Sentiment Dynamics in Search Engine Results: Unravelling the Influence on Polarisation

Kars Snijders - 2676583

University of Twente

Faculty of Behavioural Management and Social Sciences

Communication sciences

Supervisor: Shenja van der Graaf

Second Super Visor: Sikke Jansma

Word count: 15109

Abstract

This study explores search behaviour and sentiment's effect on the information-seeking and retrieval process on Google's Search Engine Ranking Position (SERP) concerning polarising topics. Previous studies' results indicate that personal- and search-engine biases appear during interaction with search engines. However, more is needed to know about the role of sentiment in the information-seeking and retrieval process in the SERP. Recognising a research gap in this area, the study aims to contribute to the academic discourse on search behaviour biases and search engine dynamics. This study is based on five propositions to get a better understanding of this influence:

- ***Proposition 1:*** *Demographic factors as age, sex, location, education level, employment, and income influence the sentiment polarity of search results.*
- ***Proposition 2:*** *The influence of political preference on the determination of sentiment polarity in search engine results may be significant.*
- ***Proposition 3:*** *Short-tail queries might lead to a broader range of results, potentially including more general, less polarised, and thus more positively toned content than the more specific long-tail queries.*
- ***Proposition 4:*** *Patterns in search behaviour can be found and linked to the nature of information seeking, particularly about polarising topics.*
- ***Proposition 5****: Polarising topics are likely to elicit stronger emotions and opinions, potentially skewing search results towards more negative sentiment polarity.*

Employing a mixed-method approach, a survey study and search engine data analysis were conducted. The survey, with 114 participants, explored the participants' general search behaviour and towards polarised topics. Subsequently, the search engine data analysis and sentiment analysis were executed on individual SERP output from Google to research the differences in search results in terms of sentiment. The data was based on 118 participants and collected through a browser extension, which participants had to install on their computers.

The survey demonstrated search behaviours and patterns during the information-seeking process on polarising topics. The study further elucidates how demographics and search query length significantly affect sentiment polarity. In addition, findings reveal a predominance of positive sentiment in SERPs for polarised topics. These findings provide new insights into the role of sentiment in algorithmic decision-making. The study underscores the importance of advancing interdisciplinary research approaches to deepen the understanding of search engines and their influence.

Keywords: Information seeking and retrieval, Search Engine Interaction, SERP, Sentiment, Polarisation

Table of Contents

1. Introduction

In modern digital society, search engines function as critical tools for daily information retrieval across a diverse array of subjects. The influence of search engines on people's attitudes has been demonstrated in multiple cases. For example, during the COVID-19 pandemic, a social panic amongst the public was caused by the spread of false information and division in society. Governments and authorities urged people to verify the truthfulness of news reports before sharing them (Apuke & Omar, 2021). Another example was during the general elections in India, where biased search rankings shifted the voting preferences of undecided voters by 20% or more (Epstein & Robertson, 2015). Furthermore, there is evidence that Russian agents bought advertisements to spread fake news on Google and Facebook to interfere with the 2016 presidential campaign in the United States (Wakabayashi, 2017). These examples underscore search engines' capacity to shape societal perspectives and possibly divide people. A search engine does not only interfere with technical matters but political matters as well these days. Therefore, understanding search engine algorithms and how information is provided to the public is essential.

This pivotal role of search engines in society is predominantly occupied by Google, with a 91.9% market share emphasising exceptional influence in the digital information landscape (Statcounter, n.d.). Google's search engine operates through three processes: crawling to discover content, indexing for organisation, and ultimately serving search results to users (Google Search Central, n.d.). This process, key to this study, underscores the influence of algorithms on what information is prioritised and how it is presented on the search engine results page (SERP) (Golebiewski & Boyd, 2019). For example, a search engine measures a webpage by importance based on the quantity and quality of links referring to that webpage. As a result, the algorithm gives more weight to information from popular websites supporting the majority's interests and values. Therefore, it is more difficult to find less popular and smaller websites (Introna & Nissenbaum, 2000; Rieder & Sire, 2014). Google's algorithms are continuously adjusted, and new ranking parameters have been added over the last few years. Besides popularity, ranking signals such as demographics, location, clicks, language, sentiment, and wording influence the Search Engine Results Page (SERP) (Kliman-Silver et al., 2015). The releases of Google's algorithms are publicly available, but the operational mechanisms and collaborative functionalities of these algorithms remain unclear (Haider & Sundin, 2019). In this digital era, in which Google's algorithmic decisions hold sway over the accessibility and prominence of online information, understanding the dynamics of its search engine becomes crucial.

The information-seeking and retrieval process is influenced by cognitive biases and the inherent biases of search engines. While users generally perceive search engine results as objective, research indicates otherwise. According to Law et al. (2006), search engine biases manipulate users' perceptions and associations with the content they search for. The architects of search engines have created such an effective algorithm that results beyond the first result page are often not even examined (Golebiewski & Boyd, 2019). The interaction between search engines and human behaviour introduces a complex dynamic where multiple cognitive biases can influence online searching (Azzopardi, 2021). As most biases operate subconsciously, it is hard for people to recognise or prevent them (Gilovich et al., 1993). This research will delve deeply into the link between sentiment and human cognitive biases, particularly how they interplay during information-seeking and retrieval processes.

Central to this research is the observation of escalating hostility and mistrust among groups that has been observed in various Western countries, such as the United States (Iyengar et al., 2019), Europe (Casal Bértoa & Rama, 2021), including the Netherlands (Albada et al., 2021; Harteveld, 2021). Recent evidence suggests an emergence of populist parties, ideas, and policies (Abramowitz & Saunders, 2008). This phenomenon, often encapsulated by the term 'polarisation,' reflects deepening divides across various issues, from political allegiance to climate change stances. Polarisation, while variably defined, generally denotes the process of separation into groups with conflicting beliefs, values, and/or behaviours (Esau et al., 2023). During the last decade, polarisation has raised about various societal topics. The public opinion is sharply divided on climate change (Treen et al., 2020), immigration (Albada et al., 2021), and vaccine acceptance (Mønsted & Lehmann, 2022). The Dutch government's initiation of an anti-polarisation campaign underscores the growing recognition of the societal impact of polarisation (Sire – Stichting Ideele Reclame, n.d.). Human cognitive biases significantly contribute to this trend, as individuals gravitate towards information that reinforces their pre-existing views, intensifying polarisation (Del Vicario et al., 2017). Polarisation, often associated with negative sentiment, forms the angle of this study. Despite widespread recognition of the impact of polarisation on society, a comprehensive relation to search engines is still lacking. Given the influential role of search engines in shaping public discourse and perception, exploring this topic is essential. While the mechanics of search results have been extensively explored, the specific interplay of sentiment within the Search Engine Ranking Position (SERP) is still unknown. The dynamic between sentiment and search engines is particularly interesting when considering polarised topics, in which the sentiment could significantly influence information retrieval and presentation of search results. The research presents a unique and relevant perspective in our digitally driven world by examining how search engines, influenced by inherent human biases, might perpetuate or even worsen polarisation. For this study, the following research question will be examined:

'To what extent does Google's Search Engine Results Page (SERP) differ for users in terms of result of information presentation and sentiment concerning polarising and non-polarising topics in the Netherlands?'.

The findings of this study present a systematic foundation for future in-depth research into the dynamics of sentiment and its role in the presentation within Search Engine Results Pages (SERP). This research illustrates how sentiment influences online content, offering valuable insights for academic discourse and practical application.

The thesis structure is as follows: the theoretical framework lays out the foundational concepts and theories that guide the study, including exploring search engines, the interplay of human biases in search behaviour, and the angle of polarisation. The research design, data collection methods and analysis strategies are detailed in the methodology, providing a comprehensive understanding of the research process and the tools employed. The results present the findings of the study, revealing patterns and insights derived from the survey and sentiment analysis. Lastly, the discussion interprets the results, providing context and exploring their implications for understanding sentiment's role in shaping search engine algorithms and their outcomes. This final section emphasizes the necessity of continued exploration in this field to enhance our comprehension of search engines and the influence of sentiment.

2. Theoretical Framework

This theoretical framework explores the dynamic between search engine algorithms, human biases in digital information seeking and retrieval, and the role of sentiment.

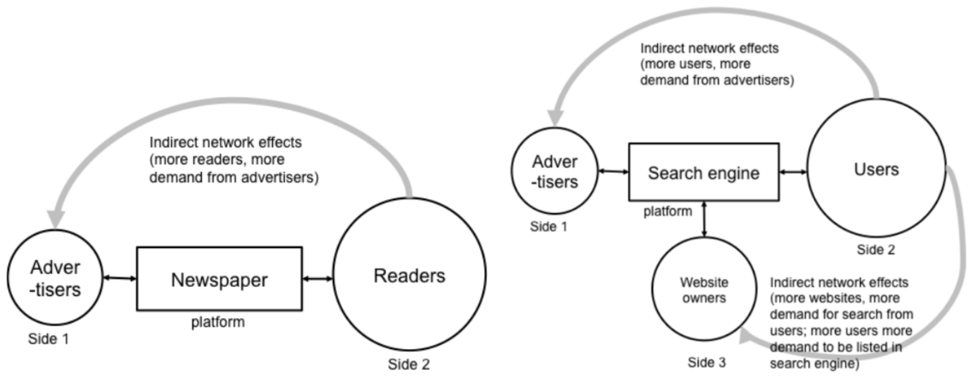
2.1. Search engines

A search engine is a software system designed to search and locate information on the internet. Because of its importance in society, a search engine can be defined as an information retrieval system in a public network, meaning that the vast majority has access to it (Van Hoboken, 2012). Search engines influence public discourse and knowledge distribution with their algorithms by curating and prioritising information based on user queries and prevailing news topics, much like traditional editorial media (Foster, 2012). Moreover, the concept of search engines as algorithmic media further enhances this understanding (Napoli, 2014). Algorithmic media refers to platforms where complex algorithms govern content curation and distribution of information instead of human editors. The intersection of technology and influence in public information distribution creates a complex dynamic that raises technical and political issues (Introna, 1999). Search engines make decisions that influence public opinion and discourse. How they manage political issues, misinformation, or controversial content can have far-reaching social- and political impacts. The ethical responsibility of search engines extends to ensuring transparency and accountability in their algorithmic choices.

Online information intermediaries such as Google and Facebook are gradually replacing traditional media, and thereby become media gatekeepers of our society (Helberger et al., 2015). Gatekeepers decide what and which information goes through the information gate to individuals or a group of people. They are high-level data decision-makers who can control the information flow of a social system. The development of search engines as gatekeepers has created a shift toward quantitative media instead of qualitative media (Van Couvering, 2017) and marked a crucial change in the media ecosystem. This shift can be understood by looking how SERP are being structured by search engines. Google's decision-making process in structuring the SERP directs users to various web pages, news, services, and products. The rise of new media platforms, such as Google, creates a shift from traditional two-sided models to multi-sided platforms. This phenomenon is also referred to as the mediatization of platforms, where mass media influence different sectors of society (Van Couvering, 2017). Figure 1 below displays the differences between a two-sided and multi-sided platform.

Figure 1

Two-sided platform versus a multi-sided platform (van Couvering, 2017, pp. 2-3)



Unlike traditional media platforms, for example newspapers, which directly connect readers and advertisers, these new multi-sided platforms rely on third-party content, separating content creation from distribution and creating a different indirect network effect. With a multi-sided media platform (e.g., social media platforms and search engines), the content is from a third party, such as a website. In this development, metadata about the platform and content is vital to the platform business. The metadata can be used for re-sale (i.e. to advertisers) as an additional product. Therefore, the stakes of new media are different than those of a traditional platform. With traditional platforms, it was just more users and more advertisements. However, the new dynamic now means more users, metadata, and content providers. This shift facilitates an environment where users engage with media in a highly filtered and personalised manner, primarily exposed to viewpoints and perspectives reinforcing their pre-existing beliefs. For platforms such as Google and Facebook, the metadata of users is a vital part of the business model (Van Couvering, 2017). The metadata is used to sell advertisements to advertisers, creating a competing dynamic between advertisements and (editorial) content on search engines. This new dynamic results in that (editorial) content directly competes with advertising on the search engine. This means that companies or organisations can buy a news spot and target a particular audience with the metadata.

This overall process is crucial to understanding how sentiment might influence information retrieval, as journalistic news values do not primarily drive the SERP's composition (Helberger et al., 2015) but rather personalized algorithms influenced by measures of popularity (Cho & Roy, 2004). This emphasis on popularity over journalistic merit could significantly reinforce biased information presentation, as it favours content that resonates more with user preferences and sentiments than with balanced, journalistic reporting. When popularity is rewarded instead of its actual content, a tension between quality and diversity arises as consumers prefer rankings based on popularity. This creates a limited set of search results that consist of frequently used websites (Helberger et al., 2015).

To attract more users to the platform, the search results must be attractive and match the user's perception of quality. However, when the organic results are high quality, it deducts the clicks on paid advertisements (Taylor, 2013). Like traditional media, a search engine business model thrives on income from advertisement (Taylor, 2013). However, in contrast to traditional media, this incentivises cannibalisation and degrades the quality of the search results. Companies are aware of this phenomenon and manipulate the search results to push their website to the top of the SERP. This can be done by paid advertisements or search engine optimisation (SEO). However, users often struggle to effectively differentiate between organic search results and advertisements (Fallows, 2005; Schultheiß & Lewandowski, 2021), which companies can take advantage of. This is substantiated by a study which disclosed that 73% of respondents from the United States demonstrated high confidence in the trustworthiness and accuracy of information obtained via search engines (Purcell et al., 2012). Therefore, it could be questionable if Google's position as a public intermediary and gatekeeper is ethical.

This concern becomes particularly pronounced considering instances where Google's influence has been manipulated for political purposes. For instance, the discovery that Russian agents exploited platforms like Google and Facebook to impact the 2016 U.S. presidential campaign underscores the potential vulnerabilities and ethical challenges (Wakabayashi, 2017). Another example is the spread of misleading advertisements concerning childhood vaccinations on Facebook and Twitter (Chiou & Tucker, 2018). These examples of interference through paid advertisements on multi-sided platform systems have profound implications for my study of the extent to which sentiment influences

Google's SERP. Such cases highlight the potential for these platforms, including search engines like Google, to facilitate the spread of targeted messages with minimal oversight. Paid advertisements can steer the audience towards specific narratives or viewpoints, thereby influencing the sentiment and biases in the SERP. The possibility of disseminating false or misleading information through advertisements and articles raises significant concerns about the role of Google in shaping public perception.

Besides Google shaping the public perception, search engine marketers use various ways to come up in the top search results. For example, the "long tail" concept, popularised by Anderson (2006), shifted trends from widespread search queries to more niche, specific search queries. Search engine marketing created a distinction between short-tail and long-tail keywords. It suggests that longer search queries are needed to drive success in the search results (Skiera et al., 2010). This study critically examines the distinction between short- and long-tail queries within the domains of search engines. Despite its widespread acceptance and application in these fields, more empirical research is needed to investigate the impact of this differentiation. By integrating this into the research it seeks to clarify whether the nature of the query, short-tail versus long-tail, yields significantly different results in sentiment. It potentially adds understanding and practices in digital marketing research.

