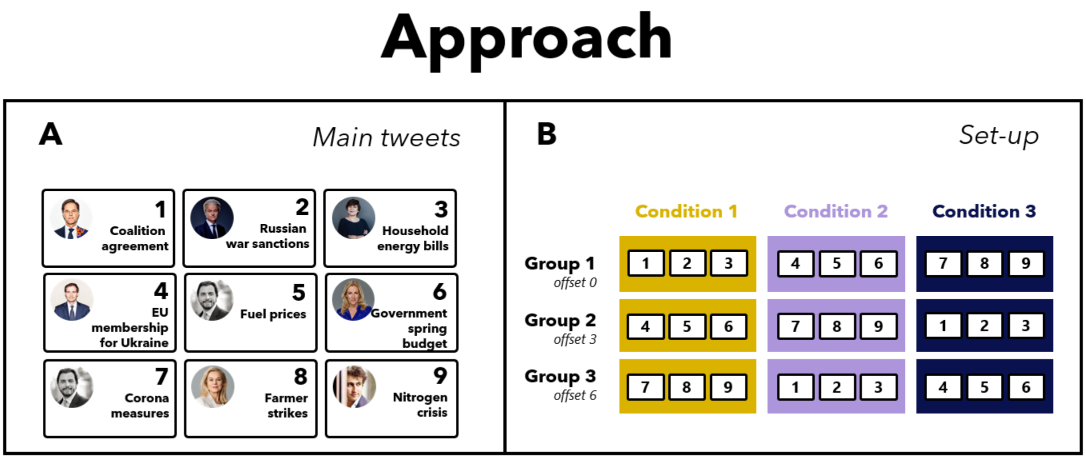


Figure 6.3: The experiment set-up approach, where (A) shows the main tweets and (B) shows the set-up approach

6.2.4 Interface

The experiment interface was aimed to be a closed environment, such that the researcher would maintain control over the interactions and the data. This section highlights two important interface design choices that impacted the experiment.

Controlled Twitter-inspired environment The main tweet was programmed to be in sight at all times, whereas the questionnaire and the opposing and supportive tweets were either visible on request, or alternating. In the third condition, the dashboard remained in sight as well. The experiment interface layouts and the descriptions of their different conditions are depicted in figures 5.2, 5.4, and 5.5.

Questionnaire infrastructure In order to proceed to the next question, participants had to answer to the current question first. Although it was possible to go back to previous questions that belonged to the same main tweet, it was not allowed to go back to a prior main tweet. This was implemented to ensure that participants would not adapt the answers given to main tweets in condition 1), after they were influenced by visualizations in conditions 2) and 3). Answers to all questions were saved in the localStorage of the researcher's browser which was used to host the experiment interface. The data in localStorage was exported in preparation of the analysis.

Chapter 7

Results

Within the overarching focus of the current study, which was to trigger analytic reasoning with information visualization, two research goals were defined; 1) create more bias awareness, and 2) mitigate possible political framing effects. In light of these research goals, the experiment was set up in a way such that research goal 1) was measured with a self-reported questionnaire responses, retrieved from the second part of the experiment, and research goal 2) was measured by using within-group and between-group comparison of participant responses, obtained in the first part of the experiment. These participant responses were registered on a 5-point Likert scale (either *agreement* or *frequency*), yet converted to numeric in order to facilitate a statistical analysis. The numerical scale that the Likert answers were mapped onto is found in table 7.1. Moreover, participants' observed behavior was analyzed, as this provided with additional insights about the way the visualizations were perceived. This was ultimately helpful in finding the answer to the research question: "How can information visualization increase political bias awareness and help mitigate the political framing effect on social media?". The analysis for these three components of the study will be summarized in this chapter. In addition, it will argument which hypotheses, as stated in chapter 3, are rejected, and which are not. These preliminary conclusions will however be vigilant, given the small sample size of the user study.

Agreement	Frequency	Value
Strongly disagree	Never	1
Disagree	Rarely	2
Neither agree nor disagree	Sometimes	3
Agree	Often	4
Strongly agree	Always	5

Table 7.1: The numeric values that correspond to the 5-point Likert scales that were featured in the questionnaires in the first part of the experiment.

7.1 Increase of bias awareness

The distributions of the survey outcomes, generated in the second part of the experiment, are visualized in figure 7.1. This section discusses them in further detail, specified per type of visualization.

7.1.1 Context indicators

The majority of participants (76,2%) reported that the context indicators made them more aware of the political bias of the main tweet, and an additional 4,8% strongly agreed to this statement.

However, less consistence was found for sentiment awareness, where 28,6% strongly disagreed that the sentiment indicator made them more aware of the main tweet's sentiment. The distribution to responses on the statement whether the context indicators triggered critical thinking were close to uniform. For the majority of participants, the context indicators neither changed their attitude towards the author of the main tweet (57,1% disagreed and 14,3% strongly disagreed), nor did the context indicators change their own political standpoint (47,6% disagreed and 19% strongly disagreed). However, the majority of participants found the context indicators insightful (52,4% agreed and 33,3% strongly agreed). The context indicators therefore seem to have met their goal, which was to create more bias awareness, yet, how they actually achieved this is not fully certain, since the sentiment indicator had little effect, and they were unable to change participants' opinion.

7.1.2 Dashboard visualization

A strong majority responded that the dashboard contained too much information(71,4% agreed and 14% strongly agreed). Most participants (57,1%) neither agreed nor disagreed on whether the dashboard was a valuable addition to the context indicators, whereas more participants disagreed (23,8%) or strongly disagreed (4,8%). Likewise, responses on the dashboard's ability to contextualize the main tweet yielded a majority of 52,4% that neither agreed nor disagreed, yet more participants disagreed (28.6%) or strongly disagreed (9.5%). Similar to the responses to the context indicators, there was no clear outcome on whether the dashboard triggered participants' critical thinking. However, though the context indicators were found insightful by a strong majority, the dashboard was not clearly perceived to give participants more insights.

7.1.3 Frame opposition

The frame opposition was found not to change the political standpoints of the participants, as 47,6% disagreed and 19% strongly disagreed. For both statements on whether the opposing frames changed the participants' view on the main tweet's subject and on whether the opposing frames provided the necessary information to form one's opinion, no clear outcome could be found. Based on these results, it is impossible to say whether the frame opposition achieved its goal, which was to mitigate framing effects.

7.1.4 Social media preferences

The survey included two questions that did not directly touch upon a visualization that was included in the experiment, but questioned a participant's social media preferences instead. This was done to gain more insights into the extent to which literature on selective exposure and filter bubble effects would hold for the current sample. Interestingly, a strong majority of participants expressed a preference to be exposed more to different political content (57,1% agreed and 23.8% strongly agreed). When asked whether participants would use a dashboard, one that is similar to the dashboard that was included in the experiment, a slight majority of 52,4% agreed and 4,8% strongly agreed. However, 23,8% disagreed. This is particularly interesting, given the fact that participant responses to the dashboard specific questions would suggest that they'd rather not use a dashboard, given its information overload.

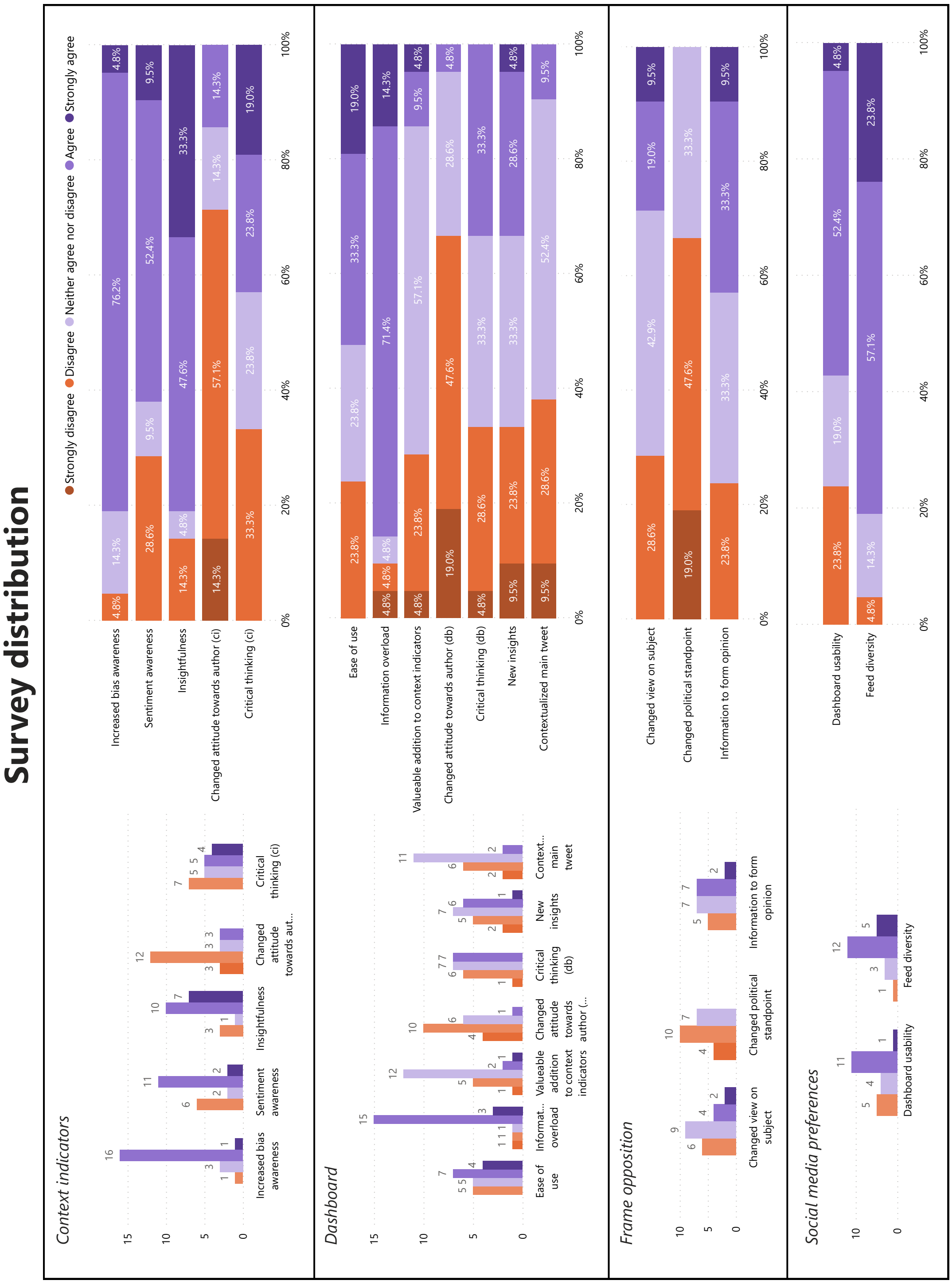


Figure 7.1: The distribution of participants' survey responses.

