The Political Use of Search Engines: Search Tendencies and Partisan Personalization in Google Search Results

by

Chau T. T. Tong

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Mass Communications)

at the
UNIVERSITY OF WISCONSIN-MADISON
2021

Date of final oral examination: 05/24/2021

The dissertation is approved by the following members of the Final Oral Committee:

Barry C. Burden, Professor, Political Science
Nicholas Diakopoulos, Associate Professor, Communication and Computer Science, Northwestern University
David Kaplan, Patricia Busk Professor, Quantitative Methods, Educational Psychology
Douglas McLeod, Evjue Centennial Professor, Journalism and Mass Communication
Hernando Rojas (Chair), Centennial Professor, Journalism and Mass Communication
Michael W. Wagner, Professor, Journalism and Mass Communication

For Lan, who made this journey possible for me,
and who I completed this for.

ACKNOWLEDGMENTS

This dissertation will never be completed without the help and support of so many people. It took me a long time ruminating on the topic, the scope of this work, and how to implement and materialize these ideas. And yet for everything I wrote, I am indebted to the following and so many others:

I deeply thank my advisor, Hernando Rojas, who has always been supportive of my ideas and the directions of my research. He has been a generous and critical first reader of all my works, giving me expert and practical guidance, answering my ruminations on life, and on being an international scholar navigating academia. I am grateful for his professional mentorship, as well as his kind check-ins and encouragements in times of personal struggles.

My sincere thanks and appreciation go to the five members of my committee: My interest in American politics and partisanship was really drummed up through *Barry Burden's* class Mass Political Behavior. Lessons from that class helped me generate and connect ideas for a lot of future research. One particular lesson that follows me until this day is to think deeply about how study limitations impact the ways we interpret the results. *Nick Diakopoulos* agreed to be in my committee despite the sudden request and was crucial for the completion of this dissertation. I am grateful for his invaluable insights into search engines and algorithms, as well as his immediate and constructive feedback whenever I reached out for questions. A huge thank you to *David Kaplan* for his guidance with statistics and quantitative methodologies. Through our discussions and interactions, I have learnt so much and become much more confident and rigorous in methods. *Doug McLeod* and *Mike Wagner* are among the academic heroes in the J-school who introduced me to political communication research, whose inputs have greatly enriched the quality and depth of this work, and for all that I am extremely indebted to them.

This text is also my thanks to numerous friends and colleagues in the J-school and UW-Madison, especially Robert Gabriel, Ayellet Pelled, Juwon Hwang, Jinha Kim, Hyesun Choung, SK Tay, Jiyoun Suk, Margarita Orozco, Chanida Jib; thanks to Lisa Aarli and all the faculty and staff in the J-school, Doug Hemken at *UW-Madison Social Science Computing Cooperative*, the *Data Science Hub*, the *Writing Center*, the two grad school coordinators Alexandra Walter and Elena Hsu.

Special thanks to Desheng Hu and the *Algorithm Auditing Research Group* at Northeastern University, Daniel Trielli and Nicholas Vincent at *Northwestern University* for giving me ideas, inspirations, and for creating the tools to conduct auditing research. Thanks to Sebastián Valenzuela for being a mentor, a cheerleader and a sounding board.

Thanks to Jeanne and Dilip, and Susan for being my guardian angels during these years. I'm extremely lucky to have met you and know you. You enriched my time in Madison with your warmth, generosity, kindness, and made such a profound impact on my growth as a person.

Lastly, my deepest love and thanks go to my little family. To my mother, Lan, who is the reason for this and for everything I will ever do. To my sister, Tien, who teaches me about work ethics, who shares with me the joys and pains of this career path, and who I will continue to turn to in all future trials and adventures.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Digital platform companies are taking increasingly important roles in the contemporary information life of citizens. Together with social media giants like Facebook and Twitter, Internet search engines are vital information intermediaries that could enable, channel, or inhibit exposure to diverse media. While social media use and their democratic impact have gained meticulous academic attention, relatively less research in the domain of mass communication has been conducted on the effects of search personalization on citizens' access to news and political information. Search engine personalization effects have not been widely studied mainly because of the heterogeneity in individuals' information-seeking habits and the opaque mechanisms of algorithmic search results. This dissertation contributes to a growing line of research on the democratic role of search media, in particular the use of search engines for political information and the potential personalization in search results, which carry far-reaching implications for contemporary issues including access to legitimate content, exposure and susceptibility to misinformation, and digital information disparities.

The dissertation examines three research problems: first, the political and informational factors that explain individual differences in search perceptions and attitudinal outcomes, which are crucial in understanding the extent to which people rely on search engines for political information; second, the extent to which information-seeking behaviors among politically engaged individuals, i.e. political partisans, reflect ideological bias and news use habits; and third, the extent to which biased search queries and political ideology lead to differences in search results.

Using a combination of methods including secondary survey analysis, original survey design and a series of crowdsourced experiments, the dissertation reports three main findings:

First, the breadth and variety of Internet use, political media use, and media trust were important determinants of search-related outcomes and political search behavior. Knowledge of search engine result determinants significantly predicted political search frequency, and significantly mediated the effects of Internet use, political interest, political media use, media trust, and search ability on political search frequency. Second, political partisans indicated preference towards specific biased terms as search queries and reported search queries that reflected issue positions under accuracy and directional motivated information-seeking goals. Third, Google search results returned from biased queries led conservatives and liberals to different sets of information, but search result differences were driven largely by specific search queries than by the political ideology of the searchers.

INTRODUCTION

Search engines have become primary conduits to access news and online information in the current digital age. Alongside social media, online search platforms like Google, Bing, Baidu, and Yahoo are pivotal information gatekeepers that drive major traffic to news sites and online outlets (Newman et al., 2020; Sullivan, 2021). In particular, Google is the dominant search engine in most Western markets, processing more than 40,000 search queries every second worldwide, and accounting for about 89% market share in the U.S. (Statcounter, 2021). With such a monumental presence, Google Search plays a tremendous role in every aspect of public life, including shaping and diverting public attention to different issues and topics.

In an increasingly networked, high-choice information environment where information flows in a hybrid media system between old and new media (Chadwick, 2017) and the public agenda can be set by more than a single entity (Gruszczynski & Wagner, 2017), it is crucial to understand the growing power that search engines like Google exert with regard to information access and equalities.