In conclusion, the role of search engines like Google in public discourse creates a shift from qualitative to quantitative media. It highlights how algorithm-driven platforms prioritise popularity and user data, potentially leading to biased information spreading. The interplay between paid advertisements and search results, especially in politically charged contexts like the 2016 U.S. election, underscores the critical need for better understanding and transparency in these powerful digital gatekeepers.

2.1.1. Search Engine bias

People often assume that search engines (results) are neutral and exclude biases. The neutral design may give users the idea that the search engine displays identical information for everybody while receiving an individualised and biased SERP (Zuiderveen Borgesius et al., 2016). This leads to a skewed or filtered representation of information. For example, the search engine excludes websites in favour of other websites or gives a website more prominence (Introna & Nissenbaum, 2000; Norocel & Lewandowski, 2023). This argument finds some support in indications that Google has 'blacklisted' specific websites, a claim that Google has denied (Grind et al., 2019).

As previously outlined, Google’s business model heavily relies on advertisement revenue, which inherently influences the user's search results (Taylor, 2013). The search process is designed to optimise search results to increase the user's experience (UX) (Law et al., 2006). For example, when Google shows results that match a person's interests and previous searches, it might feel more relevant and valuable. This is done by the selection of sources, content, views, and page ranks (Bozdag, 2013). As the search engine bias increases, the UX for a person improves (Wijnhoven & van Haren, 2021). However, this approach to information retrieval can present certain drawbacks. Using individual user data, such as historical search queries, means that Google might show a limited range of information. This increases the probability that less diverse content, sources and viewpoints will be accessible, possibly resulting in a filter bubble (Flaxman et al., 2016; Pariser, 2011). As a result, a user may only see certain web content based on account- and behaviour information delivered by the search engine's algorithm. This is crucial in understanding the role of Google's SERP in potentially reinforcing sentiment-driven biases in presenting information.

Google's algorithms are tailored to individual interests and preferences. This results in mirroring human biases and stereotypes, raising concerns regarding its implications on social perceptions (Noble, 2018). An example of this issue is the algorithm’s tendency to reinforce racial stereotypes. Research has indicated that Google's search results for job-related images mainly featured white men, suggesting a bias in the representation of different racial groups (Lam et al., 2018). Similarly, there is an observed difference in the positioning of women compared to men on search results pages, with women often appearing in less prominent positions. Furthermore, searches related to specific racial or ethnic groups have yielded troubling results, emphasising a necessary oversight in the algorithms governing search engines. For instance, Noble (2018) alarming documents that a search for 'black girls' in 2018 predominantly returned pornographic and sexualised content. This source further illustrates the prevalence of anti-Semitic content in searches related to Jewish history or people, underlining a systemic issue within search engine results that perpetuates stereotypes and discrimination.

Additionally, the case of Dylann Roof, a white supremacist responsible for the Charleston church shooting, highlights the dangerous potential of these biases in radicalization processes. Roof's search for 'black-on-white crimes' led him to websites and groups promoting hate, which played a role in his radicalization journey (Hersher, 2017). These examples underscore the profound impact of search engine algorithms in shaping and sometimes reinforcing societal stereotypes and biases. Unlike traditional media companies, search engines do not merely redistribute third-party content; they actively curate and prioritise information through their algorithms, effectively performing editorial functions in the information retrieval process. This role places an onus on search engines to be aware of and address the potential for their algorithms to perpetuate and reinforce harmful stereotypes and biases (Noble, 2018). The consequences of such biases in search algorithms extend beyond individual UX, potentially influencing broader societal perceptions and eternalising stereotypes and misinformation.

In conclusion, due to optimization strategies employed by search engines, users may encounter difficulties accessing diverse content. This phenomenon can be attributed to the prioritisation in the content curation processes, which often favours certain types of information, thereby limiting the range of perspectives readily available to the end-users. These search engine biases prevent users from accessing sources, content, and viewpoints. This opposes the architecture and the ideals and values that have driven the development and growth of the internet (Barlow, 1996; Introna & Nissenbaum, 2000). By critically examining these biases, this research underscores the urgent need for a study of the principles guiding search engine algorithms, endorsing for a recalibration of these mechanisms to align with the internet's original democratic and inclusive values.

2.2. Human biases in search behaviour

As search engines' algorithms reflect human biases, it is essential to gain a deeper understanding of these biases and their influence on the search process. During the search process, various cognitive biases can significantly affect how information is searched, discovered, and retrieved (Azzopardi, 2021). A cognitive bias is a systematic pattern of inconsistent thinking that might lead to misjudgement and faulty decision-making (Tversky & Kahneman, 1974, 1992). These biases often deviate from the regular rational decision-making models and arise from the interplay between two distinct systems in the human brain, as described by Kahneman (2011). The first system operates automatically, quickly, and intuitively, relying on heuristics.

In contrast, the second system is a more conscious, slow, and analytical approach to the decision-making process. It requires more cognitive effort and is generally more reliable. During information seeking and retrieval, the first system often dominates, aiming to reduce the cognitive load by simplifying and expediting decision-making processes (Czerlinski, 1999). This fast, intuitive system dominance can lead to decisions influenced by immediate and emotionally charged information (Kahneman, 2011), which is particularly relevant in sentiment-driven search queries. The interplay between cognitive biases and search engine algorithms can lead to a reinforcing cycle, where biased search results cater to and further grow users' pre-existing beliefs and sentiments.

When people search, they spend more time engaging with search results that affirm their (pre-existing) beliefs while swiftly bypassing content that contradicts them (Knobloch-Westerwick et al., 2015). This can be explained by confirmation bias, an acclaimed cognitive bias that influences decision-making. Individuals subconsciously tend to seek information that aligns with their pre-existing beliefs (Nickerson, 1998). It influences how people search, interpret, remember, and challenge information. It can occur when people selectively focus on information that confirms their views while discounting or ignoring data that does not (Allahverdyan & Galstyan, 2014; Frost et al., 2015). Confirmation bias plays a big part in the spread of viral media caused by communities of interest (Vicario et al., 2016) and echo chambers (Cinelli et al., 2021). These communities cause reinforcements of their beliefs, possibly one of the causes of segregation and polarisation. In addition, algorithms can create personalised information ecosystems that support confirmation bias (Del Vicario et al., 2017), leading users to engage predominantly with content that aligns with their existing beliefs.

Confirmation bias is also linked with negativity bias, where messages that support their worldview are preferred to be critical (Knobloch-Westerwick et al., 2020). Negativity bias is the tendency to pay more attention to negative events or sentiments and overlook good events. Negativity in language use has been discussed recently as one of the causes of polarisation (Hills, 2019). Negativity bias can be taken in many ways. For example, people take criticism more seriously than praise (Peeters, 1971; Taylor, 1991). In addition, negative language is perceived as more intelligent, competent, and expert than positive language is (Amabile, 1983). Research shows that online negative product reviews have a more significant effect on sales than positive reviews (Cui et al., 2012). This effect might be explained by consumers' tendency to search for negative word of mouth when consumers lack experience and information (Herr et al., 1991). Most news is written with a negative sentiment to gain more attention and be perceived as more truthful than positive news (Hilbig, 2012; Soroka et al., 2019). News about polarised topics is often written in a negative context or sentiment. For example, during COVID-19, 78% of the news published in the U.S. and U.K. was written with strongly negative sentiments (Sacerdote et al., 2020). During the pandemic, the Oxford Daily Mail reported the positive development of a COVID-19 vaccine, highlighting prior experience in MERS vaccine research. In contrast, media outlets like Fox News, CNN, The New York Times, and The Washington Post covered this story more negatively, focusing on the cautious perspectives of health officials and emphasising the low probability of a vaccine within that year. In addition, most fake news articles containing information about the government and politics are written negatively to gain more attention (Tandoc et al., 2021). Furthermore, within echo chambers, negative language dominates the language in the comments of debunking posts (Zollo et al., 2017).

Users inclined toward their pre-existing beliefs are more likely to select search results that confirm them, often favouring negative over positive content. This selective engagement reinforces personal biases and informs and refines the algorithms behind SERP, affecting future rankings and content filtering (Bozdag, 2013). Particularly in discussions around polarised topics, the predominance of negative sentiment or language can intensify these biases. Such trends in user behaviour and algorithmic response contribute to how information is perceived and processed, skewed toward reinforcing existing beliefs and amplifying negative sentiments toward opposing viewpoints. This cyclical process of reinforcement and amplification of biases and sentiments through user interaction with search results highlights the possible critical role of search engine algorithms in shaping public discourse.

2.2.1. Human bas in the SERP

Various studies have demonstrated that the way search results are displayed significantly impacts a person's attitudes, behaviour, and preferences (e.g., Haas et al., 2017; Joachims et al., 2005; Novin & Meyers, 2017; Pan et al., 2007). Users judge the usefulness of sources depending on their presentation within a SERP (Novin & Meyers, 2017). This makes the order of the search engine results crucial. When a person is searching, four main biases appear on the SERP itself: priming, anchoring, framing, and the availability heuristic (Novin & Meyers, 2017).

Firstly, priming occurs when a person is already attracted to familiar categories or texts before reading any text in the search results, and repeated text or picture cues direct our eyes to familiar information (Kahneman et al., 1992).

Secondly, anchoring occurs when a person is biased towards the first value, they perceive in a set of data (Tversky & Kahneman, 1974). The first results in the SERP can affect the user's impression of the importance of the following results. As the user reads the results in hierarchical order, their assumptions are anchored by the top results in the SERP. In addition, search engines frequently show popular results to assist users in quickly solving their search queries (Harper, 2017; Rieder & Sire, 2014). This results in users making assumptions anchored on the first popular results and makes Google's algorithm biased towards mainstream and popular information.

Thirdly, the framing effect is a cognitive bias that occurs when peoples' choices are influenced by how information is presented (Kahneman, 2003). Frames are abstractions that structure or organise a message's meaning. Within political communication, the mass media sets the frame of reference that users use to interpret and discuss public events (Scheufele, 1996). With the rise of Google as an information intermediary and taking over the mass media, Google sets the frame of reference for users. Since the release of the Hummingbird update in 2013, Google has framed specific search results at the top using semantic search (Patil et al., 2021). This results in that there is a difference between search results in how they are displayed towards a user.

Fourthly, the availability heuristic bias explains the influences on a person's perception by the ease of access to information (Tversky & Kahneman, 1974). As a person searches, it scans the first results of the SERP with greater detail than later results. Most users do not look further than the first few search results (Joachims et al., 2005; Pan et al., 2007), so the variety of perspectives in a person's search is constricted (Novin & Meyers, 2017). Consequently, this often leads to skewed judgments because the most easily recalled or accessed information is not always the most accurate or comprehensive representation of a situation or topic. As most users trust search engines more than their judgement (Pan et al., 2007), this might cause problems. When dealing with a controversial or polarised topic, a non-expert might have difficulty determining which sources have high credibility in terms of depth, objectivity, authority, and purpose due to how search results are presented in the SERP. Therefore, it is essential to have a better understanding of search engine algorithms.

2.3. Polarisation

Within the domain of science, the notion of polarisation needs a universally accepted conceptualisation, illustrating the absence of consensus regarding its definition. This concept of polarisation exhibits notable variations across disciplines such as political and media and communication studies (Esau et al., 2023). However, the concept of polarisation is multifaceted, embracing various levels and forms. Polarisation can operate at macro, meso, and micro levels and is motivated by various aspects (i.e. ideologies, emotions, positions, identities, issues, values). In addition, polarisation has different actors, such as politicians, political parties, citizens, journalists, and media outlets. However, it is still unknown if and how search engines contribute to polarisation, indicating a clear research gap.

The comprehensive approach to understanding polarisation highlights its complexity and the various ways it manifests in society. At the macro level, polarisation is observed across societal groups within political systems, characterised by ideological distances on a spectrum such as liberal to conservative. This ideological polarisation can be perceived as the degree of divergence or alignment of political ideologies and attitudes within a society (Dimaggio et al., 1996; Fiorina & Abrams, 2008). Additionally, polarisation encompasses identity and affect, where emotional responses towards in-group and out-group members, like hatred or contempt, play a significant role (Iyengar et al., 2012). In media and communication studies, polarisation is often discussed in various forms, including ideological, issue, positional, policy, and affective polarisation. However, these concepts are sometimes interchangeable (Kubin & von Sikorski, 2021).

This research adopts a bifocal perspective on polarisation, explicitly concentrating on two distinct yet related dimensions: affective- and political polarisation. Affective polarisation causes emotional dislike and distrust towards (political) out-groups (Iyengar et al., 2019). A growing hostility and mistrust toward out-groups have been observed in Western countries, such as the United States (Iyengar et al., 2019) and the Netherlands (Albada et al., 2021; Harteveld, 2021). Political polarisation refers to the movement of political opinions away from the centre and towards more ideological extremes (Dimaggio et al., 1996; Fiorina & Abrams, 2008). It is important to note that affective polarisation can also affect political subjects. Throughout this study, the term polarisation will refer to affective polarisation.

2.3.1. Insights from Social Identity and Media Theories

Part of polarisation can be explained by combining social identity and uncertainty reduction theories. Social identity theory posits that we naturally categorise the world into an ingroup, which includes those who share beliefs, and an out-group, often defined by essential attitudes like religion and politics (Currarini & Mengel, 2016). This leads to positive sentiment towards the ingroup and negative perceptions of those identified as the out-group. The uncertainty reduction theory further explains this process, suggesting that individuals seek to reduce uncertainty about their social world by aligning more closely with their ingroup, reinforcing existing beliefs and distancing themselves from the out-group (Hogg, 2007). This alignment aids in creating a more predictable social environment, intensifying the polarisation effect. However, polarisation in digital environments, mainly through search engines, presents complex dynamics beyond traditional social identity theory and uncertainty reduction. While it is true that people tend to categorise the world into 'ingroups' and 'outgroups' based on shared attributes (Billig & Tajfel, 1973), the role of search engines adds a new layer to this phenomenon. Search engines, driven by algorithms, can amplify or mitigate polarisation by influencing the information users are exposed to.

People tend to be surrounded by similar people, which is called homophily (Mcpherson et al., 2001), which has been strongly linked to ingroup bias (Currarini & Mengel, 2016). An indication of homophily clusters is the regular and selective exposure to information, where specific ideas are more likely to spread within the cluster (Anagnostopoulos et al., 2014; Bessi et al., 2016). The homophile clusters can result in a self-reinforcing cycle of increased polarisation as individuals within a group become increasingly similar within and hostile to those outside their group. It is said that algorithms can create personalised information ecosystems that support confirmation bias (Del Vicario et al., 2017), leading users to engage predominantly with content that aligns with their existing beliefs within their homophile clusters. However, a search engine's role in polarisation might be more complex. While search engines can intensify homophily, research on these algorithm-driven environments genuinely reflects or distorts societal polarisation is still lacking.