7.2 Mitigation of possible political framing effects

The political framing effect was defined in chapter 1 as the presentation of the same political issue or common problem in a different way to alter citizen's attitudes, emotions or behavior [22]. A mitigated political framing effect was defined as an increased sense of awareness at the user about the versatility of policy issues and the bias of the source as result of an intervention, where the intervention is considered to be the exposure to information visualization. Chapter 6 on experiment design explained that participants were to be exposed to nine main tweets in total, in three different conditions - yielding three tweets per condition. The questionnaires included after each individual main tweet were kept the same, such that the participant responses could be compared between conditions. Those comparisons could ultimately show whether the visualizations had any influence. This section describes this analysis.

7.2.1 Statistical testing

Four statements that were featured in a main tweet's questionnaire are of significance in testing for framing effects:

1. The tweet's subject can be viewed from multiple perspectives.
2. The tweet's author approaches the subject of the tweet from multiple perspectives.
3. I agree with the author of the tweet about this subject.
4. My own perspective on the tweet's subject has been changed.
5. The tweet got me thinking about this subject.

The answer distributions for these statements did not meet the normality assumption in any of the conditions, hence, a non-parametric test needs to be applied. Since the experiment included dependent groups, measured over three or more conditions with responses in ordinal form, the most appropriate test was found to be the Friedman's test. The Friedman's test is a non-parametric alternative to the repeated measures ANOVA that detects whether the distributions under different conditions are the same [77, 78]. The null hypothesis and the alternative hypothesis for the Friedman's test read:

H0 The median difference between conditions $= 0$.

H1 The median difference between conditions $\neq 0$ for at least one of the conditions.

The Friedman's test was performed for each of the five above mentioned statements, and executed in the statistical software called RStudio. The p-value indicates the significance of a treatment effect, and the effect size indicates the strength of the relationship by calculating Kendall's W value, which maps on a domain of [0,1][1]. The analysis was complemented with a non-parametric Mann-Whitney Wilcoxon test for pairwise comparisons of experiment conditions. No Bonferroni correction was applied to protect the family-wise error rate, since there seems to be no clear consensus in literature on whether and in what situations its application is actually most suitable.

7.2.2 Analysis

This section highlights the analysis results for each of the five statements, for which a description of the expected participant answer pattern is given upfront, followed by a discussion of the statistical test results and a conclusion. The visual report that is added to the discussions, contains a Box and Whisker plot, which is always visualized in subfigure (C). This plot displays the min and max

[1] Kendall's W, also referred to as the *coefficient of concordance*, is a non parametric statistic used to assess the level of agreement between raters on a domain of [0,1]. Here, 0 indicates no agreement, and 1 indicates perfect agreement. A value of 0 hence implied that all participants had different summed answer values [79].

value at the ultimate endpoints, first and third quartile as outlines of the quartile box, the mean as the dot inside this box, and the median as the horizontal line inside this box. The value for alpha is, in line with literature, set at 0.05[2]. Results at a p-value smaller or equal to 0.05 are complemented with one *, results at a p-value smaller or equal to 0.01 are complemented with two **, and results at a p-value smaller or equal to 0.001 are complemented with three ***.

1. The tweet's subject can be viewed from multiple perspectives.

The results to this statement are depicted in figure 7.2.

Expectation Political frames tend to highlight only one aspect of an issue, as discussed in chapter 1. Therefore, user responses to this statement were expected to be more apt to "Strongly agree" or "Agree" in conditions 2 and 3 when compared to condition 1, since the frame opposition would expose the participants to other perspectives.

Friedman's test The results of Friedman's test are depicted in figure 7.2 (A) and show that the different experiment conditions were of no statistical significance in explaining the participant responses ($p = 0.844$). Moreover, a negligible effect size of 0.0085 was reported.

Mann-Whitney Wilcoxon pairwise comparison Also at inspection of pairs of experiment conditions, the participant responses were not able to be explained by these conditions, as all pairwise p-values were larger than the alpha value of 0.05 (figure 7.2 (B)).

Conclusion Neither the context indicators and frame opposition contained in condition 2, nor the addition of the dashboard to these visualization in condition 3, triggered participants to think that a tweet's subject can be viewed from multiple perspectives.

2. The tweet's author approaches the subject of the tweet from multiple perspectives.

The results to this statement are depicted in figure 7.3.

Expectation The participant responses were expected to be more apt to "Strongly disagree" or "Disagree" in these conditions when compared to condition 1, as the exposure to different opposing tweets would prove the opposite of the statement. Moreover, The third experiment condition, featuring the dashboard, was expected to amplify this even more, as participants were able to look up more about the main tweet's author and his or her political background.

Friedman's test The results of Friedman's test are depicted in figure 7.3 (A) and show that the different experiment conditions were of no statistical significance in explaining the participant responses ($p = 0.545$). Like with the first statement, a minimal effect size was reported (0.0289).

Mann-Whitney Wilcoxon pairwise comparison The pairwise comparison yielded two statistically significant results, namely between conditions 1 and 2 ($p = 0.037$*) and between conditions 1 and 3 ($p = 0.002$**). When taking a closer look at figure 7.3 (C), one can see that the quartiles for condition 2 and 3 indeed fall out lower than for condition 1, in line with the expectation that was described earlier. Moreover, figure 7.3 (D) shows that the sum of answers for "Agree" declines in step-wise between condition 1, condition 2, and condition 3. In parallel, the same figure shows that the sum of answers for "Disagree" inclines from condition 1 to conditions 2 and 3.

[2]The alpha value is the probability of making a Type I error, which is to conclude a difference exists while there, in fact, is none. An alpha value of 0.05 means that a risk of 5% of type I errors is accepted.

"The tweet's subject can be viewed from multiple perspectives."

A *Friedman's test*

Non-parametric test for treatment effects

Y	N	Statistic	DF	P-value	Effect size
answerVal	21	0.338	2	0.844	0.0085

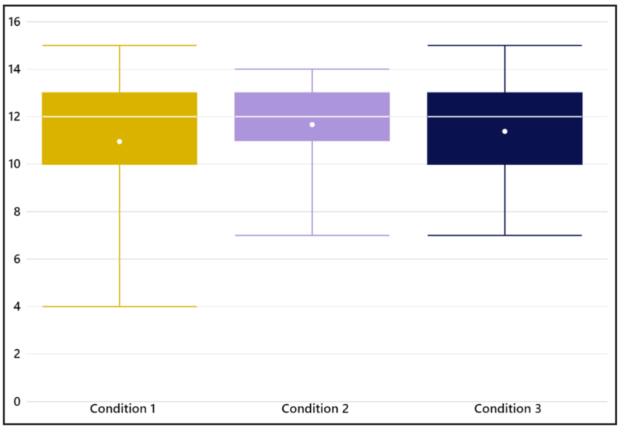
B *Mann-Whitney Wilxocon*

Pairwise comparison based on ranked sums

Y	Group I	Group II	N I	N II	Statistic	P-value
answerVal	Condition 1	Condition 2	21	21	27.5	0.060
answerVal	Condition 1	Condition 3	21	21	18	0.096
answerVal	Condition 2	Condition 3	21	21	36	0.095

C *Box and Whisker plot*

Displaying the min, max, quartiles, mean, and median values for experiment conditions



D *Answers*

Absolute distribution

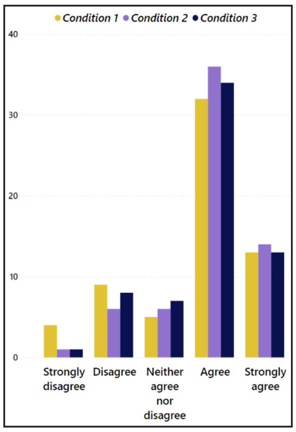


Figure 7.2: The test results for the first statement.

Conclusion The results of the Mann-Whitney Wilcoxon test and closer inspection of figures 7.3 (C) and (D) imply that with more information visualization, less participants agreed that the main tweet's author approached the tweet's subject from multiple perspectives, and more participants disagreed that the main tweet's author approached the subject from multiple perspectives.

3. I agree with the author of the tweet about this subject.

The results to this statement are depicted in figure 7.4.

Expectation In answer to this statement, participant responses were expected to be more tilted to the extremes on the 5-Point Likert scale. On the one hand, participant responses could lean more towards "Strongly disagree" or "Disagree" in conditions 2 and 3 with the exposure to more arguments and information, and on the other hand, the same holds for the answers "Agree" and "Strongly agree". More information visualization was expected to make the participants better able to critically compare arguments and decide whether or not the argument of the main tweet is probable (or not).

Friedman's test The Friedman's test yielded no statistically significant result that could reject the null hypothesis ($p = 0.959$). The effect size is negligible (0.00198).

Mann-Whitney Wilcoxon pairwise comparison The pairwise comparison of experiment conditions showed no statistically significant differences ($p > 0.005$ for all pairwise comparisons). Closer inspection of the plot in figure 7.4 (C) shows that the quartile boxes, means and medians are approximately the same in all experiment conditions.

Conclusion The results show that a participant's agreement to a tweet is not subject to change as a result of the exposure to information visualization.

4. My own perspective on the tweet's subject has been changed.

The results to this statement are depicted in figure 7.5.

Expectation Similar to the third statement, participant responses for this statement were expected to be more tilted to the extremes on the 5-Point Likert scale. Especially the addition of the dashboard in condition 3 might trigger the analytical thinking of participants with respect to the credibility of the main tweet's author, as the dashboard contained data about politicians' Twitter behavior.