Scholars who have theorized about these changes in the information environment come from two main theoretical perspectives. The first is the technologically deterministic perspective, which emphasizes how new technologies transform and dictate the ways of life and manners of communication. Concepts like “filter bubble” or the winner-takes-all “power law” express concerns about digital algorithms and personalization characterized by modern technologies, which will put people in information cocoons and reinforce the dominance of a few influential actors. For instance, algorithms designed to support search technologies, which optimize the capacity of users to receive the exact information they are looking for, could amplify audience fragmentation and individualization trends (Just & Latzer, 2017). In addition, this could lead to the reinforcement of congenial information and suppression of countervailing information, limiting access to content that can enlighten one's preexisting view. From a dystopian perspective, personalized reality constructions enabled by algorithms could make individuals and voters more susceptible to manipulation tactics and information disparities (Epstein & Robertson, 2015; Epstein, Robertson, Lazer & Wilson, 2017).

The second perspective stresses the social shaping of technology, which could include the role of active user input e.g., feedbacks and interactions in addition to algorithmic filtering as the "secondary gatekeeper" in the production and consumption of content (Singer, 2013; Wallace, 2018). According to this viewpoint, predicting the impact of technologies would be greatly amiss without understanding the role of a range of factors including social, cultural, organizational, psychological, economic in the design, reconstruction and implementation of technologies (R. Williams & Edge, 1996). Thus, search, as other types of media, must be viewed in this social process that takes into account users' interactions in the shaping of their usage, change and innovation (Dutton et al., 2017).

The inquiries in this dissertation are inspired by these fundamental theories related to the interactions of users and search technologies and their significance in the domain of politics. In particular, three empirical studies are conducted to explore several research questions regarding the role of search in the political information-seeking process.

Dissertation Overview

The dissertation consists of three parts:

The *first part*, "Chapter I: Examining the Determinants of Political Search: A Structural Equation Modelling Approach", takes advantage of a large-scale survey dataset specifically exploring the significance of online search in the political processes. A secondary analysis is conducted to examine the demographic, political, and informational determinants of political search tendencies, particularly investigating the mediating roles of search-related attitudinal and perceptional outcomes (search efficacy, search engine knowledge, perceived search accuracy) in explaining individual differences in the frequency of political search behaviors.

The *second part*, "Chapter II: Political Predispositions, Partisan Media Use and Confirmatory Search Tendency among Political Partisans", is developed through an original survey design. This part examines the relationships between political predispositions, i.e. political ideology and issue position, partisan media use and confirmatory search tendencies among political partisans in the U.S. Operationalizing "confirmatory search tendency" with two measures: open-ended self-reported search queries under two conditions of motivated reasoning goals (accuracy and directional), and ranking of biased terms as preferred search queries, this part focuses on the questions of the extent to which political liberals and conservatives engage in confirmatory search tendencies with regard to polarizing political issues (abortion, climate change, gun control, climate change), and the potential mediating role of partisan media consumption in these relationships.

The *third part*, "Chapter III: Biased Search Queries and Google Search Results for Liberals and Conservatives" systematically explores the extent to which Google Search presents political liberals and conservatives with different search results regarding three issues (election integrity, abortion, climate change), arguably creating the so-called ideological "filter bubbles" for users (Pariser, 2011). Building on the current literature on search engine auditing, this part uses real search results data submitted by users at different locations across the U.S. at three different time points, to shed light on the two mechanisms under which personalization might occur, i.e. the political identification of search users (liberals/conservatives) and users' input in the form of biased search queries (liberal vs conservative).

Significance

In the current information environment, information seeking through search is an essential mechanism through which citizens learn about politics, form political opinions, and engage in political behaviors. The technical aspects of search technologies and individuals' use tendencies can profoundly impact this process. Personalization algorithms undoubtedly play a part in creating potential ideological filter bubbles that might separate users from alternative perspectives, but biased inputs in the form of search queries can also introduce vastly different narratives in the results (Kulshrestha et al., 2019; Tripodi, 2018).

Altogether, this dissertation investigates the significance of search engines, particularly Google Search, in citizens' access to and consumption of political information (Granka, 2010). The contribution of this work is in providing insights into three conundrums: how users' information habits determine the way they use search engines, how search behaviors are colored by innate political predispositions, and how such behaviors affect the information individuals are exposed to.

CHAPTER I: EXAMINING THE DETERMINANTS OF POLITICAL SEARCH: A STRUCTURAL EQUATION MODELLING APPROACH

Together with social media, search engines are important pathways through which people encounter important information, including information about politics, government and public affairs. A large-scale survey on the importance of search in seven nations found that search engines were among the first places people reported going for information about politics (such as political candidates) or information about important national or international issues (Dutton, Reisdorf, Dubois, & Blank, 2017). Also, according to Dutton et al. (2017), search media were most often used for information about a particular topic, to navigate to sites, and to a lesser extent, to look up facts, to check the accuracy of news/information, and to find information about politics or current events. Not surprisingly, people who are more politically interested, who go online for different activities, and who use more media sources for news also tend to rely more on search engines for political purposes (Blank & Groselj, 2014; Cotter & Reisdorf, 2020). However, during high-profile events such as political elections, even those with little interest in politics and government affairs tune in to find information about the running candidates and their policy platforms. Even politicians' gaffes or trivial events could trigger search activities, leading citizens to seek out more deeper information related policy issues (Trevisan et al., 2018).

The rise of mis/disinformation on social media and the development of technologies including algorithmic filtering of online content have engendered drastic changes to the current information ecosystem, raising public distrust and apathy in their interactions with news media (Fletcher & Nielsen, 2019; Fletcher & Park, 2017; Strömbäck et al., 2020). In contrast, trust in search engines in particular Google remains extremely high (Pan et al., 2007), especially for the young generation, whose Internet experience starts with and revolves around Google (Gunter et al., 2009). In this context, understanding the conditions under which people search for political information is crucial. These include, for example, the factors that could explain public confidence and efficacy navigating information online, or the role of knowledge and trust regarding search engines that could improve the use of these tools for political information.

In short, the antecedents of search-related attitudes, knowledge and skills and how they translate into decisions to conduct politically related searches or to fact-check political information merit further attention. What is learned from such examination could help identify individuals who could benefit from support and training in order to use search more effectively, especially in light of evidence suggesting the differential effects of search results on voting preferences among different groups (Epstein, Robertson, Lazer, et al., 2017; Epstein & Robertson, 2015).