Polarisation and selective exposure can also be explained by the effects of confirmation bias (Del Vicario et al., 2017). People strongly prefer arguments that support their worldview in choice and preference in media use, resulting in ignoring dissenting information (Anagnostopoulos et al., 2014; Knobloch-Westerwick et al., 2015; Zollo et al., 2017). A study by (Anagnostopoulos et al., 2014) conducted an online field study that investigated selective exposure and its attitudinal impacts, presenting participants with search results on political topics. The study assessed attitudes across four topics, each with eight browsing intervals, featuring articles of opposing stances and varying source credibility. Results indicate a preference for attitude-consistent messages and high-credibility sources, while exposure to contradictory content yielded the opposite effect.

In addition, confirmation bias and homophily are closely linked because both theories imply that individuals are driven to surround themselves with information that supports their worldviews while avoiding information that conflicts with those worldviews (Stroud, 2010). When individuals within a group engage in confirmation bias, they are less likely to be exposed to alternative perspectives and viewpoints, leading to a reinforced information cycle that could increase polarisation. When individuals within a group engage in confirmation bias, they are inherently less exposed to alternative perspectives and viewpoints. This leads to a reinforced cycle of information that can significantly amplify polarisation. Search engines play a crucial role in this dynamic because their algorithm-driven content curation limits exposure to these alternative perspectives. This mechanism narrows the scope of information individuals receive and significantly reduces the opportunities for users to encounter or communicate with viewpoints that challenge their preconceptions, raising question about the ethical responsibilities of online platforms.

2.3.2. Technical implications of polarisation

Adopting search engines and their algorithms has simplified finding information supporting users' worldview. The broad availability and spread of user-provided content in online (social) media have brought together people of common interests, narratives, and worldviews (Sunstein, 2002; Zollo et al., 2015). A positive aspect of user-provided content on social media is democratising information. However, it is hard to control the spread of potentially fake news, hate speech, and misinformation (Zollo et al., 2015). For example, climate change has been a polarised topic for the last few years (Treen et al., 2020), as homophily and confirmation bias have been proven futile in this discussion by spreading misinformation and disinformation online.

Two famous and seminal concepts that are possible explanations for the current rise of (online) polarisation and the spreading of misinformation are echo chambers and filter bubbles (Pariser, 2011; Sunstein, 2001). Both concepts hold that people are shut out of information that contradicts their beliefs and receive only information from like-minded sources (Pariser, 2011; Sunstein, 2001). People lose sight of other viewpoints and themes due to the biased information environment intensifying polarisation. However, evidence for the existence of filter bubbles and echo chambers has been mixed (Möller, 2021). While some research has been carried out on the effect of social media on polarisation, there have been few empirical and systematic investigations into search engines with no significant results (Bruns, 2019; Haim et al., 2018). In addition, research reveals a rise in social media studies and its influence on polarisation over the past decade (Kubin & von Sikorski, 2021). Consistently, these studies indicate that media content aligning with pre-existing attitudes significantly intensifies polarisation. However, this analysis mainly analyses Twitter data and samples from American populations. Therefore, it remains unclear what effect search engines may have on polarisation.

Echo chambers and homophily are said to fuelling polarisation in modern society with the spread of online misinformation (Vicario et al., 2016). An echo chamber emerges when a group of participants choose to connect, with the exclusion of outsiders preferentially (Sunstein, 2001). One condition for an echo chamber is that content circulates within this closed group with the result of beliefs that are being amplified. In an echo chamber, people can peruse the information that reinforces their beliefs without running into opposing views and strengthening their confirmation bias. This theory often applies to social media platforms such as Facebook and Twitter. Although echo chambers are seen as a potential factor for polarisation, the concept cannot be linked to search engines as it is not a closed group or platform. Therefore, it will be excluded further from this research.

A filter bubble occurs when a person sees only online information distorted or limited by an algorithmic bias (Pariser, 2011). The filter bubble is a defined contained space like the echo chamber. However, unlike the echo chamber metaphor, which emphasises the nature of what is inside this space, the metaphor of the filter bubbles emphasises what makes up its boundaries: algorithmic filtering (Möller, 2021). A filter bubble can be distinguished into two types: a technological bubble and a societal bubble (Dahlgren, 2021). The technical effect is that every choice affects the recommended content by a personalised algorithm. The consequences of these choices can be seen in the political process and democracy over time, also referred to as the societal bubble. Both effects can be considered to have a continuum dynamic together. Within the academic literature, most research does not clearly distinguish between the terminology of these two terms, resulting in conflicting claims about the filter bubble (Dahlgren, 2021). This personalisation might skew users' perceptions, contributing to polarisation by intensifying hostility towards differing viewpoints. The more a person sees similar information, the more the filter bubble and the algorithm strengthen. Therefore, confirmation bias has been linked to the spread of fake news and misinformation on online platforms in a way that the algorithm reinforces the spread of messages (Spohr, 2017). Altogether, when individuals are exposed to a narrow range of content, they may become more likely to view those opposing views as "the other", and mistrust grows. This effect could increase hostility and animosity towards the other party, thus potentially contributing to affective polarisation.

In conclusion, the algorithmic nature of search engines such as Google can potentially reinforce human biases. This occurs when the search algorithms prioritise content that aligns with users' existing views, beliefs, or sentiments. As most polarising content is written negatively, the SERP of polarised topics is assumed to contain more negative sentiment. In a digital environment, especially on polarised topics, these algorithms may deepen the divide by consistently presenting information that confirms and strengthens existing biases. This cycle of biased search results catering to, and further entrenching users' pre-existing beliefs and sentiment towards the ‘out-group’ could be a potential factor in the polarisation in public discourse.

2.4. Research propositions

The study aims to understand better the complex dynamics between search engine algorithms and human biases in digital information seeking to polarised topics. By examining the interplay between these elements, the research seeks to understand the implications for public discourse better. Therefore, five propositions have been created:

- **Proposition 1:** Demographic factors as age, sex, location, education level, employment, and income influence the sentiment polarity of search results.
- **Proposition 2:** The influence of political preference on the determination of sentiment polarity in search engine results is significant.
- **Proposition 3:** Short-tail queries might lead to a broader range of results, potentially including more general, less polarised, and thus more positively toned content than the more specific long-tail queries.
- **Proposition 4:** Patterns in search behaviour can be found and linked to the nature of information seeking, particularly about polarising topics.
- **Proposition 5**: Polarising topics are likely to elicit stronger emotions and opinions, potentially skewing search results towards more negative sentiments.

3. Methodology

This section explains the approach to answering the research question: 'To what extent does Google's Search Engine Results Page (SERP) differ for users in terms of result of information presentation and sentiment concerning polarising and non-polarising topics in the Netherlands?'. It outlines the research design, the methods, participants, and analysis methods.

3.1. Research Design

This research employed a mixed-method methodology, integrated both qualitative and quantitative techniques, to comprehensively understand the dynamics of sentiment and search behaviour. This approach aimed to enrich the depth and breadth of understanding while facilitating rigorous cross-validation of results (Almalki, 2016). Such a methodological framework allows for a comprehensive exploration of the research subject, using qualitative insights and quantitative rigour strengths to form a cohesive analysis. The research methodology includes a survey and an analysis of search engine data. The search engine data was derived from a browser extension. In the context of this research, the browser extension served as a tool that extends the capabilities of a web browser, enabling the collection and analysis of search data. This extension facilitated the systematic search queries thereby allowing for a comprehensive examination of search results. This combination leveraged structured questionnaire data about search behaviour and actual search results to enrich the understanding of the research theme.

First, an exploratory survey was executed to gather data regarding individual search behaviours, explicitly focusing on query types related to polarising topics. The objective was to see how individuals conduct searches and to identify the polarising topics that occupy their thoughts. The survey, disseminated through social media platforms, the BMS lab, and Sona from June until July 2023, successfully engaged 114 participants. Derived from the survey, the polarised keywords were selected as input for a browser extension. In addition, a group of general and trending keywords regarding the relevant topics were added based on Google Trends. This is because 20% of all popular keywords account for, on average, 98% of the total search queries (Skiera et al., 2010). The chosen topics were Ajax, the autumn holidays, and the change of the clock to wintertime. The search queries can be found on the next below in Table 1.

Table 1
General Search Queries

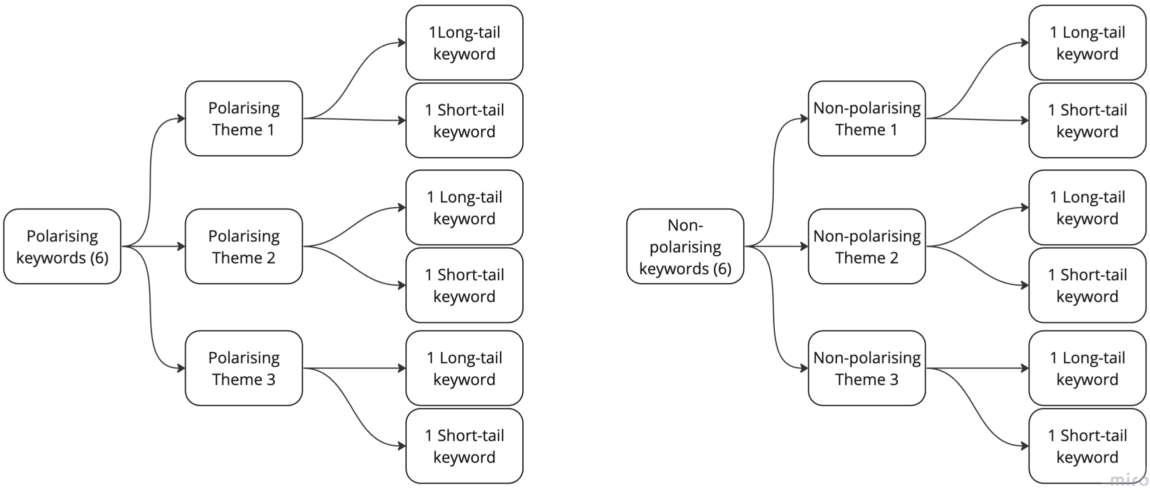
	Theme	Short-tail search query	Long-tail keyword query
General Search Queries	Football	Ajax	What is going on with Ajax?
	Holiday	Autumn Holidays	When is the autumn holidays?
	Wintertime	Wintertime	The consequences of changing the clock

The input for the browser extension categorized the queries into polarised and general categories and long-tail and short-tail classifications. This served two fundamental purposes:

1. **Distinction Between Polarised and General Queries:** It enabled a critical examination of algorithmic differences in sentiment-charged (polarised) versus more neutral (general) content. Such a comparison is instrumental in determining how much search engines might contribute to or mitigate societal polarisation.
2. **Analysis of Long-tail versus Short-tail Queries:** By differentiating between long-tail (specific, niche-focused) and short-tail (broad, high-volume) queries, this study aimed to discern the variability in search engine outputs about query specificity. This distinction is vital in assessing whether the depth and detail of a query influence the sentiment polarity.

The compilation of search queries, as displayed in Figure 2 on the subsequent page, was used in the subsequent analysis phases of the research. The second part of the study used the keywords to execute searches on participants' computers affiliated with the Digitale Polarisatie citizen science project. More information about this project can be found on the next page. It is important to note that the participants who contributed to the survey data differ from those who participated in the browser extension data collection. The final stage involved using the search data to conduct a sentiment analysis. The analyses aimed to uncover potential sentiment differences across diverse demographic groups and between polarising and general keywords. The survey and sentiment analysis will be further elaborated in the subsequent parts.

Figure 2
Schematic overview of the keywords



In collaboration with the BMS LAB at the University of Twente, this study uses a methodological approach by employing a browser extension to gain insights into the search behaviours of a diverse population within The Netherlands. This browser extension systematically conducts searches using a variety of search engines once a week. The defined search criteria were derived from this survey study. The search terms are subject to periodic revisions to maintain relevance to ongoing societal and political developments. Participation in this study is open to all residents of The Netherlands aged 18 and over. Interested individuals can contribute to this research by installing a browser extension compatible with web browsers, including Google Chrome, Mozilla Firefox, and Microsoft Edge. This methodology facilitates broad participation and ensures the collection of data representative of the general public's search patterns. The findings derived from this innovative approach are anticipated to be shared publicly on the project's website in the foreseeable future, contributing valuable insights into the digital search behaviours of the Dutch population.

Before undertaking the investigation, ethical clearance for the research was obtained from the Ethics Committee BMS under number 230687. Ethical clearance for the browser extension was obtained from the Ethics Committee BMS under number 220261. To ensure ethical clearance, the participants must be able to provide voluntary or informed consent and install the extension. All the participants are at least 18 years old and must be residents of the Netherlands. The extension installation includes a sign-up process providing demographic and consent forms. Participants can withdraw at any moment when they disable or uninstall the extension.

3.2. Survey Study

The first part of the research consisted of a survey study distributed through social media, the BMS lab, and Sona.

3.2.1. Instruments

The survey aimed to explore respondents' search behaviour concerning polarised topics on Google Search. The initial set of questions sought fundamental information about participants' search behaviour. In the subsequent part, participants selected three polarised topics they were concerned about and conducted searches using a Google Search interface. The mock-ups of the Google Search interface are located in Appendix A. Within this interface, participants could enter short- and long-tail keywords for their chosen polarising topics. These topics were predetermined and grounded in frequently recurring themes within the news. However, the participants were also free to fill in their personal polarising topics if desired. The survey data developed a list of search queries categorised into polarising and general/trending.

The survey offers several advantages and disadvantages as a core instrument in this research (Jones et al., 2013). A primary advantage lies in its ability to efficiently gather a large amount of data from many respondents. This facilitates the collection of a diverse set of responses, which is particularly valuable for examining search behaviour and attitudes toward polarising topics. Moreover, the structured nature of surveys ensures consistency in the data collected, allowing for straightforward analysis and comparison. However, there are limitations to this approach. Surveys often rely on self-reported data, which can be subject to biases such as social desirability or recall inaccuracies. This could skew the results, especially in topics of a sensitive or controversial nature. Additionally, the fixed structure of a survey can limit the depth of responses, as it confines participants to predefined options, potentially overlooking nuanced or unanticipated insights. Overall, the survey method provides a solid foundation for understanding broad trends and gathering qualitative data.

3.2.2. Participants

A total of 114 participants filled out the survey. The age and gender distribution of the participants is displayed in Table 2. The ages are widely represented, ranging from 18 to 87, with a median and mode of 28 and a mean of 34. The two most prominent groups were 18-24 (22,8%) and 25-32 (47,4%).

Table 2
The gender and age groups of the participants

	%	18-24	25-34	35-44	45-54	55-64	65-74	75+	Unknown	Total
Men	51,8%	11	31	4	4	8	0	0	1	59
Women	48,2%	15	23	5	1	7	1	2	1	55
Total	100,0%	26	54	9	5	15	1	2	2	114

In Table 3 and Figure 3, the postal codes of the participants are presented. Given that the survey was spread through a snowball sampling technique, the survey predominantly garnered responses from individuals residing in the Utrecht and Overijssel provinces. However, it should be noted that the participant demographic encompasses individuals from various regions across the Netherlands.