Friedman's test The Friedman's test result was not able to reject the null hypothesis, given a p-value of 0.397. The effect size was again minimal, at 0.0440.

Mann-Whitney Wilcoxon pairwise comparison Interestingly, a statistically significant difference did appear to exist between experiment conditions when comparing them pairwise, with a p-value of 0.025* for the comparison of conditions 1 and 2, and a p-value of 0.041* for the comparison of conditions 1 and 3. Figure 7.5 (C) shows that the median value for condition 2 and 3 is at the third quartile, and the quartile box for condition 3 is rather concentrated around the mean.

Conclusion The median values for conditions 2 and 3 fall out higher than for condition 1, which indicates that the distribution of summed answers is more tilted to "Agree" and "Strongly agree". Moreover, pairwise comparisons between the first and second, and first and third conditions were found to be significant ($p = 0.025$ and $p = 0.041$, respectively). This would indicate that with exposure to more visualizations, participants' perspectives on the main tweet's subject are more

"The tweet's author approaches the subject from multiple perspectives."

A *Friedman's test*

Non-parametric test for treatment effects

Y	N	Statistic	DF	P-value	Effect size
answerVal	21	1.22	2	0.545	0.0289

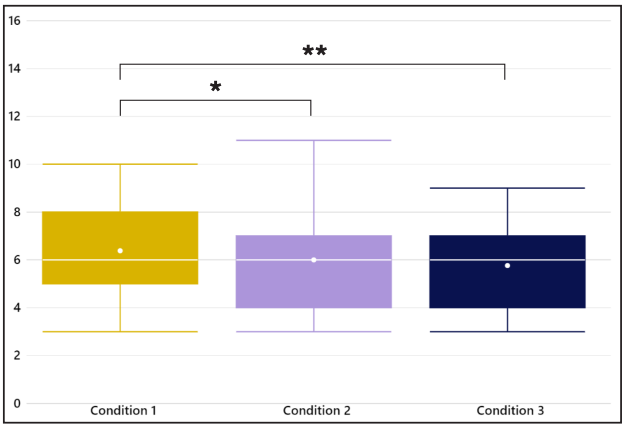
B *Mann-Whitney Wilxocon*

Pairwise comparison based on ranked sums

Y	Group I	Group II	N I	N II	Statistic	P-value
answerVal	Condition 1	Condition 2	21	21	55	0.037*
answerVal	Condition 1	Condition 3	21	21	98	0.002**
answerVal	Condition 2	Condition 3	21	21	54	0.212

C *Box and Whisker plot*

Displaying the min, max, quartiles, mean, and median values for experiment conditions



D *Answers*

Absolute distribution

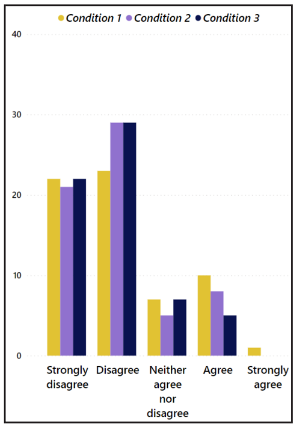


Figure 7.3: The test results for the second statement

"I agree with the author of the tweet about this subject."

A *Friedman's test*

Non-parametric test for treatment effects

Y	N	Statistic	DF	P-value	Effect size
answerVal	21	0.0833	2	0.959	0.00198

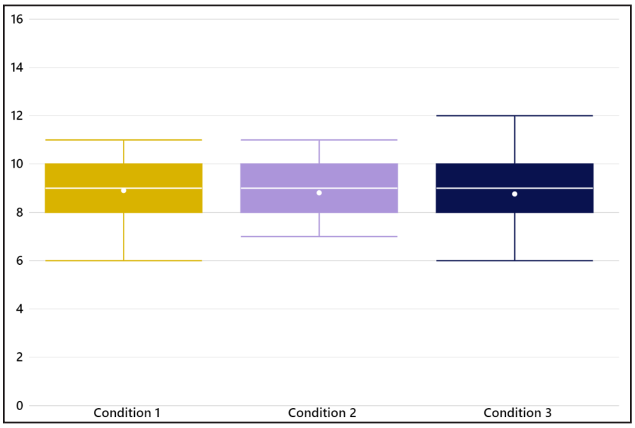
B *Mann-Whitney Wilxocon*

Pairwise comparison based on ranked sums

Y	Group I	Group II	N I	N II	Statistic	P-value
answerVal	Condition 1	Condition 2	21	21	14	0.484
answerVal	Condition 1	Condition 3	21	21	30	0.351
answerVal	Condition 2	Condition 3	21	21	25	0.790

C *Box and Whisker plot*

Displaying the min, max, quartiles, mean, and median values for experiment conditions



D *Answers*

Absolute distribution

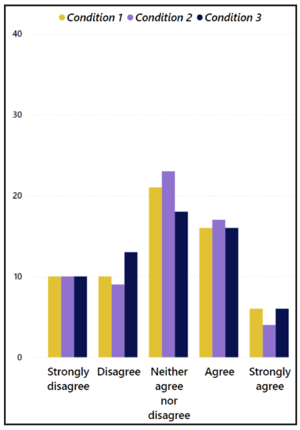


Figure 7.4: The test results for the third statement.

likely to be changed. However, this conclusion is rather tricky - figure 7.5 (D) shows that only few people actually agreed to the statement. Therefore, a more nuanced conclusion would be that with exposure to more information visualizations, less participants strongly disagreed that their perspectives changed because of the tweet.

5. The tweet got me thinking about this subject.

The results to this statement are depicted in figure 7.6.

Expectation The participant responses were expected to be tilted more towards "Agree" and "Strongly agree" in conditions 2 and 3 for the last statement, since the visualizations in these conditions were intended to trigger participants to think more analytically about the content of the main tweet. Moreover, the exposure to more information was expected to increase a participant's deliberation time, and this was found to influence participants in their social media behavior (Chapter 2).

Friedman's test The Friedman's test result was not significant ($p = 0.330$), hence supporting the null hypothesis. The effect size was again small, at 0.0528.

Mann-Whitney Wilcoxon pairwise comparison The pairwise comparisons for this statement yielded no statistically significant results. This is not strange, as the answer distribution in figure 7.6 (D) show no evident answer pattern.

Conclusion The information visualizations did not get participants to think more critically about tweet subjects.

"My own perspective on the tweet's subject has been changed."

A *Friedman's test*

Non-parametric test for treatment effects

Y	N	Statistic	DF	P-value	Effect size
answerVal	21	1.85	2	0.397	0.0440

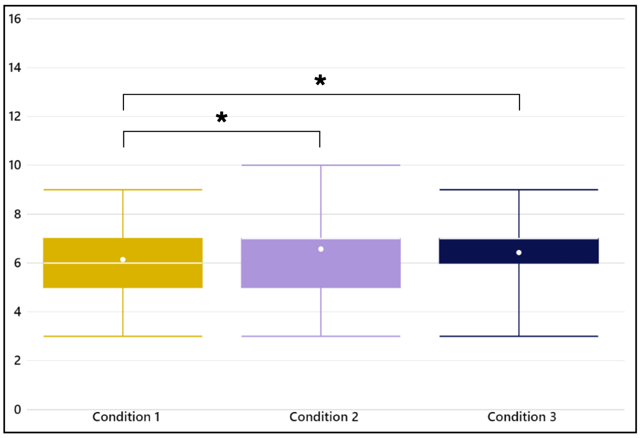
B *Mann-Whitney Wilxocon*

Pairwise comparison based on ranked sums

Y	Group I	Group II	N I	N II	Statistic	P-value
answerVal	Condition 1	Condition 2	21	21	4	0.025*
answerVal	Condition 1	Condition 3	21	21	4.5	0.041*
answerVal	Condition 2	Condition 3	21	21	24	0.407

C *Box and Whisker plot*

Displaying the min, max, quartiles, mean, and median values for experiment conditions



D *Answers*

Absolute distribution

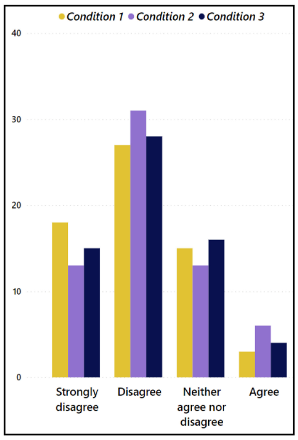


Figure 7.5: The test results for the fourth statement.

"The tweet got me thinking about this subject."

A *Friedman's test*

Non-parametric test for treatment effects

Y	N	Statistic	DF	P-value	Effect size
answerVal	21	2.22	2	0.330	0.0528

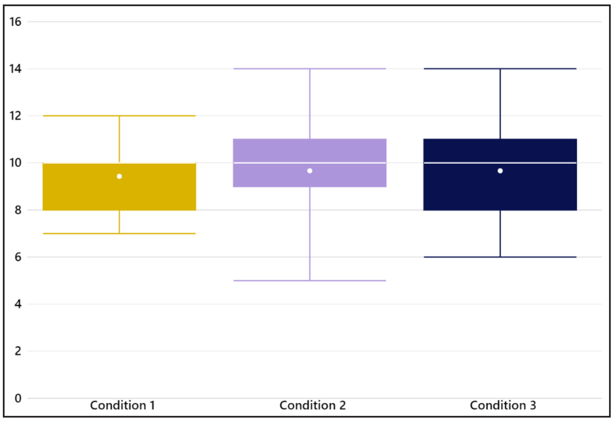
B *Mann-Whitney Wilxocon*

Pairwise comparison based on ranked sums

Y	Group I	Group II	N I	N II	Statistic	P-value
answerVal	Condition 1	Condition 2	21	21	34.5	0.444
answerVal	Condition 1	Condition 3	21	21	20.5	0.256
answerVal	Condition 2	Condition 3	21	21	27.5	1

C *Box and Whisker plot*

Displaying the min, max, quartiles, mean, and median values for experiment conditions



D *Answers*

Absolute distribution

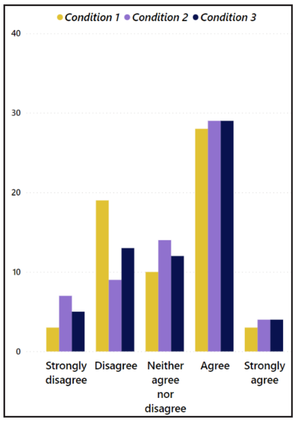


Figure 7.6: The test results for the fifth statement.