This analysis examines the relative influence of different factors that could explain individual variations in outcomes related to online search engine use. In particular, two research questions guiding such inquiry are as follows:

RQ1. What are the factors that influence outcomes related to search engine use (search efficacy, search engine knowledge, perceived accuracy of search results), and the use of search for political information?

RQ2. To what extent do search-related outcomes (search efficacy, search engine knowledge, perceived accuracy of search results) mediate the effects of political and informational factors on the use of search for political information?

The Role of Demographics, Political and Internet Dispositions

With the proliferation of Internet use, it has become clear that digital engagement and skills play a significant role in individuals' success in a range of outcomes, from professional career development to social services uptake and participation in democratic processes (Robinson et al., 2015). However, the expansive literature on digital exclusion has established that even in countries with high levels of Internet penetration and technology (e.g. smartphone) adoption, inequalities exist between groups of different social stratifications. Race and ethnicity (Mesch & Talmud, 2011), gender (Ono & Zavodny, 2003), age (Hargittai, 2010), socio-economic status (Witte & Mannon, 2010) all influence not only the amount and breadth of Internet activities but also what content people look for online (Blank & Groselj, 2014; Buente & Robbin, 2008; Helsper & Deursen, 2017; van & Helsper, 2015).

With innovations in technologies, secondary digital disparities concerning skills, efficacy, and knowledge introduce more challenges to the already disadvantaged groups and create more differentiation in terms of technological adoption for important social and political activities (Dutton & Reisdorf, 2019; Gran et al., 2020a; Klawitter & Hargittai, 2018; Monzer et al., 2020). With regard to using search engines for political information, individuals' social and economic status can affect what people search for or how they perform simple or sophisticated searches. For example, a college education may provide ones with knowledge of how to conduct research using keywords and library databases, or how to apply advanced search filters to increase the relevance of the results. Similarly, higher income might mean access to a faster Internet connection and support resources to retrieve the information ones need. Socioeconomic conditions can also affect factors that directly influence behaviors, such as the perceived self-efficacy, knowledge about how online search works, or whether people will be satisfied by search results.

The "Social Structural Model" and the "Standard Social Psychological Model" depict two causal mechanisms behind information behavior (Scheufele, Nisbet, M., Brossard & Nisbet, E., 2004). The sociological framework highlights social structural determinants such as the sociological context, and the social psychological framework foregrounds individual subjective orientations and attributes. Based on these two models, the relationships between sociodemographic variables and individual characteristics including cognitive and affective factors can be examined in tandem to assess political search behaviors.

Antecedents of Search Efficacy

As the new information environment poses increasing challenges in terms of the propagation of misinformation and unregulated information flow, modern users face a sense of "collective anxiety" and uncertainty caused by the increase of information in volume, quality, and scope (Guo, Lu, Kuang & Wang, 2020; Liang & Fu, 2017; Qiu et al., 2017). Information overload occurs when the "efficiency in using information is hampered by the amount of relevant, and potentially useful, information available" (Bawden & Robinson, 2009, 2020). As news and information on any given topic have escalated dramatically, feelings of overload and inefficiency could result in decreased engagement and avoidance (Park, 2019; Qiu et al., 2017; Skovsgaard & Andersen, 2020). Facing such changes, critical media literacy skills are needed to interpret and verify information. Individuals also need an effective approach to gather information and a sense of self-efficacy to navigate the new media ecology (Fletcher & Nielsen, 2019; Livingstone, 2019; Xiao et al., 2021).

According to social cognitive theories, self-efficacy is the belief about one's ability to perform particular behaviors which affects not only what people do, but also the efforts and invested time put into an activity (Bandura, 2001). In the context of technological adoption, self-efficacy is an important determinant of the perceived ease of use (PEOU) and perceived usefulness (PU) of technologies, two important elements in the Technology Acceptance Model

(Davis, 1989). Efficacy also plays an important role in successful information seeking (Afifi, 2017) and effective search performance (Parissi et al., 2019; Thompson et al., 2002).

Previous research has shown that greater self-efficacy in information and communication technologies is correlated with higher socioeconomic status, younger age, actual and perceived experience with technology (Hargittai & Dobransky, 2017; Olsson et al., 2019; Wickens & Miller, 2020). Self-efficacy with regard to one's own knowledge about politics also plays a key role in the link between exposure to news (online and offline), and democratic outcomes. For example, evidence from panel survey findings show that consistent use of online news media fosters individual efficacy, which in turn significantly influences turn-out and political participation (Moeller, Kuhne, & De Vreese, 2018). Similarly, individual perception of efficacy stimulates news and communication activities such as news use and discussion, which in turn positively influences political participation (Gil de Zúñiga, Diehl & Ardévol-Abreu, 2017).

The two dimensions of external and internal political efficacy suggest that efficacy relates to trust in social and government institutions (Craig, Niemi & Silver, 1990). Equally important, studies that investigate the relationship between news consumption, information efficacy and attitudes towards emerging news trends (e.g. misinformation, news personalization) suggest that in a convoluted information landscape, efficacy is contingent on how individual approach information and assign trust to different information sources (Anspach & Carlson, 2018; Bodó et al., 2019; Bradshaw, 2019).

A typology of Internet users introduced by Pew Research Center (2017) – a leading organization in tracking Americans' adoption of technologies – revealed four clusters of users with distinct attitudes and levels of engagement with information ("Confident", "Eager/Willing", "Cautious/Curious" and "Doubtful"). Despite the expectation that efficacy would greatly vary based on group membership, a deeper examination of the data revealed that feelings of information overload and struggle with online information were common even for groups with contrasting information outlooks. In particular, the “Eager/Willing” cluster reported at least some difficulties finding information online, despite having a very high interest in news and a high level of trust in national news organizations. More importantly, individuals in this cluster were predominantly non-White and tended to fall below average in access to multiple tech tools. Such feelings of inefficacy were also found in the “Cautious/Curious” cluster, which had a high interest in learning, but low levels of digital skills and some distrust of news and information sources. In contrast, feelings of struggle were least reported in the “Doubtful” group, which had very low trust in information sources and a low appetite for information.