Table 3
Postal code distribution participants

	N	%
1000-1999	11	9,6%
2000-2999	2	1,8%
3000-3999	42	36,8%
4000-4999	1	0,9%
5000-5999	1	0,9%
6000-6999	9	7,9%
7000-7999	41	36,0%
8000-8999	5	4,4%
9000-9999	2	1,8%
Total	114	100%

Figure 3
Map of the postal code areas

3.2.3. Analysis

The analysis of the survey data was conducted through a multi-step process. The survey data played a crucial role in shaping the research methodology, particularly in developing a list of search queries. The academic version of Qualtrics was utilised to design and execute the survey. It enabled the efficient creation of a structured questionnaire and facilitated its distribution, ensuring a streamlined data collection process.

The first part focuses on individual search behaviours and identifying polarising topics. Basic descriptive statistics analysed this data. In the second part, the participants could fill out search queries for their chosen topics, divided into three long-tail and three short-tail queries each. To analyse the open answers, the initial step involved recording the frequency, particularly for the short-tail queries. Given the flexibility available to participants when formulating long-tail keywords, the resulting variations can exhibit considerable diversity. Subsequently, the responses were systematically categorised into overarching themes. Keywords that are synonyms are combined, i.e. “Number asylum seekers in The Netherlands” and “How many asylum seekers in the Netherlands”. After clustering the keywords, the short-tail and long-tail keywords were selected for the three most popular themes.

3.3. Search Engine data analysis

The second part of the research consisted of a sentiment analysis of the search results derived from the browser extension.

3.3.1. Instruments

Within this study's methodology, sentiment analysis is an essential instrument. It entails a computational process to systematically evaluate the emotional tone behind a series of words to understand the search queries. Its application in this research allows for quantifying and analysing sentiments, ranging from positive to negative, thereby providing insight into the sentiment of the search descriptions.

Sentiment analysis as an instrument in this study presents its advantages and drawbacks (Taboada, 2015). A key advantage is its ability to efficiently process large volumes of text data, providing valuable insights into public opinions and attitudes toward polarising topics. This is particularly beneficial for understanding subtle nuances in language and the emotional context surrounding search queries, which can be pivotal in examining the influence of sentiment on search engine results. However, sentiment analysis also faces challenges. One is the difficulty in accurately interpreting complex linguistic elements like sarcasm, irony, or cultural references, which can lead to misinterpretations of the sentiment. Therefore, the polarity score has been chosen. It simplifies the complex array of emotions into a more straightforward positive-negative spectrum, which helps to understand a text's overall sentiment concerning a positive or negative sentiment(Kien-Weng Tan et al., 2011). In addition, it makes comparing sentiments across different texts or datasets easier by providing the polarity as a standard metric. By analysing both positive and negative dimensions, polarity scores provide a balanced perspective on sentiment, avoiding potential biases from looking at only one type of sentiment.

In conclusion, while sentiment analysis offers a valuable lens through which to examine the nuances of search behaviour and its impact on SERPs, it is crucial to acknowledge and address its inherent limitations to ensure a well-rounded and accurate interpretation of the data.

3.3.2. Participants

118 participants were actively engaged, contributing by installing a browser extension designed explicitly for data collection. Notably, these participants were distinct from those involved in the survey component of the study. The age and gender distribution of the participants is displayed in Table 4 below. Regarding the gender distribution, 61% of the participants identified themselves as male and 37,3% as female. Besides that, there are three age spikes, namely, 16-24 (26), 45-54 (25), and 65-74 (24).

Table 4

The gender and age groups of the participants

	%	16-24	25-34	35-44	45-54	55-64	65-74	75+	Total
Men	61,0%	11	10	8	15	9	14	5	72
Women	37,3%	14	3	3	10	2	10	2	44
Other	0,9%	0	0	1	0	0	0	0	1
Unselected	0,9%	1	0	0	0	0	0	0	1
Total	100%	26	13	12	25	11	24	7	118

Tables 5 and 6 below, display the distribution of income and employment status among the participants. A predominant portion of the participants report an annual income of less than 10,000 euros, which can be attributed to the fact that many are currently enrolled as students. Digitale Polarisatie is an initiative of the University of Twente that has been actively promoted among its students.

Table 5
Income of the participants

	N	%
Less than 10.000 euro	23	19,5%
10.000 to 20.000 euro	17	14,4%
20.001 to 30.000 euro	19	16,1%
30.001 to 40.000 euro	17	14,4%
40.001 to 50.000 euro	11	9,3%
50.001 to 100.000 euro	17	14,4%
Rather not say	14	11,9%
Total	118	100%

Table 6
Employment situation

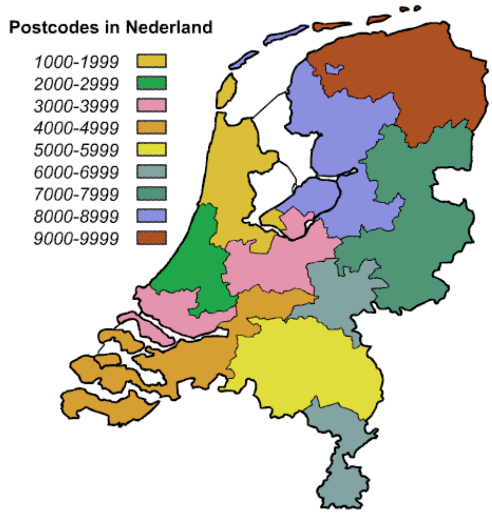
Employment situation	N	%
Full-time employment	34	28,8%
Part-time employment	17	14,4%
Unemployed	11	9,3%
Self-employed	5	4,2%
Student	23	19,5%
Retired	28	23,7%
Total	118	100%

In the following section on the next page, the postal codes of the participants are presented in Table 7 and Figure 5. Given the association of this project with the University of Twente, a significant proportion of the participants hail from the eastern region of the Netherlands, predominantly within the 7000-7999 postal code range. However, it should be noted that the participant demographic encompasses individuals from various regions across the Netherlands.

Table 7
Postal code distribution participants

Area	N	%
1000-1999	17	14,4%
2000-2999	10	8,5%
3000-3999	15	12,7%
4000-4999	10	8,5%
5000-5999	6	5,1%
6000-6999	9	7,6%
7000-7999	35	29,7%
8000-8999	7	5,9%
9000-9999	5	4,2%
Unknown	4	3,4%
Total	118	100%

Figure 4
Map of the postal code areas



The data presented in Table 8 on the next page reflects the varied political party preferences of the participants.

Table 8
Political party preference

Political Affiliation	N	%
Rather not say	26	22,0%
Groenlinks-PvdA	26	22,0%
VVD	13	11,0%
SP	10	8,5%
D66	10	8,5%
PVV	8	6,8%
Other	7	5,9%
Volt	4	3,4%
Partij voor de Dieren	4	3,4%
ChristenUnie	3	2,5%
SGP	2	1,7%
CDA	2	1,7%
BBB	2	1,7%
JA21	1	0,8%
Total	118	100,0%

Table 9
Platform usage for news

Source	N	%
TV	83	72,8%
News websites	71	62,3%
The Newspaper	43	37,7%
Radio	42	36,8%
YouTube	35	30,7%
Instagram	30	26,3%
Facebook	26	22,8%
WhatsApp	16	14,0%
Linkedin	15	13,2%
Other	14	12,3%
Twitter	11	9,6%
Telegram	4	3,5%
Reddit	3	2,6%

Finally, the subjects of the study exhibited a diversity of preferences for acquiring information and news. As indicated in Table 9 on the previous page, the participants favour traditional news sources over contemporary social media platforms. This demographic information provides the context for interpreting the results of the sentiment analysis and statistical tests presented in the subsequent sections of the Results chapter.

3.3.3. Sentiment analysis

The main target of analysis is to understand the emotional content of the collected data from the extension; therefore, a sentiment analysis was executed and followed by several statistical tests. rStudio and SPSS were used to execute the data processing for the sentiment analysis. rStudio is a programming language with a wide array of use cases. Different libraries can be installed on rStudio; every library contains modules that give rStudio additional (statistical) functionalities. These main libraries and approaches recommended by the University of Twente were used for this research. In addition, the statistical analyses were executed using SPSS. SPSS, strands for Statistical Package for the Social Sciences, is a comprehensive software suite extensively utilised for statistical analysis in diverse fields. This suite is chosen for its user-friendly functionalities, facilitating complex statistical analysis and efficient data management.

Sentiment analysis with NRC Lexicon is a technique used to extract subjective information from text by determining the attitude, sentiment, or emotional tone expressed. Sentiment analysis aims to classify text into a scale based on emotional tone. The NRC Lexicon method is used within the University of Twente and has several advantages. It has multilingual support and understands emotions. Besides that, the NRC Lexicons have been developed through academic research, ensuring high validity and reliability in their categorisation.

The process of sentiment analysis was divided into three steps. First, text pre-processing was conducted to clean and convert the text to a standard format, making extracting meaningful information from the text easier. This was done by filtering the retrieved JSON file from Digitale Polarisatie. It involved locating and removing any non-textual content unrelated to the research field from the data. Figure 5 on the next page, a single search result description is displayed.

Figure 5

Search description in Google

Once the data preparation was finished, the sentiment analysis was executed. Second, sentiment classification was performed, classifying Dutch text based on the sentiment expressed using rule-based systems or machine learning algorithms. The sentiment analysis output was generated, summarising the sentiment expressed in the text by providing an overall sentiment score or label containing the polarity score. The model generated feature vectors by assigning tags derived by subtracting the negative scores from the positive ones. This polarity differential indicates the divergence between positive and negative scores, quantifying the sentiment polarity. For each search, a polarity score was determined by aggregating the sentiment values of all words sourced from the search descriptions from the SERP, thus creating a polarity score for each search query.

The data, once processed, was subjected to a rigorous analysis using rStudio and SPSS. Various statistical tests were employed to investigate the nuances in sentiment across the dataset. The polarity scores were compared, and a selection of statistical tests was executed for the research:

1. T-tests were used to compare the means of the two groups. In sentiment analysis, this statistical test was used to compare the two groups' mean polarity or sentiment scores. This study used the t-test to compare the difference between general- and polarising search queries, short-tail- vs long-tail search queries, and genders.
2. An ANOVA compares the means of more than two groups. For sentiment analysis, ANOVA was applied to compare the mean sentiment scores of multiple demographic groups such as age, gender, education level, political preference, income, and geographical location.

4. Results

In the following section, the results of the survey study will be presented.

4.1 Results Survey

4.1.1. Search behaviour

The analysis yields insights into various dimensions of an individual's search behaviour. Understanding the dynamics of search engine usage is crucial to the study's focus on sentiment in the SERP. This survey provides valuable insights into critical aspects such as the preferred search engines, the devices predominantly used for searches, the frequency and duration of these searches, and the tendency of users to opt for either long-tail or short-tail keywords. Such information is essential to understand a search engine's information-seeking and retrieval process.

The dominance of Google as the primary search engine aligns with its widespread usage. Most participants, 88.6% (N=114), identified Google as their primary search engine. This was followed by a notably smaller percentage of users preferring DuckDuckGo and Ecosia, each accounting for 2.6% of the sample. Consequently, Google will serve as the designated search engine for this research. Most participants reported using mobile phones for searching, with 81.6% indicating this preference. This was followed by desktop computers or laptops, as noted by 63.2% of participants, and tablets, used by 10.5%. Phone usage demonstrated the highest prevalence. However, participants could select multiple devices, as shown in Figure 6. When considering multiple devices, 41.2% utilise a phone with a desktop computer or laptop to execute searches.

Figure 6
Devices used for executing searches

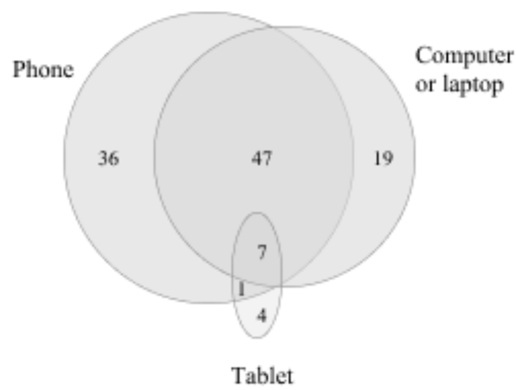


Table 10
Frequency of searches

Frequency search	N	%
Multiple times a day	74	64,9%
Daily	33	28,9%
Weekly	5	4,4%
Monthly	1	0,9%
Other, namely	1	0,9%
Total	114	100%

Table 11
Duration of searches

Duration of search	N	%
A few seconds	37	32,5%
About a minute	51	44,7%
Multiple minutes	26	22,8%
Total	114	100%

The analysis reveals that most participants (64.9%) engage in multiple daily searches, with 28.95% conducting searches at least once daily. Regarding search duration, 44.7% of participants typically spend about one minute per search, 32.5% a few seconds, and 22.8% several minutes. Most individuals (82.7%) conduct one to three searches in a session, whereas 10% perform four or more, and 7.3% conduct just a single search. These patterns imply a varying depth of information-seeking behaviour among the participants. The data reveal that during a search session, a significant segment of the population, accounting for 50.4%, predominantly engages in searches comprising one to three words. When a participant uses a phone to search, this is slightly higher (63,3%).

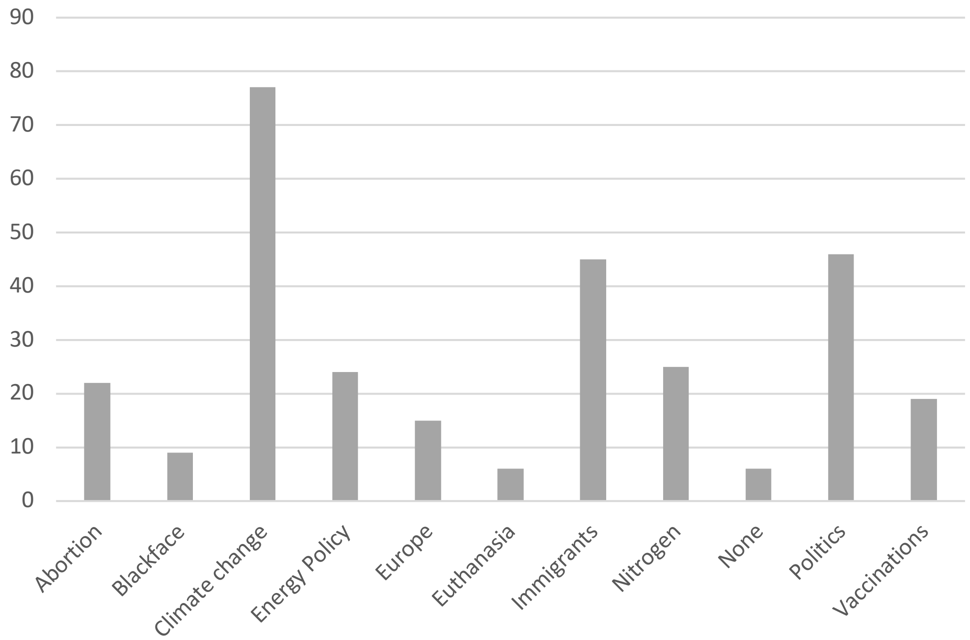
Furthermore, 36.3% of the participants utilise a blend of short- and long-tail searches. In contrast, a smaller proportion, 12.4%, is exclusively used for long-tail searches. This pattern of frequent yet brief search sessions, predominantly characterised by short-tail keywords, underscores a general preference for rapid access to information.