7.3 Observations

Table 7.2 list the most frequently occurring quotes from experiment observations, where the "count" column indicates the number of times a quote was mentioned by different participants[3]. Only quotes that were observed more than two times throughout all experiments were included. In order to structure the analysis, the quotes considered to be of importance were grouped on *type*, i.e. to which component of the study the quote applies, and *subtype*, i.e. a specification of that type. The four types identified were 1) tweet, 2) political spectrum, 3) Twitter, and 4) visualizations. This section highlights the observations specified per type in further detail.

7.3.1 Tweet

First, the participants' judgement of a main tweet's subjectivity was mainly dependent on the language used, with adjectives in particular. Besides the fact that this is a positive note on the ability of participants to identify biased language, it also indicates that the development of natural language processing for biased language detection can be of considerate value in social media context. Second, most participants had to open the links that were attached to seven out of nine main tweets in order to know what the tweet was about. This can be seen as a "shift in the burden of proof", as Twitter has no control over what external sources their users include when posting tweets. Moreover, no less than 13 participants mentioned that they needed more knowledge in order to form an opinion. Third, with respect to argumentation and persuasiveness, it appeared that politicians were unable to change participants' opinions in case they already had a strong one of their own, yet, at times this was not the case, participants either felt like they were still left without an opinion, adopted the opinion from the main or opposing tweets, or started to form their own opinion. Fourth and last, three participants mentioned they found something particularly important as soon as it felt close to them. Moreover, several participants expressed a strong disapproval towards some politicians.

7.3.2 Political spectrum

One of the statements that was included in the questionnaire after every main tweet, asked the participants to classify the main tweet's statement according to the political spectrum. This appeared to be a confusing and difficult exercise for many participants, as the statements coming from main tweets were not always perceptually congruent with this spectrum. In particular, this was the case for main tweet subjects that related to the European Union, or were written by populist politicians. This question was therefore kept out of the statistical analysis.

7.3.3 Twitter

Out of all participants, one reported to use Twitter actively, four reported to use the platform rarely, and 16 participants reported not to use it at all. Those who consider themselves active or rarely active characterized their use as passively consuming, rather than actively posting. Four participants mentioned that they would normally not open a link if they would use Twitter in daily life, on their own feed, yet because of the experiment they made an exception. Moreover, many participants would not retweet political content at all, because they actually never post - it's not how they use Twitter. However, 11 participants explicitly mentioned that they will not engage with political content in particular, as they do not want to be associated with certain politicians or political standpoints.

[3] The participants were asked to think out loud when participating in the experiment. Therefore, their quotes reflect those thoughts.

7.3.4 Visualizations

The participants gave feedback on the visualizations in the survey, and in an additional small reflection with the researcher at completion of the experiment.

Information inconsistencies First, many participants experienced that their personal interpretation of the bias in the main tweet did not resonate with the context indicators. However, some mentioned that despite or because of this discrepancy, they were activated to think more about the political bias. The frame opposition was in turn found to be helpful in understanding how issues can be viewed from multiple perspectives.

Participant's dashboard findings Second, although the results from the survey, which were discussed in section 7.1 and depicted in figure 7.1, indicated that the dashboard contained too much information, some participants were still provided with new insights when using it. Because the extent to which this happened differed widely between participants, many quotes were unique, and their count was below the threshold of 2. However, it was mentioned by three participants that the dashboard contained surprising information. For example, three participants mentioned that their expectation of a specific politician's average subjectivity was it to be high, whereas the dashboard visualization showed that, according to the algorithm performed on the data, this politician was on average more objective instead.

Complexity versus concreteness of information visualizations Third, many participants considered the dashboard an overkill of information, it being too *complex*, and three participants explicitly mentioned that the context indicators were more *concrete* and therefore more user-friendly than the dashboard for that reason. However, participants' preferences with respect to those context indicators were not consistent. Overall, it appeared that the subjectivity and political color indicator were appreciated more than the sentiment indicator, as most participants considered themselves capable to tell a tweet's sentiment by themselves. With respect to the design of the context indicators, three participants mentioned that they were easy to understand. However, some participants actually overlooked the context indicators completely.

Participant suggestions Fourth and last, some participants suggested that for possible dashboard implementations in the future, they would envision a dashboard that is standalone and not embedded as an addition to an individual tweet.

Type	Subtype	Count	Quote
Tweet	Interpretation	10	I think the tweet is subjective because of the language.
		7	I think everything can be viewed from multiple perspectives.
		5	"Mooie praatjes"
		4	The tweet is factful, but also includes a political standpoint.
		3	I think every tweet is subjective.
		3	The author doesn't state any facts.
		3	You could see the subject from multiple perspectives, but this politician does not do that.
		3	The information that was added in the link was too one-sided.
	Knowledge	15	After opening the link, I know what the article is about.
		13	Too little knowledge about the topic to have an opinion.
	Argumentation	8	I did not have an argument yet, therefore it hasn't changed.
		7	Too little arguments are provided to change my opinion.
		6	I have a strong opinion myself.
	Personal	3	I find the topic of the tweet personally important, because it relates to my daily life.
		3	Knowing what this politician usually says, I would never retweet this politician.
		3	I really don't like this politician.
Political spectrum	Interpretation	9	I find it hard to classify this tweet according to the political spectrum.
Twitter	Consumption	4	I would normally never watch clips by this politician, however, because of the experiment I would.
		3	I am not in the mood to watch a 7minute video, yet I am curious what it says.
	Engagement	11	I would never retweet a politician or political standpoint.
Visualizations	Experience	11	I disagree with the context indicator.
		3	Definitely to be viewed from other perspectives, as was shown in other tweets.
	Dashboard findings	3	I did not expect this politician to be objective and neutral compared to others.
	Feedback	8	The dashboard was an overkill.
		7	I could identify the sentiment myself, however political color and subjectivity were interesting.
		3	The dashboard was a lot of information, whereas the context indicators were much more concrete.
		3	I appreciated to see which words were biased in a tweet.
		3	I actually overlooked the context indicators.
		3	The context indicators were easy to understand.
		3	The context indicator was however triggering to think more critically, although they were often wrong.
	Future work	3	I would use the dashboard in case I would want to go vote
		3	I would like to take my time with the dashboard and inspect multiple things at once, rather than going back to the dashboard again and again

Table 7.2: Most frequently occurring quotes from experiments (count > 2)

7.4 Hypotheses

The hypotheses as formulated in chapter 3 read:

H1 Visual context indicators increase political bias awareness.

H2 Frame opposition, applied to political tweets, mitigates possible political framing effects.

H3 The interactive visualization of aggregate Twitter data by means of a dashboard, allows participants to better understand the context of political tweets.

H1 and H3 relate to the first research goal of the current study, which was to increase participants' bias awareness. Based on the evidence that was sourced by self-reported questionnaires, H1 is accepted, as the context indicators were found to provide participants with an increased sense of awareness about the main tweet's political bias. H3 is rejected, as only 9.5% of participants experienced that the dashboard helped them to put main tweets in context. However, some participants did appreciate the dashboard, but preferred to use the tool in a setting outside of the social media environment. H2 relates to the second research goal of the experiment, which was to mitigate possible political framing effects. The non-parametric Friedman's test was used to detect a statistical difference in participant responses between conditions. This Friedman's test was done for five of the questionnaire statements, and was complemented with a Mann-Whitney Wilcoxon test for pairwise condition comparisons. None of the performed Friedman's tests was significant, hence, its null hypothesis - the median differences between all conditions = 0 - was not rejected. However, for the second and fourth statement, i.e. "The tweet's author approaches the subject from multiple perspectives" and "My own perspective on the tweet's subject has been changed" respectively, statistically significant differences were found between condition 1 and conditions 2 and 3. This was, yet, not found to be enough evidence to support H2.

Chapter 8

Conclusion and Future Work

This chapter contains a discussion of the conclusion that can be drawn from the current study, as well as it argues what limitations were of impact on the experiment results. It concludes with suggestions for future work.

8.1 Conclusion

This section first discusses the hypotheses, after which it answers the research question and states a final conclusion.

8.1.1 Discussion of hypotheses

The focus of the current study was to stimulate analytical reasoning of potential social media users with information visualization, for which two research goals were set; 1) create more bias awareness, and 2) mitigate political framing effects. With respect to the first research goal, the user study (N=21) showed that participants became more aware of political bias of tweets when confronted with article-level information visualizations, accepting Hypothesis 1. However, the dashboard visualization was not found to suit the social media context. As no consistent evidence was found on the added value of the dashboard visualization, Hypothesis 3 was rejected. With respect to the second research goal, the comparison of participant responses between different conditions yielded no strong evidence on whether the different conditions influenced participants' susceptibility to framing effects. These results suggests that the participants taking part in the experiment were either already considerably resilient to framing effects, or were rather susceptible to motivated reasoning when taking part in the experiment.

8.1.2 Research question

The research question stated in chapter 1 reads:

How can information visualization increase political bias awareness and help mitigate the political framing effect on social media?

The context indicators were found to increase participants' self-reported political bias awareness. This implies that static, contextual and concrete information visualization may enhance the political bias awareness of social media users. Evidence on the influence of frame opposition for political framing effect mitigation on social media was less clear from the current study. However, a majority of participants did indicate that they would favor a more politically diversified social media feed. This implies that the sample was willing to step out of the homogeneous frames it is currently exposed to. The dashboard visualization, that could potentially contribute to both an increased sense of political bias awareness as well as mitigation of framing effects, was found to be unsuitable when embedded in a social media environment. However, participants did mention that the dashboard would help them in different settings, for which a voting support tool was mentioned as a possible application.