Altogether, the findings from this Pew Research study suggest that efficacy can clash with other factors which in turn influence technological use for political activities. Thus, it is important to investigate the antecedents of search efficacy and whether search efficacy influences the political use of search engines. Drawing on previous research, the following hypotheses are posited:

H1. While general Internet use (a), political interest (b), media trust (c) and search ability (d) are positively correlated with search efficacy, the number of media sources for political information (e) is negatively correlated with search efficacy due to induced stress of information overload.

Antecedents of Search Engine Knowledge

Search engines work by collecting hundreds of billions of web pages and indexing them with meta information, such as the keywords found in the page’s content, the freshness of the page, and previous user engagement with the page. When users enter a search query into a search engine, search engine algorithms look for relevant pages in the index and hierarchically rank these pages to produce a set of search results. In addition to the search queries, other data are also used to return the results, including location, language settings, previous search history, and the device where the search originates (Bozdag, 2013; Kitchin, 2017; Kliman-Silver et al., 2015; Latzer, 2016). The political and commercial aspects of modern search engines directly concern the display and ranking of search results, which could be affected by advertisers buying ads to target queries relevant to their products, optimization efforts to maximize visibility and improve ranking in search result pages, as well as what search algorithms infer from users' demographic information and online activities (Bradshaw, 2019; Giomelakis & Veglis, 2015).

Given the diverse use of search engines for different purposes including politics, it is essential to understand how search engines work and what influences the search results. However, it has been shown that the general public does not have a good understanding of how search engines work (e.g. Cotter & Reisdorf, 2020). This is a common lack of awareness about how algorithms on digital platforms and services work in general (Eslami et al., 2016; Eslami & Karahalios, 2017). For instance, Eslami et al.'s study (2015) using test experiments and interviews on 40 Facebook users found that 62.5% of the participants were not aware that Facebook posts on their feeds are algorithmically filtered. Another survey study on Facebook users (Rader & Gray, 2015) found wide variations in the perceptions and understanding of filtering algorithms.

Being aware of how algorithms influence what people see, however, is a crucial meta-skill that can enhance other skills and digital benefits. At the beginning of the twentieth century, only experienced and sophisticated users were aware that search engines return sponsored results (Fallows, 2005). Two decades later, knowledge of search personalization based on location and sponsored ads has greatly improved; however, terms and concepts like "search optimization" (i.e. the strategies websites use to maximize their visibility and ranking) are still hardly grasped by casual and novice searchers (Pew Research Center, 2019).

As knowledge of search engines can be considered a subset of knowledge about digital technologies, prior research has focused on socioeconomic status and demographic factors in explaining the search engine knowledge gap. For instance, Cotter & Reisdorf (2020) found that awareness of the political aspects of search engines varied by education levels and the breadth of Internet use for different daily-life activities.

As knowledge can be built over time through practice and accumulation of experience, it is arguable that factors like interest in politics, multiple media use, and search ability increase awareness of how search engines work. For example, being politically engaged and keeping up with the news regularly can inform individuals about issues like website tracking, or how search engines keep track of users' search activities and tailor results based on such information. Similarly, online search experiences allow individuals to develop search strategies that would help them become more effective in retrieving needed information, such as formulating and reformulating queries, adopting advanced search features, using Boolean terms. Thus, a second set of hypotheses is formed:

H2. General Internet use (a), political interest (b), media trust (c), search ability (d), and the number of media sources for political information (e) will all be positively correlated with search engine knowledge.

Antecedents of Perceived Accuracy of Search

Discerning the accuracy of search results is important in deciding whether search media will be trusted for important behaviors, such as fact-checking activities. If search results are deemed accurate and reliable, the likelihood of using search for political information will be higher; in contrast, lack of trust in search engines and results will reduce their perceived utility.

Previous research establishes a relationship between trust in the information coming from different sources and the use or non-use of such sources. Unsurprisingly, individuals trust their own sources and feel generally skeptical of the media they don't feel familiar with. In particular, lack of trust in the news media was found to be associated with increased use of alternative and non-mainstream sources, including social media, digital-only sources, or news aggregators (Fletcher & Park, 2017; Hoff & Bashir, 2015; Monzer et al., 2020). Trust was also influenced by political views, and the general public showed a declined trust in a wide range of democratic institutions (Strömbäck et al., 2020; A. E. Williams, 2012).

Regarding trust in technology, to a certain extent, people's trust in technological systems represents their trust in the designers of such systems (Parasuraman & Riley, 1997). Thus, trust (and distrust) in social media and search engines could be extended to trust assigned to tech companies, who hold increasingly powerful control over online information. Public sentiment indicates that the general Americans remain wary of the role social media sites play in delivering credible information, and trust in social media is significantly lower compared to sources like professional outlets or friends and family (Pew Research Center, 2020). Major tech companies have been accused of political bias and stifling open discussion; in particular, the belief that tech companies intentionally censor conservative political viewpoints is commonly held among Republicans and conservatives (Byrnes, 2020). These perceptions of bias have led the political right to react in adversarial ways towards the mainstream media and big tech companies.

Extant research suggests that trust in search results and the consequent decisions to utilize search for political information is a matter of political identity as much as other contextual and motivational factors. Accordingly, two hypotheses are formed as follows:

H3. Media trust (a), search ability (b), and the number of media sources for political information (c) will be positively correlated with the perceived accuracy of search results.

H4. There will be a main effect of political identity, i.e. identifying as a) liberal Democrats and b) conservative Republicans on the perceived accuracy of search results. In particular, liberal Democrats would perceive search results to be more accurate (positive association), and conservative Republicans would perceive search results to be less accurate (negative association).

Finally, to consider the mediating role of search-related outcomes (search efficacy, search engine knowledge, perceived accuracy of search results) on the political use of search, a research question is posed:

RQ. To what extent do search-related outcomes (search efficacy, search engine knowledge, perceived accuracy of search results) mediate the effects of political and informational factors on the use of search for political information?

Methods

Data

The data was collected by the Quello Center at Michigan State University as part of a larger project titled "Search and Politics: The Uses and Impacts of Search in Britain, France, Germany, Italy, Poland, Spain and the United States" (Dutton et al., 2017). Access to the data was provided by the authors, and for this work, a secondary analysis was conducted on the U.S. dataset (N = 2,018). The data consisted of a random sample of American Internet users aged 18 and older using pre-qualified panels. The selection of respondents was done using quotas based on age, gender, and region. The resulting sample was weighted using age, gender and region so that it matched known national population proportions. The questionnaire contained items for approximately 244 variables and was fielded in January 2017.