4.1.2. Polarising topics

The participants were asked to choose at least three topics from a predefined list they believed to be polarising. In addition to the predefined list, participants were allowed to choose their own topics. Furthermore, participants were allowed to use chosen topics as input for the survey. The distribution of choices is displayed in Figure 7 on the next page. In the context of this research, 'None' displayed in Figure 7 indicates a scenario where a participant does not express any concerns related to a polarising topic.

Figure 7

Frequency of chosen polarised topics from the survey



The survey results indicate that climate change is viewed as the most polarising issue, with 77 respondents. Such findings underscore the profound divergence in viewpoints and the intensity of opinions within the community concerning climate change. Additionally, politics was highlighted by 46 participants as a topic of significant dispute, reflecting the varied and often conflicting perspectives prevalent in political discourse. The theme of immigration, chosen by 45 participants, similarly illustrates the considerable variation in attitudes towards immigration policies and related societal issues.

The survey findings reveal that 25 participants identified Nitrogen Policy and 24 identified Energy Policy as polarising topics. Lesser recognition of issues such as Euthanasia and Blackface, each selected by 6 participants. The opportunity for participants to provide comments or additional insights on their chosen topics yielded answers predominantly centred around human rights issues and the meat industry. Based on the frequency of selections in the survey, Climate Change, Politics, and Immigration were emerged as the three most frequently identified polarising topics, which will be further explored in the subsequent stages of this research.

4.1.3. Results open answers

After participants selected their top three topics, they were tasked with providing search queries related to these chosen subjects. Specifically, each participant was required to insert a long- and short-tail keyword for each selected. The research into the nature of search queries across the three chosen topics. Initially, it was observed that individuals searching for short-tail topics tend to employ a focused approach, typically combining a thematic keyword with an additional descriptive term. This pattern suggests a preference for specificity and relevance in the search efforts. Long-term search queries were often written in a more specific matter or as a question.

In the climate change domain, the prevalent search queries predominantly revolve around the current status of climate change, the consequences, and forecasts for future developments. Political inquiries, particularly those related to political parties, demonstrate a notable inclination towards specificity, with most searches directed at individual parties. Furthermore, when exploring political party-related searches in more depth, it was revealed that long-tail search queries extend beyond mere identification, incorporating the party's stance on various issues. Lastly, search queries concerning immigrants primarily focus on their origins, demographic presence in the Netherlands, and prevailing immigrant policy.

Based on all the search queries eventually the chosen keywords are displayed in Table 12 below. These were translated from Dutch to English.

Table 12
Chosen polarising search queries derived from the survey

	Theme	**Short-tail search query**	**Long-tail keyword query**
Polarising Search Queries	Climate Change	Climate Change	Effects of climate change in the Netherlands
	Politics	Political Parties Netherlands	What are the views of party BBB?
	Immigrants	Asylum seekers	How many asylum seekers are coming to the Netherlands?

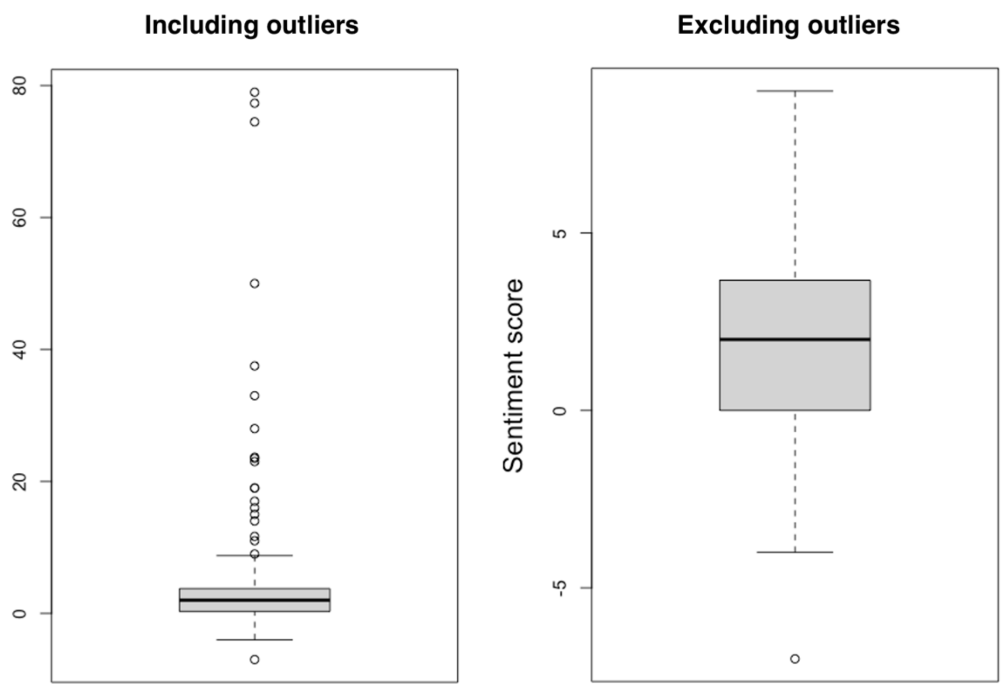
4.2. Sentiment Analysis

The sentiment analysis explored potential differences in Google's Search Engine Results Page (SERP). This aspect of the study was crucial to determine if there is a distinguishable variation in sentiment, which could imply a bias in how information is presented in search results.

4.2.1. Descriptive Statistics

At the outset of the analysis, it became evident that outliers were present within the dataset. These outliers were search results containing a significantly higher count regarding search descriptions. By closer inspection, the outliers appeared to have 17 to 20 search descriptions compared to the median range of 8 to 10. This implies that the browser extension did not limit the data extraction from the SERP, thereby capturing a greater number of search descriptions than regular searches. Consequently, these results were excluded from the analysis. After the data was processed and cleaned, 1.913 search queries with 20.304 search result descriptions, as shown in Figure 8 below.

Figure 8
Search results before and after removing the outliers



However, it is essential to note that some participants were recorded multiple times throughout the study, as the data collection extension operated weekly. Therefore, not all participants were subjected to repeated measurements over time. During the data collection, participants joined the project. Therefore, to ensure comparability of the demographic data, the scores from these multiple entries per participant were averaged per search query, providing a more consistent and representative analysis of each participant's responses. After careful data cleaning and parsing, the total search query count is 867. All search queries, including the tag in Table 13, are displayed below.

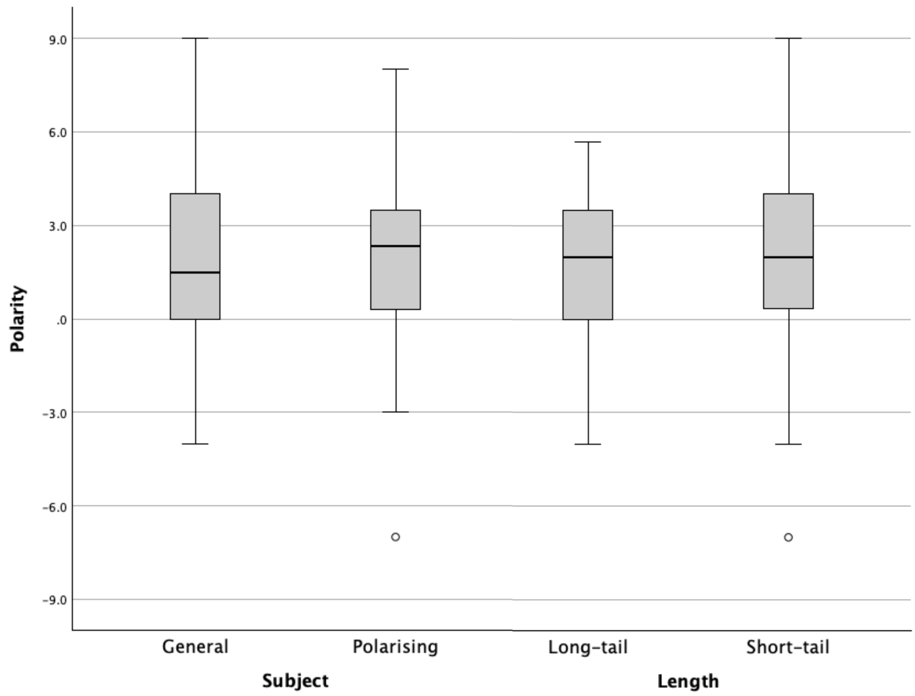
Table 13

All search queries

Search query	Polarising	Length	Tag	N
When are the autumn holidays?	General	Long-tail	LG1	77
The consequences of changing the clock	General	Long-tail	LG2	72
What is going on with Ajax?	General	Long-tail	LG3	83
Autumn Holidays	General	Short-tail	SG1	78
Wintertime	General	Short-tail	SG2	77
Ajax	General	Short-tail	SG3	85
How many asylum seekers are coming to the Netherlands?	Polarising	Long-tail	LP1	98
What are the views of party BBB?	Polarising	Long-tail	LP2	103
Effects of climate change in the Netherlands	Polarising	Long-tail	LP3	107
Asylum seekers	Polarising	Short-tail	SP1	31
BBB	Polarising	Short-tail	SP2	30
Climate Change	Polarising	Short-tail	SP3	26

An extensive analysis was undertaken to investigate the impact of sentiment on the search results generated by Google's Search Engine Results Page (SERP) algorithm. This was done by measuring the polarity, which involved subtracting the negative score from the positive score for each search result. A higher polarity score indicates a more positive sentiment of the keyword. Figure 9 below and Table 14 on the following page, show that the search queries exhibit a neutral to positive sentiment in general.

Figure 9
Box plot comparing subject and length



A closer inspection of Table 14 summarises this study's descriptive statistical analysis of the sentiment scores for all regular- and polarising search queries. The upper segment of the table outlines the statistics related to the types of search queries, whereas the lower segment focuses on the length of the search query.

Table 14
Search descriptive statistics of the polarity scores

Variable	Mean	Std. Deviation	Median	Min	Max	Skewness	Kurtosis	N
All search queries	1,99	2,27	2	-7	9	0,12	0,09	867
Regular search queries	1,83	2,44	1,5	-4	9	0,16	-0,35	472
Polarising search queries	2,17	2,04	2	-7	8	0,16	0,92	395
Short-tail search queries	2,34	2,60	2	-7	9	0,28	-0,15	540
Long-tail search queries	1,77	2,02	2	-4	5,667	-0,31	-0,42	327

Table 14 reveals that the mean sentiment score of all search queries is 1,99. The skewness suggests a slightly positively skewed distribution. The mean values across all categories fluctuate from 1,77 to 2,34. Notably, short-tail search queries have the highest average polarity score, indicating that these keywords are associated with more positive sentiments. Upon further examination of the box plots and accompanying table, and considering the variation indicated by the standard deviation, there appears to be a wide spread of data points around the mean value. The most pronounced difference is observed between long-tail and short-tail search queries. The data indicates that long-tail queries exhibit a more negative skew, whereas short-tail queries are predominantly skewed towards a positive sentiment. Short-tail search queries exhibit the highest average polarity, implying that they generally possess a more positive sentiment than other queries.

Within the dataset, the subset of polarising search queries displayed a higher average sentiment score of 2,17 compared to regular search queries, with a mean of 1,83. This indicates that polarising search results contain a more positive sentiment. An in-depth box plot illustrating the distribution of polarity scores for each search query type is displayed in Figure 10 on the subsequent page.

Figure 10
Polarity scores of the sentiment

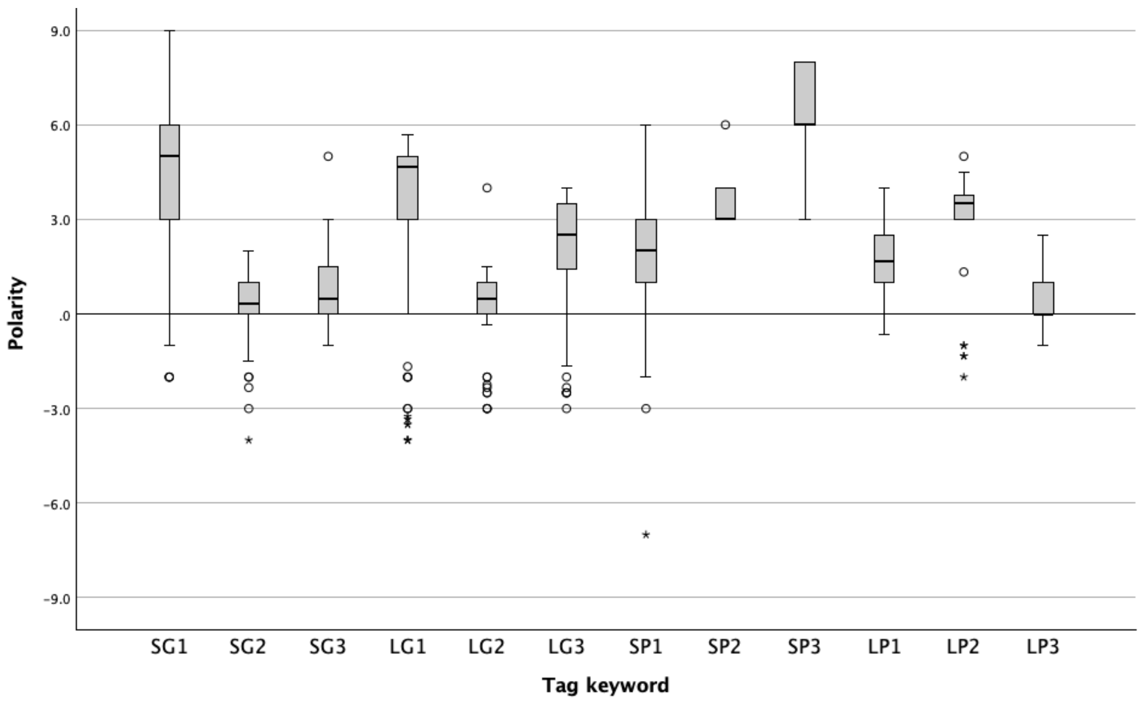
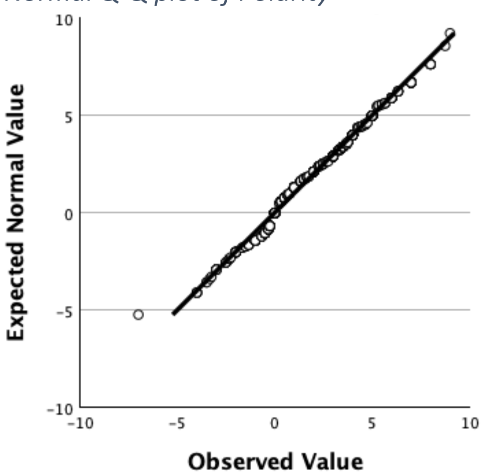


Figure 11 presents a QQ plot that reveals that most data points closely align with the reference line, indicating that the observed values conform well to a normal distribution, especially in the central part of the distribution (between approximately -2 and 2 on the expected average value scale). However, there are deviations in the tails: the lower tail (bottom left of the plot) shows that the observed values are higher than expected under normality, while the upper tail (top right of the plot) shows that the observed values are lower than expected. This is indicative of positive kurtosis.