8.1.3 Conclusion

These findings showcase the complexity of the social media landscape, as even within the demographically narrow sample, participants reacted differently to both the main tweets as well as the different visualizations. Furthermore, the impact of a political tweet was found to go beyond its limit of 280 characters. As many politicians include links to external websites, social media users are exposed to even more stimuli when they leave the platform to visit those links. All these stimuli can influence one's perception in many possible ways. Moreover, the the consumption of information on social media has become increasingly volatile. Low-effort and concrete visualizations, e.g. context indicators, seem to win in applicability over complex ones, e.g. a dashboard.

8.2 Limitations

This section discusses the limitations of the current study, which impacts were present on seven levels; 1) cognitive biases that were not tested for, 2) the sample, 3) the experiment set-up, 4) the questionnaires, 5) the Twitter dataset, 6) the realization of the visualizations, and 7) the statistical testing procedure.

Cognitive biases that were not tested for The scope of the current study did not allow to go too much in-depth with testing for the cognitive bias of motivated reasoning and deliberation time. First, the political preference of participants was included as a control variable, yet only used for descriptive purposes. Second, since participants' "usual" social media consumption behavior was not known, the within-group comparison cannot be made to "baseline" social media consumption behavior. Therefore, the influence of deliberation time on information interpretation was not tested, although it may have occurred.

Sample First and foremost, the sample of the current study is small and homogeneous. A larger sample could yield more robust results on statistical significance or insignificance, where the latter would mean that the visualizations were not impactful. The homogeneity in participants reduced the demographic variance of the sample. The current study would have benefited from a more representative sample in relation to the problem statement (chapter 1), as only few participants were active on Twitter, a minority would generally share political content on social media, and a vast majority of participants identified most as left-progressive. Furthermore, as all of the participants were university students, a strong age and education bias is visible in the results. Importantly, many participants have already been exposed to or have yet interacted with dashboards in their studies or jobs, whereas that does not hold for any arbitrary Dutch citizen, whom would not intuitively know how a dashboard works. The results are, therefore, descriptive for the population that the sample represents, but cannot be generalized to other demographic samples.

Set-up First, although the layout of the main tweets, opposing tweets and supportive tweets were a direct copy of how a tweet is styled on Twitter, the dynamics of the platform were missing from the experiment interface. Being able to see multiple posts in a personal feed, interact with others and explore trending topics; all are important aspects that complete the full social media experience. Second, since every main tweet was accompanied with a questionnaire, participants likely disassociated from their "normal" social media behavior, and inherently became more critical as the questionnaire forced them to. Since this normal behavior of participants was unknown, no comparison was possible between one's behavior within and outside of the experiment setting.

Questionnaires Making the questionnaires identical between main tweets allowed for a comparison of participant responses in different conditions. However, this approach is less common in framing effect literature, as testing for issue-specific framing effects is not possible. Moreover, the participants were divided into three groups of seven, such that every tweet was seen in one of the conditions. The offset, i.e. the sequence of main tweets participants were exposed to, may have influenced their responses.

Twitter dataset The selection of the included eight politicians was based on the number of votes in the '21 elections. However, the political portfolio's of these politicians, the extent to which these politicians are active on Twitter and the amount of followers they have, were not considered in the selection process. This seemed the most straightforward way to construct the dataset, yet, it proved that the dataset featured some politicians that were hardly active, or had significantly less followers than some politicians that were excluded from the dataset. Therefore, prominent and influential politicians in the social media landscape were potentially missed. More in particular, inspection of Twitter at times of the experiment revealed that the official Ministry accounts that some politicians maintain besides their personal Twitter accounts, were showing more activity than their personal accounts. These types of official and edited Twitter accounts therefore might have been worthwhile to include rather than exclude, which choice was argued in chapter 4.

Realization of the visualizations The most important limitation in the information visualization has been the unreliable results of pattern-nl and its sentiment analysis. The most important reason to choose pattern-nl over a machine learning model was to easily retrieve what biased words the subjectivity and sentiment scores are based on. Although this seemed to be appreciated by some participants, a larger group of participants either disagreed with the scores, or got confused. This may have worked out for the better; some participants became more critical when they started to see that the visualized scores were not consistent with what they perceived the tweet to be. Another limitation comes from the political spectrum, which was criticized by some participants. Since no assumptions were made about the participants' political knowledge, the spectrum seemed the most comprehensible option to include. However, populist parties do not get classified easily according to this metric, and hence, coordinates can be misleading instead of assisting. Last, the choice to facilitate the dashboard with Microsoft PowerBI has constrained the technical possibilities of condition 3. A full linkage between main tweet and dashboard data, e.g. in D3.js, such that participants can click-through instead of interacting in two different interfaces, would have been a valuable addition.

Statistical testing procedure Although the current study has three observations per participant per condition, as every participant was exposed to three tweets in each condition, the statistical tests chosen for the analysis of user responses (i.e., Friedman and Mann-Whitney Wilcoxon) only take one value per condition. Therefore, the numerically converted participant responses were summed. A more appropriate analysis would therefore have been a multi-factor Friedman's test, however, no literature was found on how to mathematically perform this multi-factor test.

8.3 Future work

Future studies would benefit from an even more contextualized approach to the social media domain, possibly in a controlled setting where participants would take the experiment as if they would be seeing their own, personal feed. Above all, more participants are needed for a thorough analysis and statistically significant evidence on the influence that visualizations may or may not have on political framing effects. A larger sample would allow to make groups of participants that would be exposed to different visual encodings. This set-up would provide with more in-depth feedback on what visualization may benefit what purposes. Although testing for demographic factors was left out of scope for the current study, some quotes by participants do indicate strong political preferences that have influenced them. In addition, the current sample was largely inactive on Twitter, and did not engage with political content out of principle. Therefore, a final suggestion for future researchers in this domain is to account for social media behavior and preferences when defining sample characteristics.

Bibliography

[1] S. Kruikemeier, G. Van Noort, R. Vliegenthart, and C. H. De Vreese, "The relationship between online campaigning and political involvement," *Online Information Review*, vol. 40, no. 5, pp. 673–694, 2016. 1

[2] K. Liebhart and P. Bernhardt, "Political storytelling on instagram: Key aspects of Alexander van der Bellen's successful 2016 presidential election campaign," *Media and Communication*, vol. 5, no. 4, pp. 15–25, 2017. 1

[3] P. Varis, "Trump Tweets the Truth: Metric Populism and Media Conspiracy," *Trabalhos em Linguística Aplicada*, vol. 59, no. 1, pp. 428–443, 2020. 1

[4] M. Hameleers, "Populist disinformation: Exploring intersections between online populism and disinformation in the us and the netherlands," *Politics and Governance*, vol. 8, no. 1, pp. 146–157, 2020. 1

[5] D. Wakabayashi, "What is Section 230? Legal Shield for Websites is Targeted by Lawmakers - The New York Times," may 2020. 1

[6] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, pp. 1146–1151, mar 2018. 1, 8

[7] A. D'Ulizia, M. C. Caschera, F. Ferri, and P. Grifoni, "Fake news detection: A survey of evaluation datasets," *PeerJ Computer Science*, vol. 7, pp. 1–34, jun 2021. 1

[8] K. Sharma, F. Qian, H. Jiang, N. Ruchansky, M. Zhang, and Y. Liu, "Combating fake news: A survey on identification and mitigation techniques," *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 3, 2019. 1, 9, 10

[9] A. Bondielli and F. Marcelloni, "A survey on fake news and rumour detection techniques," *Information Sciences*, vol. 497, pp. 38–55, sep 2019. 1, 4

[10] G. Pennycook, Z. Epstein, M. Mosleh, A. A. Arechar, D. Eckles, and D. G. Rand, "Shifting attention to accuracy can reduce misinformation online," *Nature*, vol. 592, 2021. 1, 3, 8, 9

[11] Oxford Advanced Learner's Dictionaries, "Bias, noun." 2

[12] K. Lazaridou, R. Krestel, and F. Naumann, "Identifying media bias by analyzing reported speech," *Proceedings - IEEE International Conference on Data Mining, ICDM*, vol. 2017-Novem, pp. 943–948, 2017. 2

[13] D. J. Sude, G. D. Pearson, and S. Knobloch-Westerwick, "Self-expression just a click away: Source interactivity impacts on confirmation bias and political attitudes," *Computers in Human Behavior*, vol. 114, p. 106571, jan 2021. 2

[14] N. Guimaraes, A. Figueira, and L. Torgo, "Profiling Accounts Political Bias on Twitter," *Iberian Conference on Information Systems and Technologies, CISTI*, jun 2021. 2

[15] J. Bai, Q. Kong, L. Li, L. Wang, and D. Zeng, "Exploring cognitive dissonance on social media," *2019 IEEE International Conference on Intelligence and Security Informatics, ISI 2019*, pp. 143–145, jul 2019. 2

[16] M. Jeong, H. Zo, C. H. Lee, and Y. Ceran, "Feeling displeasure from online social media postings: A study using cognitive dissonance theory," *Computers in Human Behavior*, vol. 97, pp. 231–240, aug 2019. 2

[17] C. Sindermann, A. Cooper, and C. Montag, "A short review on susceptibility to falling for fake political news," *Current Opinion in Psychology*, vol. 36, pp. 44–48, dec 2020. 2, 8

[18] G. Pennycook and D. G. Rand, "Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning," *Cognition*, vol. 188, pp. 39–50, jul 2019. 2, 8

[19] C. Martel, G. Pennycook, and D. G. Rand, "Reliance on emotion promotes belief in fake news," *Cognitive Research: Principles and Implications*, vol. 5, no. 47, 2020. 2

[20] D. P. Calvillo, R. J. Garcia, K. Bertrand, and T. A. Mayers, "Personality factors and self-reported political news consumption predict susceptibility to political fake news," *Personality and Individual Differences*, vol. 174, may 2021. 2, 9

[21] E. P. Baumer, J. Snyder, and G. K. Gay, "Interpretive impacts of text visualization: Mitigating political framing effects," *ACM Transactions on Computer-Human Interaction*, vol. 25, aug 2018. 2, 5, 10

[22] E. Amsalem and A. Zoizner, "Real, but Limited: A Meta-Analytic Assessment of Framing Effects in the Political Domain," *British Journal of Political Science*, vol. 52, pp. 221–237, jan 2022. 2, 6, 51