Analysis strategy

Structural equation modeling technique (SEM) was used to explore the factors that influence individual variations in attitudinal and behavioral outcomes related to search engine use. SEM is a multivariate extension of regression in which all multiple predictors and outcomes (with underlying covariance structure) are examined simultaneously. SEM also provides the innovation of examining the latent structure behind “search engine knowledge” and “search efficacy” – the two potential mediators of interest that were not observed but measured by several indicator variables. Lastly, SEM allows testing the indirect effects of multiple exogenous variables on the outcome variable “political use of search” by computing the product of all individual paths that constitute the mediation.

The exogenous variables included in the model are demographic characteristics (age, gender, race, education, income), political orientations (general political interest and ideology), Internet and media use (general Internet use, political media use, media trust), and search engine variable (self-rated search ability). Three endogenous variables are considered attitudinal outcomes, including feelings of efficacy when using search, knowledge of different factors influencing search results, and perceived accuracy of search results. Finally, the use of search engines for political purposes, i.e. the political use of search, is the endogenous outcome variable.

Measures

Exogenous variables

General Internet use. was the average of the frequency of Internet use for different purposes, such as to watch movies/TV, buy/order products or food, read/send/receive emails and calls/messages, look for news online, or to investigate topics of personal interest (M = 3.75, SD = 0.94, Cronbach alpha = .86).

Political interest. was measured by asking respondents how interested they are in politics (from 1: not at all interested to 5: extremely interested) (M = 2.81, SD = 0.91).

Political use of media. was the frequency of using the following sources as online sources of political information: social media sites, online-only news sites, online sites of print media, email, political websites, online video platform (M = 2.68, SD = 0.98, Cronbach alpha = .88).

Media trust. was respondents' rating of the reliability and accuracy of information found in traditional media sources, including newspapers, TV, radio, and online versions of those (M = 3.47, SD = 0.77, Cronbach alpha = .86).

Search ability. was the self-rated ability to use search engines from 1: bad to 5: excellent (M = 4.46, SD = 0.65).

Political ideology. was measured by respondents' self-placement on a scale from 1 being "very right-wing" to 7 "very left-wing" (M = 3.94, SD = 1.74). The question wording explained that socialist parties would be considered "left-wing" while conservative parties would be considered "right-wing". Another question asked with which political party the respondent identified with. From these two measures, two dummy variables were created to indicate Liberal Democrat (N = 363) and Conservative Republican (N = 388).

Demographic variables. age, gender (dummy variable), race (Caucasian = 1), education, income.

Endogenous variables

Search efficacy. Respondents were asked for their feelings when they search online for information about politics, in particular, to indicate their agreement with four items: *Impossible* ("There is so much information about politics online that it is IMPOSSIBLE to find what I am interested in"), *Helpless* ("I sometimes feel HELPLESS trying to find specific information about politics online"), *Frustrated* ("I often feel FRUSTRATED when searching online for political information"), *Lucky* ("If I am successful, it is probably only because I was lucky"). The factor structure underlying these four items was examined to determine if they could be combined to reduce the degrees of freedom in the SEM model.

Search results determinants. Six items asking respondents about the perceived influence of six factors on search engine results (location, past search history, optimization, advertising, relevance and site popularity) from having "no influence" to "strong influence". Similar to Search efficacy, exploratory factor analysis was performed first to examine the factor structure underlying these latent variables.

Perceived accuracy of search results. was measured by respondents' answers to how reliable and accurate they rate the information in search engine results (from 1 to 5) (M = 3.57, SD = .80).

Main dependent variable of interest

Political use of search. The average frequency of using search to find information about politics or current events, to look up facts, answer a factual question, check the accuracy of news or information (M = 3.51, SD = 0.88, Cronbach alpha = .80).

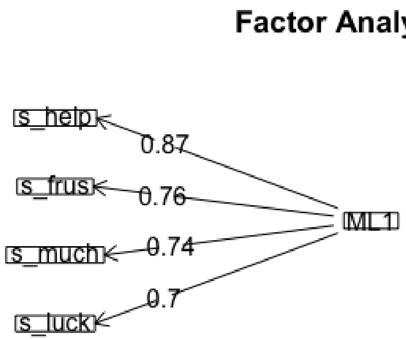
Analysis

Exploratory factor analysis with Search efficacy and Search results determinants. In order to test the underlying constructs in the measurements of Search efficacy and Search results determinants, and reduce the degrees of freedom that could potentially be lost, an exploratory factor analysis using principle axis factoring with promax rotation to extract factors was performed using the 'psych' package in R. The purpose of this procedure was to examine whether answers to multiple questions regarding Search efficacy and Search results determinants were correlated in such a way that suggested these two concepts had more than one factor in their factor structure. In the following, the results were presented first for Search efficacy, followed by Search results determinants.

* *Search Efficacy:*

First, the hypothesis of whether four measures of Search Efficacy belonged to one factor was tested. Parallel analysis and scree plot suggested that the number of factors underlying Search Efficacy was 1 (Figure 1). The Tucker Lewis Index of factoring reliability was 0.952 (indicating a good fit), and RMSEA was 0.118 with 90% CI [.093, .145], BIC value = 42.71 which indicated a poor fit. For the SEM model, the four measures were reverse coded and then averaged to create an index of Search Efficacy (Cronbach's alpha = .854).