Figure 11
Normal Q-Q plot of Polarity



4.2.2. Inferential Statistics

The t-tests and Analysis of Variances (ANOVA) were conducted in inferential statistical analysis. An independent-sample t-test was conducted to compare polarity in general and polarising search queries. There was a significant difference in the polarity scores between general search queries (M = 1.83, SD = 2.44) and polarising search queries (M = 2.17, SD = 2.04); t(865) = -2.14, p = 0.027, indicating higher polarity scores in polarising search queries. In addition, an independent-sample t-test was conducted to compare polarity in short-tail search queries and long-tail search queries. There was a significant difference in the polarity scores between short-tail search queries (M = 2.34, SD = 2.60) and long-tail search queries (M = 1.77, SD = 2.02); t(865) = 3.06, p = 0.01, indicating higher polarity scores in short-tail search queries.

The analysis of the search result content revealed that regular search queries yielded a more negative polarity than polarising search queries. Nevertheless, the overall search results demonstrated a predominance of positive terminology.

Demographic Differences Between Search Results

A series of t-tests and ANOVA tests were conducted to understand how the average sentiment scores in search results vary among different demographic groups. These tests were designed to examine whether participants from different demographic backgrounds – categorised by age, gender, education level, political preference, income, employment status, and geographical location – received search results with varying sentiment tones.

Age: The sentiment scores were analysed across seven distinct age groups using a one-way ANOVA (Analysis of Variance). This statistical test revealed that age significantly affected the sentiment polarity of search results [F(3, 863) = 3.595, p = 0.013]. This indicates that the sentiment in search results varied depending on the user's age group. Upon further examining Table 15, it becomes clear that the 18-24 age group has the lowest sentiment polarity of 1,367 and thus most negative search results, closely followed by the 45-54 age group, which displays a polarity score of 1,789. Contrariwise, the 65-74 age bracket has the highest sentiment polarity, registering a score of 2,435.

Table 15
Polarity score Age

Age	Mean	N
18-24	1,367	184
25-34	2,038	78
35-44	2,234	118
45-54	1,789	156
55-64	2,238	92
65-74	2,435	182
75+	2,099	57

Gender: Due to the insufficient representation of the categories 'Other' and 'Prefer Not to Say', they were excluded from gender-related analysis. An independent-sample t-test was used to compare the sentiment polarity in search queries between men and women. Results showed a statistically significant difference, with women receiving search results with lower polarity scores compared to men [men ($M = 2.11$, $SD = 2.25$), women ($M = 1.74$, $SD = 2.28$); $t(852) = 3.06$, $p = 0.02$], which means that women receive more negative search results in general.

Education Level: Another one-way ANOVA was conducted to assess how different educational backgrounds influenced sentiment polarity in search results. Significant variations were found across various educational levels [$F(3, 863) = 3.595$, $p = 0.013$]. Table 16 demonstrates that high school students have the lowest sentiment polarity (1,605). Contrarily, vocational education (MBO) students exhibit the highest sentiment polarity with 2,307, indicating that their search results are more positive.

Table 16
Polarity score Education

Education	Mean	N
Bachelor	2,112	177
High school	1,605	232
MBO	2,307	176
University	2,021	282

Income: A one-way ANOVA was performed for different income groups to explore how income impacts the sentiment polarity in search results. This analysis found a significant effect of income level on sentiment polarity [$F(6, 860) = 3.419$, $p = 0.002$]. A closer inspection of Table 17 shows that people with less than €10.000, - as income, have the lowest polarity score, with a score of 1,373. Remarkably, the income groups after that €10.000, - to 20.000, - have the highest polarity score (2,381).

Table 17
Polarity score Income

Income (€)	Mean	N
<10.000	1,373	148
10.000-20.000	2,381	124
20.001-30.000	2,268	158
30.001-40.000	1,956	113
40.001-50.000	2,304	106
50.001-100.000	1,799	135
Unselected	1,892	83

Employment Status: The sentiment polarity in search results also varied significantly with employment status, as indicated by another one-way ANOVA [$F(5, 861) = 3.843$, $p = 0.002$]. Table 18 on the next page, clarifies that students manifest a sentiment score of 1,342, marking the lowest score among the groups analysed. Conversely, individuals identified as unemployed register the highest sentiment score, amounting to 2,319.

Political Preferences: When examining the effect of political preferences on sentiment polarity, no significant differences were found across 15 political categories [$F(14, 852) = 5.804$, $p = 0.327$]. Due to the absence of a statistically significant difference, further examination will not be undertaken.

Table 18
Polarity score Employment

Employment	Mean	N
Full-time employment	1,998	270
Part-time employment	2,181	111
Retired	2,255	214
Self-employed	2,059	34
Student	1,342	168
Unemployed	2,319	70

Geographical Location: The impact of geographical location was analysed by grouping postal codes into clusters, as illustrated in Figure 5, page 27. A one-way ANOVA showed a significant effect of location on sentiment polarity [$F(8, 858) = 3.695$, $p < .001$]. Upon examining Table 19, it is noted that Area 7, encompassing parts of Overijssel and Gelderland, exhibits the lowest sentiment score, recorded at 1,503. In contrast, Area 8, which includes portions of Overijssel, as well as parts of Flevoland and Friesland, demonstrates the highest sentiment score, amounting to 2,586. Given the overlap among provinces, drawing a definitive conclusion from this analysis proves challenging.

Table 19
Polarity score Employment

Area	Mean	N
1	2,355	135
2	1,909	84
3	2,547	120
4	1,709	75
5	2,259	37
6	1,988	75
7	1,504	254
8	2,586	56
9	1,621	31

In conclusion, these test results highlighted that demographic factor such as age, gender, education, income, employment status, and geographical location could be of significant influence when it comes to the sentiment polarity of search results. This reflects how personal characteristics shape online information experiences, and every individual's experience can be different based on these demographic differences.

4.3. Overview

Four propositions were created for this study; every proposition will be discussed shortly based on the results. The study has shed light on significant variations in sentiment polarity across various dimensions through descriptive statistics and conducted statistical tests. The different propositions will further explain the key results:

Proposition 1: *Demographic factors as age, sex, location, education level, employment, and income influence the sentiment polarity of search results.*

The ANOVA tests conducted in the study provided valuable insights into how sentiment responses vary across different demographic groups. The findings highlight the influence of factors as age, location, education level, and income on differences in sentiment polarity, thus supporting Proposition 1 partly.

Proposition 2: *The influence of political preference on the determination of sentiment polarity in search engine results may be significant.*

The study's findings did not show a significant effect of political preference on sentiment polarity. This suggests that while political preference may have been hypothesized to influence sentiment polarity in search results, the evidence did not support this hypothesis statistically significantly.

Proposition 3: *Short-tail queries might lead to a broader range of results, potentially including more general, less polarised, and thus more positively toned content than the more specific long-tail queries.*

The study's findings indicate that search query length plays a more significant role than the subject matter in determining sentiment polarity. It was found that short-tail search queries generally exhibit a more positive sentiment, while long-tail search queries display a broader distribution with a slightly more negative sentiment. This finding aligns with Proposition 4, emphasizing the importance of query length in shaping the sentiment polarity of search results.

Proposition 4: *Patterns in search behaviour can be found and linked to the nature of information seeking, particularly about polarising topics.*

The study identified three main polarising topics: climate change, politics, and immigration chosen by the participants, which seemed significant. In addition, it reveals the dominant use of Google as the primary search engine. The frequent and short searches underscore a pattern of how people seek information online in Google.

Proposition 5*: Polarising topics are likely to elicit stronger emotions and opinions, potentially skewing search results towards more negative sentiment polarity.*

The study revealed that contrary to Proposition 1, the overall sentiment polarity in Search Engine Results Pages (SERPs) for polarised topics tends to be more positive than negative on average. This finding challenges the initial expectation that polarising topics evoke predominantly negative sentiments in search results. The initial expectations of (polarised) news is often written in negative sentiment (Hilbig, 2012; Sacerdote et al., 2020; Soroka et al., 2019).

5. Discussion

The primary objective of this study was to investigate the existence of a correlation between sentiment, specifically in terms of polarity and Search Engine Results Page (SERP) outcomes.

5.1. Main Findings

In the following part, the interpretations of the results will be discussed. The influence of demographic factors like age, location, education level, and income on the sentiment polarity of search results aligns with the theoretical framework's emphasis on the role of human biases in information retrieval (Kliman-Silver et al., 2015). The research highlights significant differences in sentiment polarity across various demographics, such as age, location, education level, and income. This finding aligns with proposition one, that these factors influence sentiment polarity in search results. Notably, differences in sentiment between students and employed participants may reflect contrasting lifestyles and interests, underscoring the role of socioeconomic factors in shaping online search experiences. This suggests that these biases are not just individual but can be influenced by broader societal and demographic factors. Proposition one builds upon the understanding that search engine results are not merely technical outputs but might be shaped by complex interactions between algorithmic processing and human socio-economic contexts (Haider & Sundin, 2019; Kliman-Silver et al., 2015).

Furthermore, the study uncovered that the length of search queries (distinguishing between short-tail and long-tail) significantly influences sentiment polarity. This insight adds another layer of complexity to understanding how sentiment influences SERPs, suggesting that the relationship between sentiment and information presentation in SERPs is complex but multifaceted. The data indicate a distinct difference in sentiment polarity associated with the length of search queries. Short-tail queries exhibited a more positive sentiment polarity than long-tail queries. This observation substantiates that short-tail queries might lead to less polarised and more positively toned content. The breadth of interpretation allowed by short-tail keywords grants search algorithms more flexibility, influencing the sentiment polarity of the results (Skiera et al., 2010). This finding adds another dimension to understanding search algorithms, suggesting that query length significantly shapes search outcomes. This is closely linked to the theoretical framework's discussion on the mechanics of search results and how the nature of search queries might influence sentiment within the SERP.

Overall, a more neutral polarity score across search results aligns with previous research on search engine algorithms, indicating that these systems are increasingly designed to provide users with a balanced and diverse set of information sources (Patil et al., 2021). Contrary to proposition five, that polarising topics would elicit predominantly negative sentiments in search results, the study's findings reveal a tendency towards more positive sentiment polarity. The discovery contradicts previous research, which often posits that polarised or contentious topics are associated with negative narratives (Hilbig, 2012; Soroka et al., 2019). However, a study conducted by Fletcher and Jenkins (2019) into the dynamics of polarisation within news media challenges prevalent assumptions about media consumption and polarisation. Their study observed minimal evidence to support the concept that increased exposure to digital media significantly contributes to widespread polarisation, whether it aligns with or opposes one's views. This finding is critical in re-evaluating the common belief that platforms act as echo chambers or filter bubbles, predominantly exposing users to homogeneous viewpoints.

The findings contribute to a nuanced understanding of the media's role in polarisation, emphasizing the complexity of the relationship between media consumption patterns, information seeking, and the development of polarised attitudes. A possible explanation is that the Netherlands has a less polarised media landscape and less selective exposure towards polarised news (Trilling et al., 2017). A study by Trilling et al. (2017) examined the impact of news media exposure on public attitudes toward immigration in the Netherlands. Contrary to findings from the United States, the study posited that selective exposure occurs but does not necessarily lead to polarisation. The researchers postulated that the differential effects observed in the Netherlands, as compared to the USA, might be attributable to the Netherlands' less polarised media landscape, which could reduce the likelihood of individuals encountering highly partisan news. This is further supported by recent annual research by the Reuters Institute that indicates that Dutch news maintains high trust levels, ranking equally fourth among 46 countries surveyed (Newman et al., 2023). In particular, the public broadcaster NOS emerges as the most used and trusted brand, closely followed by its commercial counterpart, RTL Nieuws, and local and regional newspapers. These results highlight a trend of trust in traditional and established news sources. Contrariwise, the report reveals that more partisan media outlets and tabloid press are less trusted, suggesting a possible correlation between perceived media bias and trustworthiness. These observations may support the hypothesis that Dutch media perceived as less partisan or sensationalist tend to be more trusted by the public (Newman et al., 2023; Trilling et al., 2017).

5.2. Implications of the Study

Despite the relatively modest significance of the findings in this research, the study still holds important implications for the understanding of search engines and polarisation. The variability in sentiment polarity across search queries and demographic details underscores an interaction between search algorithms and user demographics (Haider & Sundin, 2019; Kliman-Silver et al., 2015). This variability explains a dimension of SERPs that could be subject to the subtle influences of algorithmic biases and user profiles (Pariser, 2011). These findings, therefore, catalyse future research, particularly in investigating the intersection of search engine algorithms with socioeconomic and demographic differences among users. Furthermore, the study illustrates the complex nature of information-seeking and retrieval processes, emphasising the need for a more profound understanding. Despite its limitations, the methodological approach of this study provides a valuable template for future investigations, offering a unique perspective on the interaction between user behaviour, search engine algorithms, and the resultant sentiment polarity of search queries.

This study contributes to the ongoing discourse of the (technological) filter bubble (Flaxman et al., 2016; Möller, 2021). However, the mechanics of a filter bubble remain unclear. It invites further inspection and debate, particularly in challenging or refining existing theories or assumptions about filter bubbles. The findings underscore the importance of continued exploration of polarisation, especially in the context of technological influence. This suggests a need for more comprehensive, perhaps even interdisciplinary, approaches to research in this domain.

Moreover, there is a pressing requirement for quicker research efforts, given technology's continuous and rapid evolution (Patil et al., 2021). The implications of these findings are profound, especially considering the pivotal role of search engines in information dissemination and public opinion formation. The bias in search engine results can contribute to the digital polarisation of society by perpetuating filter bubbles, where users are predominantly exposed to information that aligns with their algorithmic profile. This has significant implications for democratic discourse, as it can lead to a more divided and less informed public.

In conclusion, this research contributes to the ongoing discourse on polarisation and the role of technology in information distribution. The tested propositions demonstrated the subtle variations in sentiment polarity. It questions the neutrality of search engines, suggesting a more complex and dynamic interplay between user profiles, algorithmic processing, and content presentation. Therefore, while the findings of this study may not be ground-breaking, they are instrumental in shaping a more nuanced understanding of the digital landscape, particularly in the context of search engines and polarisation.

5.3. Limitations of the Study

While providing valuable insights, this study has its limitations. First and foremost, the scope of the research was constrained by the sample size and sample representation. This limitation potentially affects the generalizability of the findings and suggests the need for caution when extending these results to broader contexts. Another notable limitation concerns demographic diversity and geographical coverage. Most participants are from the East of the Netherlands due to the location of the University of Twente. The study's findings are based on a sample of many students, which may only partially represent the wider population. The demographic composition of the sample does not accurately reflect the Dutch society. Thus, future research should include a more diverse and representative sample to validate and expand upon these findings. In addition, while comprehensive, the analytical framework used in this study may only encompass some possible perspectives or variables relevant to Dutch society. Therefore, the interpretations and conclusions are among many possible cultural contexts, encouraging further exploration and discussion in other regions.