[23] M. N. Lukacovic, "Framing Events of Political Violence within Peace Journalism and Internet User-Generated Comments: Effects on Readers' Emotions and Contemplation of Broader Implications," *Partecipazione e Conflitto*, vol. 13, no. 2, pp. 1022–1041, 2020. 3

[24] J. Kaiser, *Emphasis Framing Effects in Political Communication-An Experimental Approach to the Role of Issue-Specific Argument Strength, Cross-Thematic Salience Emphasis Frames, and Political Value Preferences in Citizens' Attitude Formation.* Zurich: Zurch: University of Zurich, Faculty of Arts, 2020. 3, 44

[25] R. Berman and Z. Katona, "Curation Algorithms and Filter Bubbles in Social Networks," *https://doi.org/10.1287/mksc.2019.1208*, vol. 39, pp. 296–316, mar 2020. 3

[26] P. Liu, K. Shivaram, A. Culotta, M. A. Shapiro, and M. Bilgic, "The interaction between political typology and filter bubbles in news recommendation algorithms," *The Web Conference 2021 - Proceedings of the World Wide Web Conference, WWW 2021*, vol. 11, pp. 3791–3801, apr 2021. 3

[27] G. M. Lunardi, G. M. Machado, V. Maran, and J. P. M. de Oliveira, "A metric for Filter Bubble measurement in recommender algorithms considering the news domain," *Applied Soft Computing*, vol. 97, p. 106771, dec 2020. 3

[28] M. Cinelli, G. de Francisci Morales, A. Galeazzi, W. Quattrociocchi, and M. Starnini, "The echo chamber effect on social media," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, mar 2021. 3

[29] M. Boler and E. Davis, "The affective politics of the "post-truth" era: Feeling rules and networked subjects," *Emotion, Space and Society*, vol. 27, no. December 2017, pp. 75–85, 2018. 3, 5

[30] S. C. Woolley and D. R. Guilbeault, "Computational Propaganda in the United States of America: Manufacturing Consensus Online," in *Computational Propaganda in the United States of America: Manufacturing Consensus Online*, 2017. 3

[31] Detrow Scott, "What Did Cambridge Analytica Do During The 2016 Election? : NPR," mar 2018. 3

[32] X. Teng, Y.-R. Lin, W.-T. Chung, A. Li, and A. Kovashka, "Characterizing User Susceptibility to COVID-19 Misinformation on Twitter," tech. rep., University of Pittsburgh, sep 2021. 4

[33] U. K. H. Ecker, S. Lewandowsky, J. Cook, P. Schmid, L. K. Fazio, N. Brashier, P. Kendeou, E. K. Vraga, and M. A. Amazeen, "The psychological drivers of misinformation belief and its resistance to correction," *Nature Reviews*, vol. 1, 2022. 4

[34] Twitter, "Permanent suspension of @realDonaldTrump," jan 2021. 4, 18

[35] S. A. Thompson and D. Alba, "In Ukraine's Information War, a Blend of Fact and Fiction - The New York Times," mar 2022. 4

[36] E. Culliford, "Facebook to label all posts about COVID-19 vaccines — Reuters," mar 2021. 4

[37] K. Paul, "Twitter targets Covid vaccine misinformation with labels and 'strike' system — Twitter — The Guardian," mar 2021. 4

[38] V. Irwin, "Misinformation effect: Spotify adds COVID labels on podcasts - Protocol," feb 2022. https://www.protocol.com/bulletins/spotify-covid-warnings-podcasts. 4

[39] X. Zhang and A. A. Ghorbani, "An overview of online fake news: Characterization, detection, and discussion," *Information Processing and Management*, vol. 57, mar 2020. 4

[40] P. Bharadwaj and Z. Shao, "Fake News Detection with Semantic Features and Text Mining," *International Journal on Natural Language Computing*, vol. 8, pp. 17–22, jun 2019. 4

[41] G. Gadek and P. Guélorget, "An interpretable model to measure fakeness and emotion in news," *Procedia Computer Science*, vol. 176, pp. 78–87, jan 2020. 4

[42] K. Clayton, S. Blair, J. A. Busam, S. Forstner, J. Glance, G. Green, A. Kawata, A. Kovvuri, J. Martin, E. Morgan, M. Sandhu, R. Sang, R. Scholz-Bright, A. T. Welch, A. G. Wolff, A. Zhou, and B. Nyhan, "Real Solutions for Fake News? Measuring the Effectiveness of General Warnings and Fact-Check Tags in Reducing Belief in False Stories on Social Media," *Political Behavior*, vol. 42, pp. 1073–1095, dec 2020. 4

[43] N. Pezzotti, J.-D. Fekete, T. Höllt, B. Lelieveldt, E. Eisemann, A. Vilanova, and B. P. F. Lelieveldt, "Multiscale Visualization and Exploration of Large Bipartite Graphs," *Computer Graphics Forum*, vol. 37, no. 3, p. 12, 2018. 4

[44] T. Spinde, F. Hamborg, K. Donnay, A. Becerra, and B. Gipp, "Enabling News Consumers to View and Understand Biased News Coverage: A Study on the Perception and Visualization of Media Bias," *ACM/IEEE-CS Joint Conference on Digital Libraries*, 2020. 5, 10, 11

[45] B. Plüss and A. de Liddo, "Democratic replay: Enhancing TV election debates with interactive visualisations," *Proceedings of the Annual Hawaii International Conference on System Sciences*, vol. 2018-Janua, pp. 1711–1720, 2018. 5

[46] L. South, M. Schwab, N. Beauchamp, L. Wang, J. Wihbey, and M. A. Borkin, "DebateVis: Visualizing Political Debates for Non-Expert Users," *Proceedings - 2020 IEEE Visualization Conference, VIS 2020*, pp. 241–245, oct 2020. 5

[47] C. Gosse and J. Burkell, "Critical Studies in Media Communication ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rcsm20 Politics and porn: how news media characterizes problems presented by deepfakes," *Critical Studies in Media Communication*, 2020. 5

[48] S. van der Linden, "Misinformation: susceptibility, spread, and interventions to immunize the public," *Nature Medicine*, vol. 28, pp. 460–467, mar 2022. 7, 8

[49] K. Bryanov and V. Vziatysheva, "Determinants of individuals' belief in fake news: A scoping review determinants of belief in fake news," *PLoS ONE*, vol. 16, jun 2021. 7, 8, 9

[50] L. K. Fazio, "Knowledge does not protect against illusory truth," *Journal of Experimental Psychology: General*, vol. 144, no. 5, pp. 993–1002, 2015. 8

[51] B. E. Weeks, "Emotions, Partisanship, and Misperceptions: How Anger and Anxiety Moderate the Effect of Partisan Bias on Susceptibility to Political Misinformation," *Journal of Communication*, vol. 65, pp. 699–719, 06 2015. 9

[52] F. J. Z. Borgesius, D. Trilling, J. Möller, C. H. De Vreese, and N. Helberger, "INTERNET POLICY REVIEW Journal on internet regulation Should we worry about filter bubbles?," *Internet Policy Review*, 2016. 9

[53] J. Oosterman and N. Tintarev, "Operationalizing Framing to Support Multiperspective Recommendations of Opinion Pieces; Operationalizing Framing to Support Multiperspective Recommendations of Opinion Pieces," *ACM Reference Format: Mats Mulder*, 2018. 9

[54] N. Tintarev, E. T. Sullivan Delft EESullivan-Mumm, T. Dror Guldin, S. Qiu, D. Odjik Blendle Research, E. Sullivan, D. Guldin, and D. Odjik, "Algorithmic Diversification of Viewpoints in News," *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*, 2018. 9

[55] F. Loecherbach, K. Welbers, J. Moeller, D. Trilling, and W. Van Atteveldt, "Is this a click towards diversity? Explaining when and why news users make diverse choices," *ACM International Conference Proceeding Series*, pp. 282–290, jun 2021. 9, 34

[56] C. S. Traberg and S. van der Linden, "Birds of a feather are persuaded together: Perceived source credibility mediates the effect of political bias on misinformation susceptibility," *Personality and Individual Differences*, vol. 185, p. 111269, feb 2022. 9

[57] T. Spinde, C. Jeggle, M. Haupt, W. Gaissmaier, and H. Giese, "How do we raise media bias awareness effectively? Effects of visualizations to communicate bias," *PLOS ONE*, vol. 17, p. e0266204, apr 2022. 9, 10, 14

[58] U. K. Ecker, S. Lewandowsky, and D. T. Tang, "Explicit warnings reduce but do not eliminate the continued influence of misinformation.," *Memory and Cognition*, vol. 38, pp. 1087–1100, dec 2010. 10

[59] A. Patankar, J. Bose, and H. Khanna, "A Bias Aware News Recommendation System," *Proceedings - 13th IEEE International Conference on Semantic Computing, ICSC 2019*, pp. 232–238, 2019. 10, 11

[60] F. F. Cheng and C. S. Wu, "Debiasing the framing effect: The effect of warning and involvement," *Decision Support Systems*, vol. 49, pp. 328–334, jun 2010. 10

[61] S. Lecheler, C. De Vreese, and R. Slothuus, "Issue Importance as a Moderator of Framing Effects:," *http://dx.doi.org/10.1177/0093650209333028*, vol. 36, pp. 400–425, apr 2009. 10, 44

[62] J. Kaiser, "Disentangling the Effects of Thematic Information and Emphasis Frames and the Suppression of Issue-Specific Argument Effects through Value-Resonant Framing," *https://doi.org/10.1080/10584609.2019.1658662*, vol. 37, pp. 1–19, jan 2019. 10

[63] P. Borah, "Competitive framing, emotion and heterogeneity: framing effects on Twitter in the case of an alt-right rally," *Online Information Review*, vol. 45, no. 6, 2021. 10

[64] M. Dörk, C. Collins, P. Feng, and S. Carpendale, "Critical InfoVis: Exploring the Politics of Visualization," *Extended Abstracts*, 2013. 14, 22