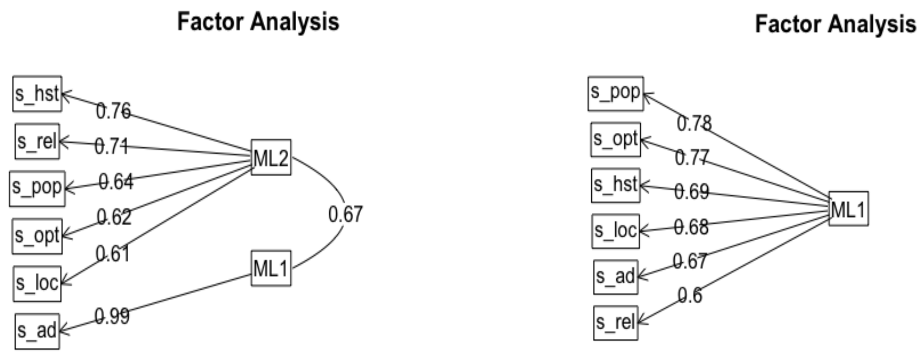
Figure 1: Factor analysis diagram of Efficacy, measured by 4 measures (helpless, frustrated, too much information, "successful only because I'm lucky")



* **Search results determinants***:*

Second, the potential underlying constructs in a set of six measurements of Search results determinants (perceived influence of location, search history, optimization, advertising, query relevance and site popularity) were examined. Scree plot analysis suggested that the number of factors was 1. The Tucker Lewis Index of factoring reliability was 0.956, and RMSEA was 0.081 with 90% CI [.063, .1], BIC value = 26.43 which indicated an adequate fit. The two-factor model was also run to test whether it would provide a better fit compared to the one-factor model. The BIC value of the two-factor model was bigger (BIC = 94.75) than the one-factor model (BIC = 26.43), the RMSEA index also increased (RMSEA = 0.092), suggesting the two-factor model was not any better. Both models were rerun with Maximum Likelihood estimation to account for the underlying distribution of the data. With the two-factor model (Figure 2), the standardized loadings based on the correlation matrix table showed that knowledge about the influence of advertising fees paid to the search engine loaded separately from knowledge of all other factors. The correlation between two factors was .67, Tucker Lewis Index of factoring reliability was 0.944, RMSEA 0.092, and BIC value 93.13. In comparison, the one-factor model had likelihood Chi-square value of 56.87 (df = 15, p < .000), Tucker Lewis Index of factoring reliability 0.958, RMSEA = 0.08, BIC = 24.83.

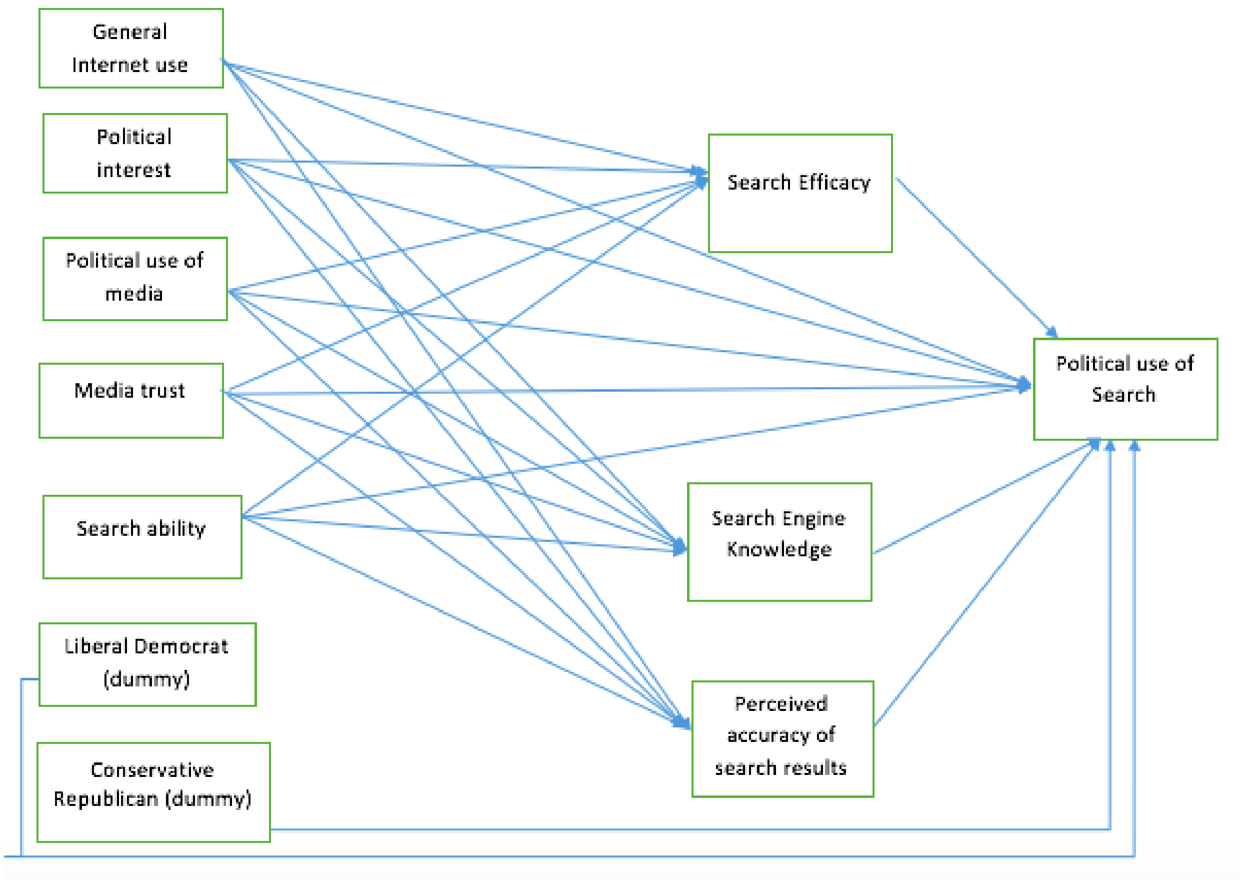
Figure 2. Loadings of the Search results determinants factor structure: two-factor model vs one-factor model



(Tucker Lewis Index of factoring reliability = 0.944, RMSEA = 0.092, BIC = 93.13).	Tucker Lewis Index of factoring reliability = 0.958, RMSEA = 0.08, BIC = 24.83).

Since the one-factor model indicated a better and logical fit with the data, I decided to go with the one-factor model for Search results determinants, and for the SEM model, the six measures were averaged to create an index of Search results determinants (Cronbach's alpha = .855).