A limitation of this study lies in its temporal scope. The study did not incorporate a longitudinal dimension, restricting its ability to assess changes and trends over time. Multiple entries from individual were consolidated into a single record. Consequently, while the study provides valuable insights into the state of the subject matter at a specific point in time, its findings may only partially represent the dynamic nature of the topic under study.

Therefore, future research would benefit from incorporating a longitudinal approach to capture the temporal variations and provide a more comprehensive understanding of the subject matter. Furthermore, the choice of survey distribution channels (social media, BMS lab, and Sona) and the demographic makeup of the participants (predominantly from the East of the Netherlands) influence the representativeness of the data, impacting the generalizability of the findings.

While effective for this research, the chosen trending general search queries about Ajax employed for this study may have introduced specific biases. This is primarily attributable to the prevalent negative sentiment surrounding Ajax during the study period. In addition, owing to errors in the extension of Digitale Polarisatie, there has been an underrepresentation of short-tail polarising keywords. On the other hand, long-tail polarising keywords have been overrepresented, which mitigates this imbalance in polarising keywords.

The reliance on sentiment analysis impacted the data's accuracy or depth. While the lexicon-based method from the University of Twente was employed for sentiment analysis, it is essential to note that more advanced methodologies exist, such as Google's BERT algorithm implemented through Python. This sophisticated machine learning approach examines the interrelationships between words within a sentence rather than assigning scores to individual words, offering a potentially more nuanced analysis. Nevertheless, it must be acknowledged that using algorithms or machine learning techniques, including those as advanced as BERT, has limitations. These methods often stumble to accurately interpret nuances such as sarcasm, cultural references, or unconventional expressions of sentiment, which can lead to inaccuracies in sentiment analysis.

Finally, specific unforeseen challenges, such as web extension software development issues, were encountered during the research. This research highlights different areas for improvement of the web extension for future research. In addition, the observed search results may contain bias from the inherent predispositions of the algorithm, which could be influenced by the nature of the earlier executed searches by the extension. For example, if the extension searches every week for polarised topics, it could influence the algorithm and, therefore, the search results.

In conclusion, while this study has contributed new information to the field of search engines, the limitations outlined above should be carefully considered. They offer a framework for interpreting the results with appropriate caution and a roadmap for future research endeavours seeking to build upon this work.

5.4. Recommendations for Further Studies

The outcomes of this study not only augment our current understanding of the field of search engine algorithms but also unveil numerous opportunities for future research. Reflecting upon the limitations and insights gained, several vital recommendations emerge for guiding future investigations.

A primary suggestion is the expansion of the research scope. Future studies should include a broader spectrum of more generalizable participants, offering a more holistic perspective on political influence. Additionally, exploring different political preferences within varied contexts, types of search queries and faster iterating would provide invaluable comparative insights and deepen the comprehension of the influencing variables. In addition, faster iterating would prevent a biased search engine by repeatedly executing search queries weekly. Future studies should undertake a longitudinal approach to understand evolution and trends better. This would provide a deeper insight into the dynamic nature of the subject matter, offering a more comprehensive understanding of how phenomena develop and change.

The options for diverse methodologies present another avenue ripe for exploration. Adopting alternative approaches, including different types of sentiment analysis models and datasets, could reveal new perspectives and surmount some of the constraints faced in this study. Applying the BERT algorithm might give a better understanding of the sentiment. Longitudinal studies represent a crucial strategy for understanding the temporal dynamics of search results. By tracking the evolution of the search engine results page over time, such studies can shed light on long-term trends and impacts, offering a richer understanding of the subject matter. This can be done with the new data the University of Twente project collects.

Moreover, embracing interdisciplinary approaches could significantly enrich future research. Integrating theories and methodologies from search queries and results can gain new theoretical and practical insights. This confluence of perspectives, especially from communications and computer sciences, is essential for developing a comprehensive understanding of the complex interactions of digital information seeking and retrieval. Finally, addressing the specific gaps identified in this study is paramount. Future research endeavours should focus on the specific differences between the demographic characteristics of people, which currently represent underexplored areas within the literature. Exploring the demographics separately, which has received relatively limited attention, could yield significant contributions to the field. In addition, conducting similar studies in different cultural and linguistic settings to compare the impact of sentiment on search engines globally could also contribute.

In summary, these recommendations aim to steer future research towards enriching knowledge about algorithms and sentiment. Building on the groundwork laid by this study, pursuing these suggested directions is anticipated to advance our knowledge and understanding of polarisation and sentiment significantly, fostering further scholarly inquiry and practical applications.

As last, this thesis underscores the critical need to understand and address the sum of sentiment on digital polarisation, particularly in search engines. The findings highlight the need for further research to promote a more balanced and less polarised digital landscape.

6. References

Abramowitz, A. I., & Saunders, K. L. (2008). Is polarization a myth? *Journal of Politics*, *70*(2), 542–555. https://doi.org/10.1017/S0022381608080493

Albada, K., Hansen, N., & Otten, S. (2021). Polarization in attitudes towards refugees and migrants in the Netherlands. *European Journal of Social Psychology*, *51*(3), 627–643. https://doi.org/10.1002/ejsp.2766

Allahverdyan, A. E., & Galstyan, A. (2014). Opinion dynamics with confirmation bias. *PLoS ONE*, *9*(7). https://doi.org/10.1371/journal.pone.0099557

Almalki, S. (2016). Integrating Quantitative and Qualitative Data in Mixed Methods Research—Challenges and Benefits. *Journal of Education and Learning*, *5*(3), 288. https://doi.org/10.5539/jel.v5n3p288

Amabile, T. M. (1983). Brilliant but Cruel: Perceptions of Negative Evaluators. In *JOURNAL OF EXPERIMENTAL SOCIAL PSYCHOLOGY* (Vol. 19).

Anagnostopoulos, A., Bessi, A., Caldarelli, G., Del Vicario, M., Petroni, F., Scala, A., Aldo Moro Roma, P., Fabiana Zollo, I., San Ponziano, P., & Walter Quattrociocchi, I. (2014). *Viral Misinformation: The Role of Homophily and Polarization General Terms Misinformation, Virality, Attention Patterns*.

Anderson, C. (2006). *Long Tail: Why the Future of Business Is Selling Less of More*.

Apuke, O. D., & Omar, B. (2021). Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telematics and Informatics*, *56*. https://doi.org/10.1016/j.tele.2020.101475

Azzopardi, L. (2021). Cognitive Biases in Search: A Review and Reflection of Cognitive Biases in Information Retrieval. *CHIIR 2021 - Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*, 27–37. https://doi.org/10.1145/3406522.3446023

Barlow, J. P. (1996). *A declaration of the independence of cyberspace*. https://www.eff.org/cyberspace-independence.

Bessi, A., Petroni, F., Vicario, M. Del, Zollo, F., Anagnostopoulos, A., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Homophily and polarization in the age of misinformation. *European Physical Journal: Special Topics*, *225*(10), 2047–2059. https://doi.org/10.1140/epjst/e2015-50319-0

Billig, M., & Tajfel, H. (1973). *Social categorization and similarity in intergroup behaviour'*.

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, *15*(3), 209–227. https://doi.org/10.1007/s10676-013-9321-6

Bruns, A. (2019). *Are Filter Bubbles Real?*

Casal Bértoa, F., & Rama, J. (2021). Polarization: What Do We Know and What Can We Do About It? *Frontiers in Political Science*, *3*. https://doi.org/10.3389/fpos.2021.687695

Chiou, L., & Tucker, C. (2018). *Fake News and Advertising on Social Media: A Study of the Anti-Vaccination Movement*. http://www.nber.org/papers/w25223

Cho, J., & Roy, S. (2004). *Impact Of Search Engines On Page Popularity*. Association for Computing Machinery.

Cinelli, M., De, G., Morales, F., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). *The echo chamber effect on social media*. https://doi.org/10.1073/pnas.2023301118/-/DCSupplemental.y

Cui, G., Lui, H. K., & Guo, X. (2012). The effect of online consumer reviews on new product sales. *International Journal of Electronic Commerce*, *17*(1), 39–58. https://doi.org/10.2753/JEC1086-4415170102

Currarini, S., & Mengel, F. (2016). Identity, homophily and in-group bias. *European Economic Review*, *90*, 40–55. https://doi.org/10.1016/j.euroecorev.2016.02.015

Czerlinski, J. (1999). *Simple Heuristics That Make Us Smart Evaluation Techniques for Uncertainty Representation and Reasoning View project Demonstration-Guided Deep Reinforcement Learning for Sensor Allocation View project*. www.mpib-berlin.mpg.de/abc

Dahlgren, P. M. (2021). A critical review of filter bubbles and a comparison with selective exposure. In *Nordicom Review* (Vol. 42, Issue 1, pp. 15–33). Sciendo. https://doi.org/10.2478/nor-2021-0002

Del Vicario, M., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2017). Modeling confirmation bias and polarization. *Scientific Reports*, *7*. https://doi.org/10.1038/srep40391

Dimaggio, P., Evans, J., Bryson, B., Bartels, L., Hochschild, J., Nee, V., Smith, T., Starr, P., Taylor, H., & Western, B. (1996). *Have Americans' Attitudes Become Polarized?*

Epstein, R., & Robertson, R. E. (2015). The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(33), E4512–E4521. https://doi.org/10.1073/pnas.1419828112

Esau, K., Choucair, T., Vilkins, S., Svegaard, S., Bruns, A., O'connor, K., & Lubicz, C. (2023). *Destructive Political Polarization in the Context of Digital Communication-A Critical Literature Review and Conceptual Framework*.

Fallows, D. (2005). *How Women and Men Use the Internet Summary of Findings*. http://www.pewinternet.org

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annual Review of Political Science*, *11*, 563–588. https://doi.org/10.1146/annurev.polisci.11.053106.153836

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, *80*(Specialissue1), 298–320. https://doi.org/10.1093/poq/nfw006

Fletcher, Richard., Jenkins, Joy., & European Parliament. European Parliamentary Research Service. Scientific Foresight Unit. (2019). *Polarisation and the news media in Europe : a literature review of the effect of news use on polarisation across Europe*.

Frost, P., Casey, B., Griffin, K., Raymundo, L., Farrell, C., & Carrigan, R. (2015). The Influence of Confirmation Bias on Memory and Source Monitoring. *Journal of General Psychology*, *142*(4), 238–252. https://doi.org/10.1080/00221309.2015.1084987

Gilovich, T., Kerr, M., Husted Medvec, V., Dunning, D., Hayes, A., Madey, S., McElwee, B., Regan, D., & Story, A. (1993). Effect of Temporal Perspective on Subjective Confidence. In *Journal of Personality and Social Psychology* (Vol. 64, Issue 4).

Golebiewski, M., & Boyd, D. (2019). *Data Voids: Where Missing Data Can Easily Be Exploited*. https://www.researchgate.net/publication/356909935

Google Search Central. (n.d.). *In-depth guide to how Google Search works*. Retrieved July 22, 2022, from https://developers.google.com/search/docs/advanced/guidelines/how-search-works

Grind, K., Schechner, S., McMillan, R., & West, J. (2019, November 15). *How Google Interferes With Its Search Algorithms and Changes Your Results*. The Wall Street Journal. https://www.wsj.com/articles/how-google-interferes-with-its-search-algorithms-and-changes-your-results-11573823753

Haas, A., Unkel, J., & Munich, L. (2017). *Ranking versus Reputation: Perception and Effects of Search Result Credibility*.

Haider, J., & Sundin, O. (2019). Invisible search and online search engines: The ubiquity of search in everyday life. In *Invisible Search and Online Search Engines: The Ubiquity of Search in Everyday Life*. Taylor and Francis. https://doi.org/10.4324/9780429448546

Haim, M., Graefe, A., & Brosius, H. B. (2018). Burst of the Filter Bubble?: Effects of personalization on the diversity of Google News. *Digital Journalism*, *6*(3), 330–343. https://doi.org/10.1080/21670811.2017.1338145

Harper, T. (2017). The big data public and its problems: Big data and the structural transformation of the public sphere. *New Media and Society*, *19*(9), 1424–1439. https://doi.org/10.1177/1461444816642167

Harteveld, E. (2021). Fragmented foes: Affective polarization in the multiparty context of the Netherlands. *Electoral Studies*, *71*. https://doi.org/10.1016/j.electstud.2021.102332

Helberger, N., Kleinen-von Königslöw, K., & van der Noll, R. (2015). Regulating the new information intermediaries as gatekeepers of information diversity. *Info*, *17*(6), 50–71. https://doi.org/10.1108/info-05-2015-0034

Herr, P. M., Kardes, F. R., & Kim, J. (1991). *Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective*. http://jcr.oxfordjournals.org/

Hersher, R. (2017, January 10). *What Happened When Dylann Roof Asked Google For Information About Race? : The Two-Way : NPR*. NPR. https://www.npr.org/sections/thetwo-way/2017/01/10/508363607/what-happened-when-dylann-roof-asked-google-for-information-about-race

Hilbig, B. E. (2012). Good things don't come easy (to mind): Explaining framing effects in judgments of truth. *Experimental Psychology*, *59*(1), 38–46. https://doi.org/10.1027/1618-3169/a000124

Hills, T. T. (2019). The Dark Side of Information Proliferation. *Perspectives on Psychological Science*, *14*(3), 323–330. https://doi.org/10.1177/1745691618803647

Hogg, M. A. (2007). Uncertainty-Identity Theory. In *Advances in Experimental Social Psychology* (Vol. 39, pp. 69–126). https://doi.org/10.1016/S0065-2601(06)39002-8

Introna, L. D. (1999). *Sustaining the Public Good Vision of the Internet: The Politics of Search Engines*. https://www.researchgate.net/publication/24116930

Introna, L. D., & Nissenbaum, H. (2000). Shaping the web: Why the politics of search engines matters. *Information Society*, *16*(3), 169–185. https://doi.org/10.1080/01972240050133634

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). *The Origins and Consequences of Affective Polarization in the United States*. https://doi.org/10.1146/annurev-polisci-051117

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. In *Public Opinion Quarterly* (Vol. 76, Issue 3, pp. 405–431). https://doi.org/10.1093/poq/nfs038

Joachims, T., Granka, L., Pan, B., Hembrooke, H., & Gay, G. (2005). *Accurately Interpreting Clickthrough Data as Implicit Feedback*.

Jones, T. L., Baxter, M., & Khanduja, V. (2013). A quick guide to survey research. *Annals of the Royal College of Surgeons of England*, *95*(1), 5–7. https://doi.org/10.1308/003588413X13511609956372

Kahneman, D. (2003). A Perspective on Judgment and Choice: Mapping Bounded Rationality. In *American Psychologist* (Vol. 58, Issue 9, pp. 697–720). https://doi.org/10.1037/0003-066X.58.9.697

Kahneman, D. (2011). Thinking, fast and slow. In *Thinking, fast and slow.* Farrar, Straus and Giroux.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The Reviewing of Object Files: Object-Specific Integration of Information. In *COGNITIVE PSYCHOLOGY* (Vol. 24).