[65] Tweepy, "Tweepy." https://www.tweepy.org/. 18

[66] N. Project, "NLTK :: Natural Language Toolkit." https://www.nltk.org/. 21

[67] "Lemmatisation - Wikipedia." 21

[68] P. Delobelle, T. Winters, and B. Berendt, "RobBERT: a Dutch RoBERTa-based Language Model," *Findings of the Association for Computational Linguistics Findings of ACL: EMNLP 2020*, pp. 3255–3265, jan 2020. 21

[69] E. Kokalj, J. S. International, B. Š. Jožef, S. International, N. L. Jožef, S. P. Jožef, and M. Robnik-Šikonja, "Proceedings of the EACL Hackashop on News Media Content Analysis and Automated Report Generation, pages 16-21 BERT meets Shapley: Extending SHAP Explanations to Transformer-based Classifiers," in *Proceedings of the EACL Hackashop on News Media Content Analysis and Automated Report Generation, pages 16-21 BERT meets Shapley: Extending SHAP Explanations to Transformer-based Classifiers*, 2021. 22

[70] L. Gatti, J. Van Stegeren, and J. E. V. Nl, "Improving Dutch sentiment analysis in Pattern," *Computational Linguistics in the Netherlands Journal*, vol. 10, pp. 73–89, 2020. 22

[71] "Kieskompas Tweede Kamerverkiezingen 2021 - Kieskompas." https://www.kieskompas.nl/nl/tools/tweedekamer2021/. 26

[72] R. Mehrotra, S. Sanner, W. Buntine, and L. Xie, "Improving LDA Topic Models for Microblogs via Tweet Pooling and Automatic Labeling," tech. rep., NICTA, 2013. 26

[73] T. Munzner and E. G. artist) Maguire, *Visualization analysis and Design*, vol. 16. CRC Press, 2014. 30, 31, 33

[74] K. Kucher, C. Paradis, and A. Kerren, "The State of the Art in Sentiment Visualization," *COMPUTER GRAPHICS forum*, vol. 37, no. 1, pp. 71–96, 2018. 36

[75] C. Schemer, W. Wirth, and J. Matthes, "Value Resonance and Value Framing Effects on Voting Intentions in Direct-Democratic Campaigns:," *American Behavioral Scientist*, vol. 56, pp. 334–352, nov 2011. 44

[76] S. Clifford, "How Emotional Frames Moralize and Polarize Political Attitudes," *Political Psychology*, vol. 40, pp. 75–91, feb 2019. 44

[77] R. H. Riffenburgh, "Chapter Summaries," *Statistics in Medicine*, pp. 533–580, 2006. 51

[78] Wikipedia, "Friedman test." 51

[79] S. Glen, "W Statistic (Coefficient of Concordance) ." https://www.statisticshowto.com/w-statistic/. 51

Appendix A

Wordlist-based algorithm

Wordlists Exploratory topic modelling was performed on the cleaned text of tweets in order to ease the manual tweet selection. The wordlists for the first three topics, as shown in table A.1, were constructed by combining two types of words, i.e. 1) outputted words by the LDA modelling algorithm, and 2) topical words found online. The fourth topic, i.e. the nitrogen crisis in the Netherlands, was added at later stage, hence the wordlists for this topic only contain words of the second type. An overview of these wordlists is given in A.1.

Algorithm In order to label the tweets with a topic, in case present, a wordcount matrix was constructed. This matrix contained one identifier column with the unique status id for each tweet, four columns for each of the main topics, and two columns for all corresponding subtopics, yielding 1 + 4 + (4x2) = 13 columns in total. For every clean text of a tweet, the number of words that occurred in the main topic and subtopic wordlists were counted and added to the matrix. In other words, a high wordcount of a tweet for main topic X and subtopic Y, meant that the tweet was likely to contain arguments about those topics. With all the wordcounts collected in the matrix, an algorithm checked whether there was a unique main topic by inspecting whether there was a unique maximum in the set of four main topic wordcounts[1]. If this was the case, the tweet's main topic label was set to this main topic, and the algorithm then checked whether there was a unique maximum present among the corresponding subtopics for that tweet's main topic. If this was the case, the tweet's subtopic label was set to this subtopic. In case there were no main topics or subtopics found, the tweet received empty labels for main topic and subtopic[2]. The described process is depicted in figure A.1. The distribution of topics and their subtopics is displayed in figure A.2.

[1]The subtopic wordcounts were added to the main topic wordcounts when checking for a main topic maximum.
[2]For some tweets, a main topic was found, however, no subtopic was found. In these cases, the main tweet's topic was labelled, yet the subtopic of the tweet remained empty.

Main topic words	Subtopic words	
War in Ukraine	Violence	Consequences
rusland, oekraine, ukraine, russia, kiev, marioepol, russisch, oekraiens, poetin, zelenski, oorlog, nato, navo, support, steun, security, cooperation, situation, europe, europa	invasie, leger, militair, gevecht, vechten, aanval, strijd, belagen, bomb, gewond, ziekenhuis	vluchteling, burger, wereld, sanctie, inflatie, gas
General policy criticism	Social inequalities	Sustainability
rutte, kaag, kabinet, kabinet , binnenhof, den haag, samenleving, nederland, minister, tweede kamer, europa, democratie, partij, euro, unie, veiligheid, ons land, onze mensen, verkiezing, kamerlid, kamervraag, motie, miljoen, beleid, belang, gesprek, wet, onwettig, besluit, debat, kamervragen	bezuiniging, bezuinigen, gemeente, onze zorg, belasting, belasting , buurt, aow, inflatie, inflatie , inflatie, pensioen, actie, werkloosheid, toeslagen, toeslagenschandaal, toeslagenaffaire, generatie, wegcijferen, ouder, rijk, vermogen, arm, bezuiniging, werken, koopkracht, koopkracht , huurwoning, eerlijk, wonen, minimumloon, betaalbaar, petitie, onrecht, progressief, plan, sociaal, voorstel	energie, energierekening, stikstof, klimaat, groen, groningen, groninger, welkomgasberaad, sluiting, elektriciteit
Coronavirus	Healthcare consequences	Lawmaking
corona, coronavirus, delta, omicron, omikron, variant, coronagolf, virus, covid, rivm, nederlanders, structureel, rutte, kamervraag, kamervragen, samenleving, mens, mensen, hugodejonge, vaccin, vaccinatie	zorg, samen, ic, artsen, verpleegkundig, investeren, ziekenhuis, zorgmedewerker, verpleeghuis, volwassen, besmetten	lockdown, coronatoegangsbewij, pfizer, booster, vrijheidsbeperking, vrijheid, ingeperkt, coronapas, kuiper, pandemie
Nitrogen crisis	Governance	Innovation
stikstof, stikstof, stikstof-, stikstofcrisis, stikstof crisis, agrarische sector, boeren, raad van staate, raad van state, voedselveiligheid, voedsel veiligheid, grondkwaliteit, stikstofverbindingen, ecosysteem, veehouderij, veestapel, traktor, landbouw, landbouw en klimaat, klimaat	wanbeleid, stikstofmaatregelen, noodzakelijk, debat, oppositie, coalitie, van der wal, staghouwer, protest, boerenprotest, malieveld, stikstofrechten, boerenprotest, onteigenen, compens, woede, stikstofplannen, tweede kamer, elite	tovervloer, innovatie, samenwerken, stikstofplannen

Table A.1: Wordlists specified for the three topics that were found with LDA topic mining, as well as the fourth topic that was added at later stage

	War in Ukraine	Violence	Consequences	General policy criticism	Social inequalities	Sustainability	Coronavirus	Healthcare consequences	Lawmaking	Nitrogen crisis	Governance	Innovation
Topic count key	t1	t11	t12	t2	t21	t22	t3	t31	t32	t4	t41	t42
1510925110894600192	5	0	2	2	0	0	1	1	0	0	0	0

Figure A.1: The topic labelling process. A row represents a tweet and a column represents a wordlist. Every cell contains a count of how many of the tweet's words were also featured in the corresponding wordlist. The main topic of the example tweet is *War in Ukraine*, and its subtopic is *Consequences*.

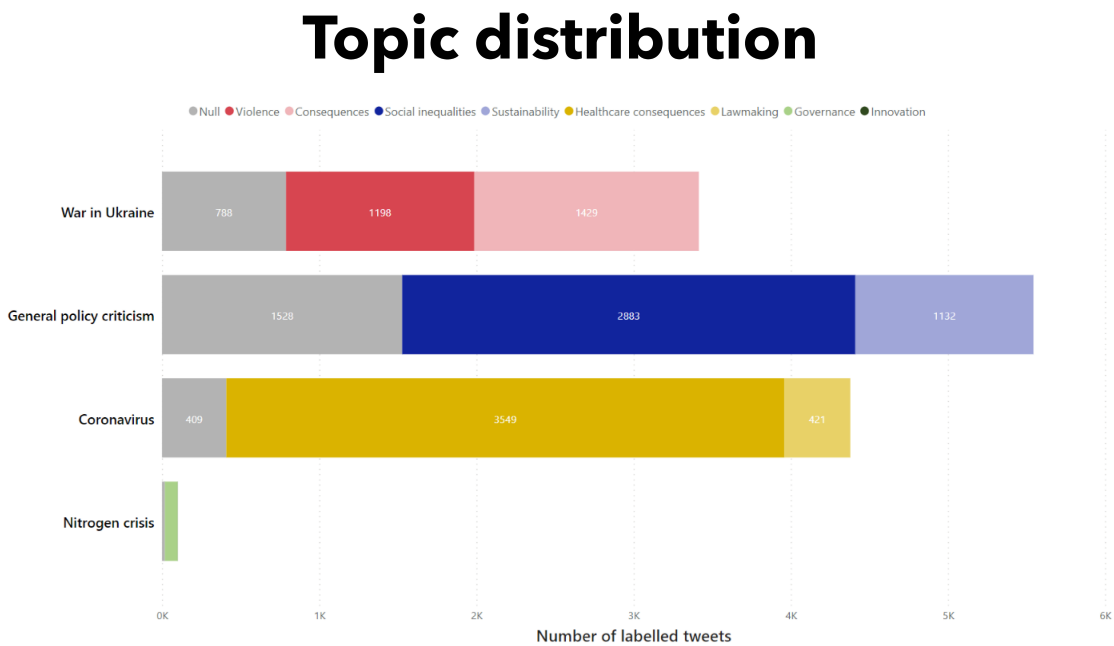


Figure A.2: The topic distribution of the current dataset. The majority of tweets (50152 tweets in total) was, however, left without a main topic label. This figure only displays the tweets for which the main topic was found.