Figure 3. Graphical display of the proposed SEM model

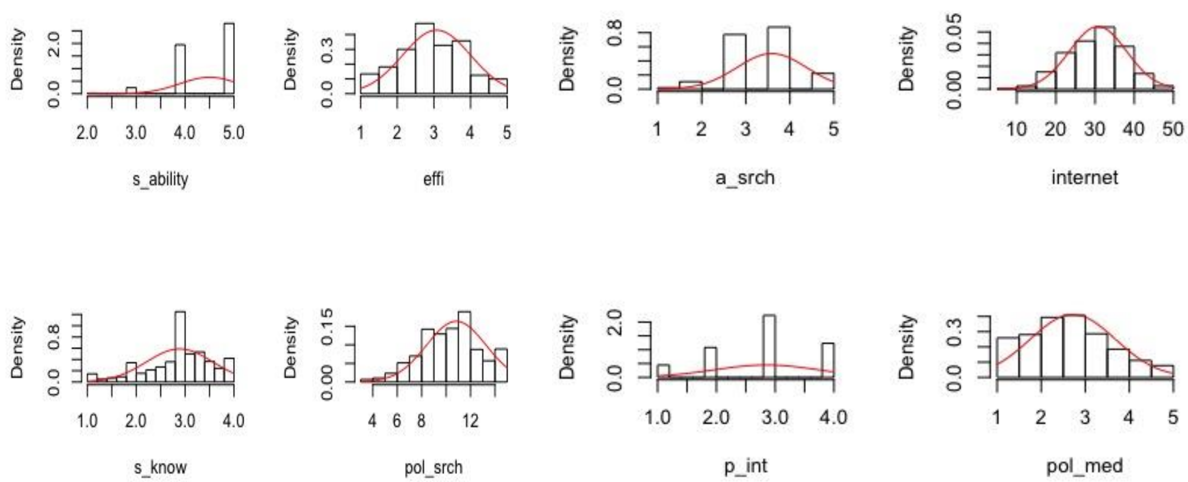


Note: For clarity purposes, the exogenous demographics variables (age, gender, race, education, income) were not included in the graphical display of the model. However, they were included in the specification of all regression paths predicting the endogenous variables (Search Efficacy, Search Engine Knowledge, Perceived accuracy of results, and Political use of search).

Checking assumptions of non-normality and missing data. Figure 3 presents the initial SEM model with a graphical display of the structural relationships between exogenous and endogenous variables. To examine whether the assumptions for structural equations modeling were met, diagnostics tests were performed to check for any serious deviations such as multivariate non-normality and missing data. Mardia's MVN test was used to calculate Mardia's multivariate skewness and kurtosis coefficients and their statistical significance. Both tests indicated non-normality, indicating that the data did not follow a multivariate normality distribution (Table 1). Univariate plots indicated that the Search ability variable was heavily skewed to the right.

Table 1. Results of Mardia's MVN test for multivariate non-normality

Test	Statistic	p-value	Result
Mardia Skewness	1404.79	1.390e-171	NO
Mardia Kurtosis	2.325	0.02004	NO
MVN	<NA>	<NA>	NO



With regard to missing data, Table 2 indicated some issues of missingness, especially with Search engine knowledge with 15.31% missing data.

Table 2. Descriptive statistics for all observed variables and missing data percentage

Variable	Mean	SD	min	max	% Missing
Search ability	4.46	0.65	1	5	1.88
Political interest	2.81	0.91	1	4	2.33
Perceived accuracy of results	3.57	0.8	1	5	3.42
Political use of media	2.68	0.98	1	5	3.57
Political use of search	3.51	0.88	1	5	3.57
General interest use	3.75	0.94	1	6	6.19
Search efficacy	3.1	0.92	1	5	7.48
Media trust	3.47	0.77	1	5	7.92

Search results determinants	2.87	0.7	1	4	15.31

To address the issues of missing data and multivariate non-normality, the SEM model was estimated with the "lavaan" package in R using robust maximum likelihood "MLR" as estimator to address non-normality and full information maximum likelihood "FIML" approach to address non-random missing data. In the model, all three variables Search Efficacy, Search results determinants, and Perceived accuracy were regressed on age, gender, race, education, income. In addition, Efficacy, Search results determinants and Perceived accuracy were regressed on five exogenous variables: general Internet use, political interest, political use of media, media trust, and search ability. Finally, the main outcome variable of interest Political use of search was regressed on three endogenous variables (Search efficacy, Search results determinants, perceived accuracy), all exogenous variables (demographics, five exogenous variables of interest, and two dummy variables indicating Liberal Democrats and Conservative Republicans).

Results

The model converged successfully after 105 iterations. The likelihood ratio chi-square which assessed the overall fit and the discrepancy between the sample and fitted covariance matrices indicated that the model did not fit perfectly (Chi-square = 34.40, df = 9, p = .000). Other fit indices however indicated an adequate fit (Robust CFI = .98, Robust TLI = .92, Robust RMSEA = .04 with 90% CI range .03 ~ .06, SRMR = .01).

The modification indices were then examined to identify potential misspecifications in the model. In particular, the post-hoc power (cut off value .80) was computed for each modification index, using delta value (effect size) of 0.1 and alpha value (significance level) of 0.05. Two paths with the biggest MI values and diagnosis of "misspecification" were: (1) search efficacy being regressed on the dummy variable Liberal Democrat (MI = 8.42), and (2) perceived accuracy of results being regressed on the dummy variable Conservative Republican (MI = 7.28). Model modification was then performed by adding these paths into the model step by step and observing the changes in the fit indices. The results showed that when both paths were added, the model performed slightly better with improved fit indices (Robust CFI = .99, Robust TLI = .95, Robust RMSEA = .03 with 90% CI range .02 ~ .05, SRMR = .008) and a slight decrease in the BIC value (from 14,896 to 14,888). The Chi-square statistic however still indicated the model did not fit perfectly (Chi-square = 18.88, df = 7, p = .009).

Table 3 presents the results from the lavaan package output of the effects of exogenous variables on each of the endogenous variables.

RQ1: What are the factors that influence attitudinal outcomes (search efficacy, search results determinants, perceived accuracy of search results) related to search engine use, and the use of search for political information (political use of search)?

First, age ($\beta = 0.009$), political interest ($\beta = 0.068$), search ability ($\beta = 0.299$), were found to have significantly positive effects on search efficacy, whereas higher level of political use of media ($\beta = -0.332$), and media trust ($\beta = -0.104$), were negatively correlated with search efficacy (H1b, d, e were supported). There were also strong main effects of identifying as liberal Democrats ($\beta = 0.163$) on efficacy, indicating high degree of confidence in using search engines among these individuals.

Turning to Search engine knowledge as the outcome, education ($\beta = 0.063$), Internet use ($\beta = 0.023$), political use of media ($\beta = 0.087$), political interest ($\beta = 0.052$), media trust ($\beta = 0.114$), and search ability ($\beta = 0.093$) were positively correlated with having higher knowledge of factors influencing search results (H2 was supported).