Kien-Weng Tan, L., Na, J.-C., Theng, Y.-L., & Chang, K. (2011). Sentence-Level Sentiment Polarity Classification Using a Linguistic Approach. In *LNCS* (Vol. 7008). www.blogspot.com

Kliman-Silver, C., Hannak, A., Lazer, D., Wilson, C., & Mislove, A. (2015). Location, location, location: The impact of geolocation on web search personalization. *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, *2015-October*, 121–127. https://doi.org/10.1145/2815675.2815714

Knobloch-Westerwick, S., Johnson, B. K., & Westerwick, A. (2015). Confirmation bias in online searches: Impacts of selective exposure before an election on political attitude strength and shifts. *Journal of Computer-Mediated Communication*, *20*(2), 171–187. https://doi.org/10.1111/jcc4.12105

Knobloch-Westerwick, S., Mothes, C., & Polavin, N. (2020). Confirmation Bias, Ingroup Bias, and Negativity Bias in Selective Exposure to Political Information. *Communication Research*, *47*(1), 104–124. https://doi.org/10.1177/0093650217719596

Kubin, E., & von Sikorski, C. (2021). The role of (social) media in political polarization: a systematic review. *Annals of the International Communication Association*, *45*(3), 188–206. https://doi.org/10.1080/23808985.2021.1976070

Lam, O., Wojcik, S., Broderick, B., & Hughes, A. (2018). *Gender and Jobs in Online Image Searches* (Vol. 17). www.pewresearch.org.

Law, S. C., Commons, D., & Goldman, E. (2006). *Search Engine Bias and the Demise of Search Engine Utopianism*.

http://digitalcommons.law.scu.edu/facpubshttp://digitalcommons.law.scu.edu/facpubs/76

Mcpherson, M., Smith-Lovin, L., & Cook, J. M. (2001). *Birds of a Feather: Homophily in Social Networks*. https://doi.org/10.3410/f.725356294.793504070

Möller, J. (2021). Filter bubbles and digital echo chambers. In *The Routledge Companion to Media Disinformation and Populism* (pp. 92–100). Taylor and Francis Inc. https://doi.org/10.4324/9781003004431-10

Mønsted, B., & Lehmann, S. (2022). Characterizing polarization in online vaccine discourse—A large-scale study. *PLoS ONE*, *17*(2 February). https://doi.org/10.1371/journal.pone.0263746

Napoli, P. M. (2014). Automated media: An institutional theory perspective on algorithmic media production and consumption. *Communication Theory*, *24*(3), 340–360. https://doi.org/10.1111/comt.12039

Newman, N., Fletcher, R., Eddy, K., Robertson, C. T., & Kleis Nielsen, R. (2023). *Reuters Institute Digital News Report 2023*. https://doi.org/10.60625/risj-p6es-hb13

Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. In *Review of General Psychology* (Vol. 2, Issue 2).

Noble, S. (2018). *Algorithms of Oppression*. NYU Press.

Norocel, O. C., & Lewandowski, D. (2023). Google, data voids, and the dynamics of the politics of exclusion. *Big Data and Society*, *10*(1). https://doi.org/10.1177/20539517221149099

Novin, A., & Meyers, E. (2017). Making sense of conflicting science information: Exploring bias in the search engine result page. *CHIIR 2017 - Proceedings of the 2017 Conference Human Information Interaction and Retrieval*, 175–184. https://doi.org/10.1145/3020165.3020185

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, *12*(3), 801–823. https://doi.org/10.1111/j.1083-6101.2007.00351.x

Pariser, E. (2011). *The Filter Bubble: What the Internet is Hiding from You by* (Vol. 1). Penguin Press. https://doi.org/10.5070/D482011835

Patil, A., Pamnani, J., & Pawade, D. (2021, April 2). Comparative Study of Google Search Engine Optimization Algorithms: Panda, Penguin and Hummingbird. *2021 6th International Conference for Convergence in Technology, I2CT 2021*. https://doi.org/10.1109/I2CT51068.2021.9418074

Peeters, G. (1971). *The positive-negative asymmetry: on cognitive consistency and positivity bias*. https://doi.org/https://doi.org/10.1002/ejsp.2420010405

Purcell, K., Brenner, J., & Rainie, L. (2012). *Seach Engine Use 2012*.

Rieder, B., & Sire, G. (2014). Conflicts of interest and incentives to bias: A microeconomic critique of Google's tangled position on the Web. *New Media and Society*, *16*(2), 195–211. https://doi.org/10.1177/1461444813481195

Sacerdote, B., Sehgal, R., & Cook, M. (2020). *Why Is All COVID-19 News Bad News?* https://statsiq.co1.qualtrics.com/public-

Scheufele, D. A. (1996). *Framing as Media Effects Framing as a Theory of Media Effects*.

Schultheiß, S., & Lewandowski, D. (2021). How users' knowledge of advertisements influences their viewing and selection behavior in search engines. *Journal of the Association for Information Science and Technology*, *72*(3), 285–301. https://doi.org/10.1002/asi.24410

Search Engine Market Share Worldwide | Statcounter Global Stats. (n.d.). Retrieved July 21, 2022, from https://gs.statcounter.com/search-engine-market-share

Sire - Stichting Ideele Reclame. (n.d.). Retrieved April 5, 2023, from https://sire.nl/

Skiera, B., Eckert, J., & Hinz, O. (2010). An analysis of the importance of the long tail in search engine marketing. *Electronic Commerce Research and Applications*, *9*(6), 488–494. https://doi.org/10.1016/j.elerap.2010.05.001

Soroka, S., Fournier, P., & Nir, L. (2019). Cross-national evidence of a negativity bias in psychophysiological reactions to news. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(38), 18888–18892. https://doi.org/10.1073/pnas.1908369116

Spohr, D. (2017). Fake news and ideological polarization: Filter bubbles and selective exposure on social media. *Business Information Review*, *34*(3), 150–160. https://doi.org/10.1177/0266382117722446

Stroud, N. J. (2010). Polarization and partisan selective exposure. *Journal of Communication*, *60*(3), 556–576. https://doi.org/10.1111/j.1460-2466.2010.01497.x

Sunstein, C. (2001). *Echo Chambers*. Princeton University Press.

Sunstein, C. R. (2002). The Law of Group Polarization*. In *The Journal of Political Philosophy* (Vol. 10, Issue 2).

Taboada, M. (2015). *Sentiment Analysis: An Overview from Linguistics*. https://doi.org/10.1146/annurev-linguistics-011415-040518

Tandoc, E. C., Thomas, R. J., & Bishop, L. (2021). What is (Fake) news? analyzing news values (and more) in fake stories. *Media and Communication*, *9*(1), 110–119. https://doi.org/10.17645/mac.v9i1.3331

Taylor, G. (2013). Search quality and revenue cannibalization by competing search engines. *Journal of Economics and Management Strategy*, *22*(3), 445–467. https://doi.org/10.1111/jems.12027

Taylor, S. E. (1991). Asymmetrical Effects of Positive and Negative Events: The Mobilization-Minimization Hypothesis. In *Psychological Bulletin* (Vol. 110, Issue l).

Treen, I., P Williams, H. T., & O, S. J. (2020). *Online misinformation about climate change*. https://doi.org/10.1002/wcc.665

Trilling, D., Van Klingeren, M., & Tsfati, Y. (2017). *Exposure, Political Polarization, and Possible Mediators: Evidence From the Netherlands*.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. In *New Series* (Vol. 185, Issue 4157).

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. In *Journal of Risk and Uncertainty* (Vol. 5). http://about.jstor.org/terms

Van Couvering, E. (2017). The Political Economy of New Media Revisited. *Proceedings of the 50th Hawaii International Conference on System Sciences (2017)*. https://doi.org/10.24251/hicss.2017.220

Van Hoboken, J. (2012). *Search Engine Freedom*.

Vicario, M. Del, Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(3), 554–559. https://doi.org/10.1073/pnas.1517441113

Wakabayashi, D. (2017, October 9). *Google Finds Accounts Connected to Russia Bought Election Ads - The New York Times*. https://www.nytimes.com/2017/10/09/technology/google-russian-ads.html

Wijnhoven, F., & van Haren, J. (2021). Search Engine Gender Bias. *Frontiers in Big Data*, *4*. https://doi.org/10.3389/fdata.2021.622106

Zollo, F., Bessi, A., del Vicario, M., Scala, A., Caldarelli, G., Shekhtman, L., Havlin, S., & Quattrociocchi, W. (2017). *Debunking in a world of tribes*. https://doi.org/10.1371/journal.pone.0181821

Zollo, F., Novak, P. K., Del Vicario, M., Bessi, A., Mozetič, I., Scala, A., Caldarelli, G., Quattrociocchi, W., & Preis, T. (2015). Emotional dynamics in the age of misinformation. *PLoS ONE*, *10*(9). https://doi.org/10.1371/journal.pone.0138740

Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., De Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review*, *5*(1). https://doi.org/10.14763/2016.1.401

7. Appendix

7.1. Appendix A: Survey mimics search screen google

Figure 12

Google Search in Qualtrics on desktop

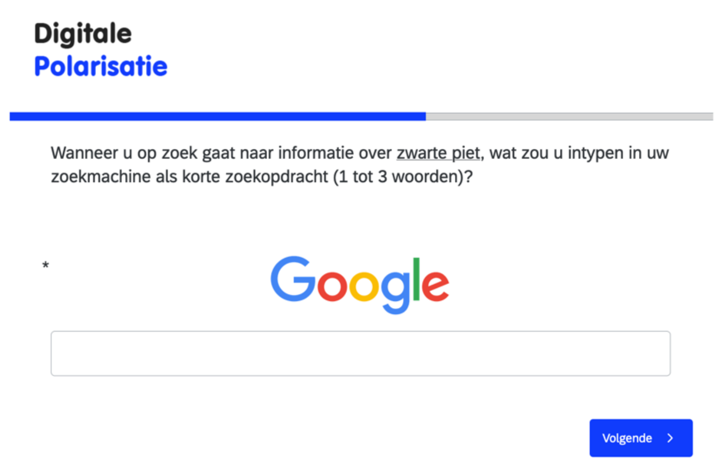


Figure 13

Google Search in Qualtrics on Mobile

7.2. Appendix B: Results Survey

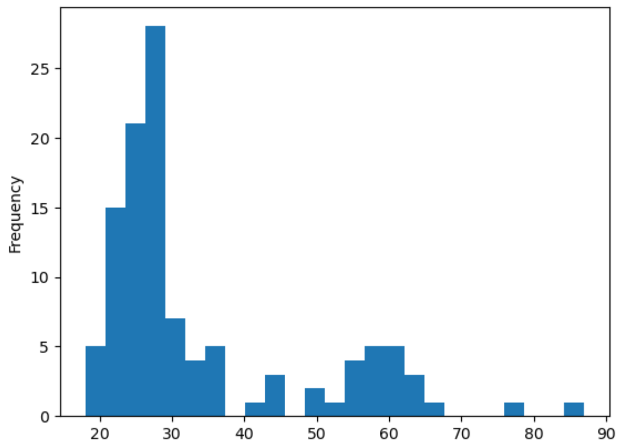
Below all the questions and answers are stated.

Q1.1 What is your gender?

Field	Count (%)
Man	59 (51.75%)
Woman	55 (48.25%)
Different	0 (0.00%)
Prefer not to say	0 (0.00%)

Q1.2 What is your age?

Age distribution



Median 28.0

Mean 34.14

Field	Count (%)
18-24	25 (22.32%)
24-34	55 (49.11%)
35-44	9 (8.04%)
45-55	5 (4,46%)
55-64	15 (13.39%)
64-74	1 (0.79%)
75+	2 (1.79%)

Q1.3 What are the first four numbers of your ZIP code?

Map distribution

Q1.4 About which polarising issues are you most concerned about? (Maximum 3 answers possible)

Field	Choice count
Climate change	77 (67,54%)
Politics	46 (40,35%)
Immigrants	45 (39,47%)
Nitrogen	25 (21,93%)
Energy Policy	24 (21,05%)
Abortion	22 (19,30%)
Vaccinations	19 (16,67%)
Europe	15 (13,16%)
Blackface	9 (7,89%)
Euthanasia	6 (5,26%)
None	6 (5,26%)
Total	294

SEARCH BEHAVIOUR

Q2.1 Which search engine do you typically use to perform online searches?

Field	**Choice count (n=114)**
Google	101 (88.60%)
Bing	1 (0.88%)
Yahoo!	1 (0.88%)
DuckDuckGo	3 (2.63%)
Brave	0
Startpage	1 (0.88%)
AOL search	0
Anders, namelijk:	7 (6.14%)

Q2.2 On which device do you typically use to perform online searches?

Device Type	Choice count (%)
Desktop computer or laptop	72 (63,16%)
Phone	93 (81,58%)
Tablet	12 (10,53%)
Other, namely:	0

Device Type	Choice count (%)
Desktopcomputer or laptop, Telefoon	47 (41.23%)
Telefoon	36 (31.58%)
Desktop Computer of laptop	19 (16,67%)
Desktop, computer of laptop, Telefoon, Tablet	7 (6.14%)
Tablet	4 (3,51%)
Telefoon, Tablet	1 (0,88%)

Q2.3 Frequency – How often do you use a search engine?

Field	Choice Count (%)
Multiple times a day	74 (64.91%)
Daily	33 (28.95%)
Weekly	5 (4.39%)
Monthly	1 (0.88%)
Other, namely:	1 (0.88%)

Q2.4 On average, how long do you spend on a search?

Field	Choice Count (%)
Several Seconds	37 (32.46%)
About one minute	51 (44.74%)
Multiple minutes	26 (22.81%)
Other, namely:	0 (0.00%)

Q2.5 When you search for one specific topic, how many different searches do you perform?

Field	Choice count
One search	9 (7.27%)
One to three searches	91 (82.73%)
Four or more searches	11 (10.00%)
Other, namely:	0

Q2.6 AI Chat – Did you ever asked a question to chatbots such as ChatGPT (Open AI)?

Field	Count (%)
Yes	79 (69.30%)
No	32 (28.07%)
I don’t know what that is	3 (2.63%)

Q2.7 AI Chat Freq. - How often does a chatbot such as ChatGPT use?

Field	Count (%)
Multiple days	7 (9.09%)
Daily	8 (10.39%)
Weekly	31 (40.26%)
Monthly	20 (25.97%)
Other, namely:	11 (14.92%)

Q2.7 During a search, you can use short or long searches. For example: Short search: "Pancakes recipe" Long search: "How do I make batter for pancakes?" Do you often use short or long searches during a search?

Field	Count (%)
Short searches (one to three words)	57 (50,44%)
Long searches (sentences)	14 (12.39%)
Short as well as long searches	41 (36.28%)
Other, namely:	1 (0.88%)