Appendix B

Main tweets

The experiment followed a 3x3 design, as was described in chapter 6.2 and visually clarified in figure 6.3. The main tweets were hence grouped in three clusters, each containing three main tweets. The main tweets that belonged to the same cluster, were exposed under the same experiment condition - which condition that was, depended on the offset that the participant was assigned - i.e., the sequence of main tweets. Figures B.1, B.2, B.3 depict experiment screenshots of the main tweets in their corresponding cluster. The numbering of main tweets was kept consisted with figure 6.3.

Main tweet cluster 1

1

Political party:	VVD
Seated:	Coalition
Positioned:	Right-conservative
Tweet subject:	Coalition agreement

2

Political party:	PVV
Seated:	Opposition
Positioned:	Left-conservative
Tweet subject:	Russian war sanctions

3

Political party:	PVDA
Seated:	Opposition
Positioned:	Left-progressive
Tweet subject:	Household energy bills

Figure B.1: The main tweets contained in the first cluster.

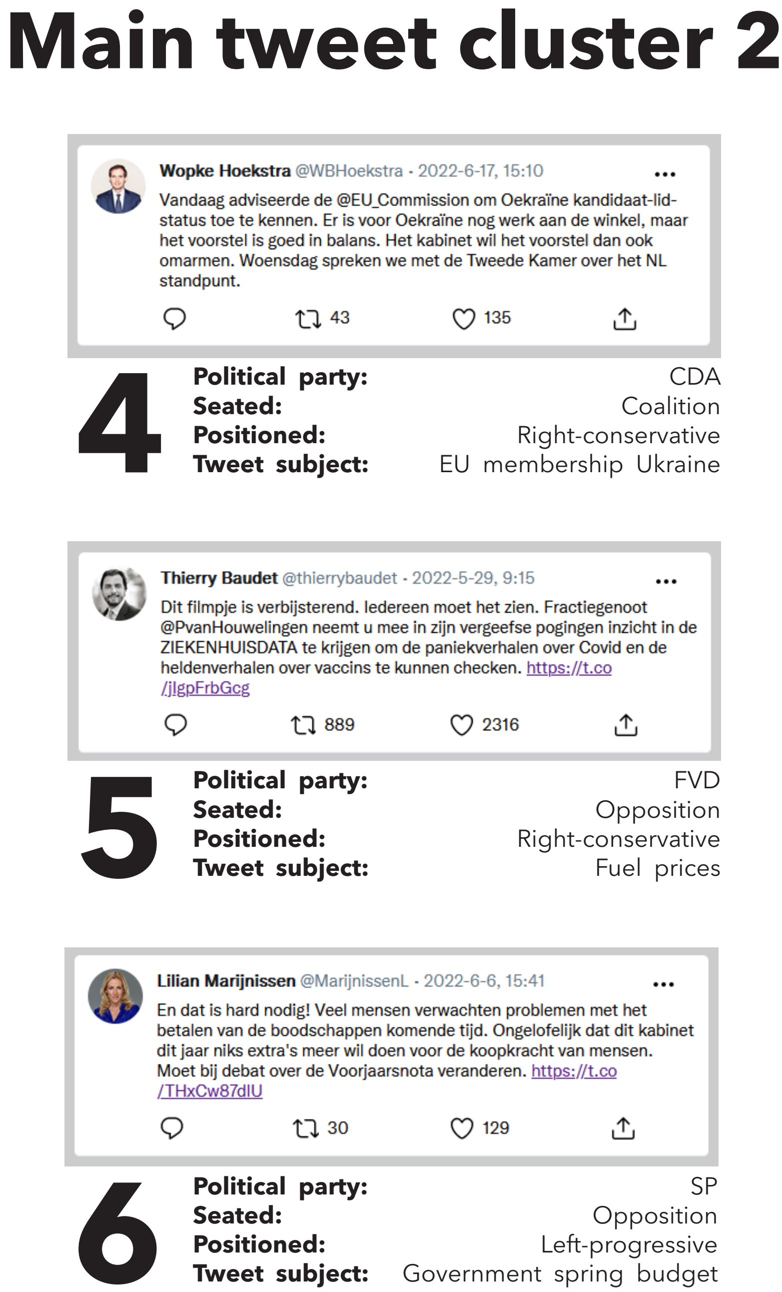


Figure B.2: The main tweets contained in the second cluster.

Figure B.3: The main tweets contained in the third cluster.

Appendix C

Experiment instruction

This appendix features a translation of the one-page intro that participants had to read before starting the experiment.

Translation

This experiment concerns a study to the interpretation to political tweets. The experiment consists of two parts. During the first part, you will answer questions after having seen one tweet. During the second part, you will answer questions about your overall experience in this experiment. At the end of the experiment, you will be asked to answer some final questions about yourself. These are control variables. These variables will be included in the analysis, yet they will be anonymous - this means that your identity can not be inferred from the answers you give.

In the first part, you will see nine main tweets. The main tweet will be highlighted on top of the page at all times. In some cases, you will see other tweets next to the main tweet, and/or a dashboard you may interact with. The other tweets and the dashboard concern the same topic as the main tweet, yet, they are supportive. The questions in the first part will always concern the main tweet.

All tweets you will see were obtained via Twitter. The dataset contains tweets from 8 political parties, namely those which received the most votes during the Dutch general elections in 2021, and seven national news outlets. The timeframe is set between December 2021 and June 2022. In case your interpretation of any artifact is asked for, please answer according to how you feel about it at this very moment. The table on the next page provides with an overview of the politicians and news outlets that were included.

During the experiment, you are allowed to ask questions to the researcher.

Appendix D

Questionnaires

The experiment featured three questionnaires; 1) questions to be asked after every main tweet in the first part, referred to as *tweet questionnaire*, 2) a survey about one's overall experience in the second part, referrerd to as *survey*, and 3) control variables. In most cases, a 5-Point likert scale was used for participant answers. In case of *Agreement*, the answer scale is as follows; *Strongly disagree, Disagree, Neither agree nor disagree, Agree, Strongly agree.* In case of *Frequency*, the scale is as follows; *Never, Rarely, Sometimes, Often, Always.* The tweet questionnaire is depicted in table D.1, the survey in table D.2, and the control variables in table D.3.

Tweet question	Likert-scale
I know what the tweet is about.	Agreement
The tweet contains information and/or arguments I did not know before.	Agreement
The subject of the tweet is important to me.	Agreement
The tweet is subjective.	Agreement
I know which political party the author of this tweet belongs to.	Custom[1]
The political statement in the tweet matches best with...	Custom[2]
The tweet's subject can be viewed from multiple perspectives.	Agreement
The tweet's author approaches the subject of the tweet from multiple perspectives.	Agreement
I agree with the author of the tweet about this subject.	Agreement
The tweet got me thinking about this subject.	Agreement
My own perspective on the tweet's subject has been changed.	Agreement
I would give this tweet a retweet.	Agreement
I would give this tweet a like.	Agreement
In general, I would retweet/like tweets by the author of the main tweet.	Agreement

Table D.1: Tweet questionnaire

Survey question	Likert scale
The context indicators made me more aware of the possible political bias of the main tweet.	Agreement
The sentiment indicator made me more aware of the possible sentiment of the main tweet, despite of its occasional incorrectness.	Agreement
The context indicators were insightful.	Agreement
The context indicators changed my attitude towards the author of the main tweet.	Agreement
Because of the context indicators, I took on a more critical view towards the main tweet's argument.	Frequency
The added tweets coming from other politicians and news media changed my view on the main tweet's subject.	Frequency
The added tweets coming from different news media gave me more information and context, which I was in need of in order to form my own opinion.	Frequency
The dashboard was easy to use.	Agreement
The dashboard contained too much information.	Agreement
The information I was able to get from the dashboard was a valuable addition to the context indicators.	Frequency
The dashboard gave me new insights about the subject and context of the main tweet.	Frequency
The information I was able to find in the dashboard has changed my attitude towards the main tweet's author.	Frequency
The interaction with the dashboard helped me to put the main tweet in context.	Frequency
Because I was able to get more contextual information as a result of interacting with the dashboard, I became more critical towards the main tweet's standpoint.	Agreement
I would use a similar dashboard if it was embedded in my social media feed.	Agreement
I would like to see more different (political) content in my own feed.	Agreement

Table D.2: Survey

Control variable	Answer possibilities
Highest level of education	LO / MBO / HBO / WO / Other
Age category	<21 / between 21 and 24 / between 25 and 29 / between 30 and 49 / older than 50
Political preference	Left-progressive / Left-conservative / Right-progressive / Right-conservative / None of the above
Political party I identify with	VVD / D66 / CDA / PVV / PVDA / SP / GL / FVD / Other / I do not want to share
I have a strong preference for a specific politician.	Yes / No
I follow politicians on social media.	Yes / No
Rating of my own political knowledge	Very poor / Poor / Fair / Good / Excellent
I am up-to-date about today's topicalities.	Strongly disagree / Disagree / Neither agree nor disagree / Agree /Strongly agree
I keep myself informed about Dutch politics.	Never / Rarely / Sometimes / Often / Always
Most used news outlet	De Telegraaf / Algemeen Dagblad (AD) / De Volkskrant / NRC Handelsblad / Trouw / NOS / NU.nl / Sociale media / RTL Nieuws / BNR Nieuwsradio / Other / I purposefully skip reading the news
Most used news consumption mode	Social media / Mobile apps by news outlets on my smartphone, tablet, or laptop / Websites / A paid subscription (content available via smartphone, tablet, or laptop) / A paid paper subscription (content delivered by post) / Radio / Television
I am active on Twitter.	Not at all / Rarely / Regularly / Active / Very active
I have used Twitter before.	Yes / No

Table D.3: Control variables