Regarding perceived accuracy of search results as the outcome, those with more frequent use of the Internet (β = 0.012), who possess more trust in the media (β = 0.461), rate themselves higher in search ability (β = 0.141) and Conservative Republicans in particular (β = 0.115) were more likely to perceive search results as reliable and accurate. Reversely, higher income (β = -0.018) were found to negatively relate to perceived accuracy of search results. H3a, c and H4b were supported.

Last, using search for political purposes were strongly correlated with Internet use (β = 0.095), political use of media (β = 0.591), political interest (β = 0.410), search ability (β = 0.494). The relationship between media trust and political use of search was not found to be significant. Of the three hypothesized mediators, only Search results determinants was shown to have a strong positive effect on the political use of search (β = 0.767), whereas the effects of search efficacy and perceived search accuracy were not statistically significant. The main effects of identifying as liberal Democrats and conservative Republicans were also not statistically significant.

RQ2. To what extent do attitudinal factors (search efficacy, search results determinants, perceived accuracy of search results) mediate the effects of political and informational factors on the political use of search?

Table 4 indicates the Indirect effects and Total effects of Internet use, Political interest, Political use of media, Media trust and Search ability on Political use of search through the three endogenous variables (Search Efficacy, Search results determinants and Perceived accuracy of search results). In general, it can be seen that Efficacy and Perceived accuracy were not the significant mediational paths through which different factors influenced the frequency of political search. In contrast, Search results determinants was found to significantly mediate the effects of all exogenous variables. In particular, the effects of general use of the Internet, political interest, political use of the media sources, and search ability on political use of search were partially mediated through knowledge of search engine result determinants. The effect of media trust on political use of search, in contrast, was fully mediated through search engine result determinants.

Table 3. Effects of Exogenous on Endogenous Variables (Efficacy, Search results determinants, Perceived Accuracy, Political Use of Search)

	Search Efficacy	Search Results Determinants	Perceived Accuracy of Search Results	Political Use of Search
Age	0.009***(0.002)	-0.001(0.001)	0.002(0.001)	0.007(0.004)
Female	0.018(0.045)	0.052(0.032)	-0.058(0.034)	0.173(0.097)
White	0.005(0.059)	0.047(0.042)	-0.085(0.048)	-0.030(0.125)
Education	0.054(0.028)	0.063**(0.021)	-0.006(0.023)	-0.096(0.065)
Income	-0.07(0.010)	-0.004(0.007)	-0.018*(0.008)	0.011(0.022)
General Internet use	-0.003(0.004)	0.023***(0.003)	0.012***(0.003)	0.095***(0.010)
Political use of media	-0.332***(0.036)	0.087***(0.025)	0.052(0.03)	0.591***(0.081)
Political interest	0.068*(0.030)	0.052*(0.024)	-0.004(0.024)	0.410***(0.068)
Media trust	-0.104**(0.035)	0.114***(0.025)	0.461***(0.028)	-0.114(0.082)
Search ability	0.299***(0.036)	0.093**(0.032)	0.141***(0.033)	0.494***(0.095)
Liberal Democrats (dummy)	0.163** (0.058)	-	-	0.200 (0.118)
Conservative Republicans (dummy)	-	-	0.115** (0.042)	-0.035 (0.124)
Search efficacy	-	-	-	0.028 (0.059)
Search engine knowledge	-	-	-	0.767*** (0.103)
Perceived accuracy of search results	-	-	-	0.137 (0.081)

Notes: Entries are regression estimates with standard errors in parentheses.

Table 4. Indirect effects and total effects

	Estimate	SE	Z-value	*p*
Indirect effect of INTERNET USE on POLITICAL SEARCH through Efficacy	-0.000	0.000	-0.383	0.702
….. through Search results determinants	0.017***	0.003	5.284	0.000
….. through Perceived accuracy of search results	0.002	0.001	1.536	0.125
Total effect	0.114***	0.011	10.751	0.000
Indirect effect of POL INTEREST on POLITICAL SEARCH through Efficacy	0.002	0.004	0.474	0.636
….. through Search results determinants	0.040*	0.019	2.103	0.036
….. through Perceived accuracy of search results	-0.001	0.003	-0.159	0.874
Total effect	0.452***	0.071	6.338	0.000
Indirect effect of POL USE OF MEDIA on POLITICAL SEARCH through Efficacy	-0.009	0.020	-0.482	0.630
…. through Search results determinants	0.067**	0.021	3.173	0.002
…. through Perceived accuracy of search results	0.007	0.006	1.171	0.242
Total effect	0.656***	0.081	8.110	0.000
Indirect effect of MEDIA TRUST on POLITICAL SEARCH through Efficacy	-0.003	0.006	-0.475	0.635
……. through Search results determinants	0.088***	0.023	3.789	0.000
……. through Perceived accuracy of search results	0.063	0.038	1.682	0.093
Total effect	0.034	0.080	0.425	0.670
Indirect effect of SEARCH ABILITY on POLITICAL SEARCH through Efficacy	0.009	0.018	0.480	0.631
……. through Search results determinants	0.071**	0.026	2.733	0.006
……. through Perceived accuracy of search results	0.019	0.012	1.569	0.117
Total effect	0.593***	0.094	6.302	0.000

***$p < .001$, **$p < .01$, *$p < .05$

Conclusion

Search engines are widely used for a variety of purposes including the monitoring and fact-checking of news and information. What remains unknown from past research, however, is the relative influence of factors crucial to the political contexts in which search engines are used (e.g. media trust, political identity), as well as how attitudes and perceptions underlying search behavior (e.g. perceived efficacy, search results determinants, perceived accuracy in search results) are shaped by external factors and how they, in turn, influence the political use of search. The purpose of this analysis is thus to examine individual differences in search perceptions and attitudinal outcomes which are important in understanding how and to what extent people rely on search engines for political information.

The findings indicate that the breadth and variety of Internet use, political media use and media trust are important determinants of search-related outcomes and political search behavior. Among these, the negative correlation between media trust and search efficacy is noteworthy. As previous findings on the linkage between media trust and selective (non)exposure to media sources suggest, it is possible that those with lower trust in traditional media sources did not feel less efficacious because they already relied on alternative online outlets (Kalogeropoulos, Suiter, Udris, Eisenegger, 2019; Mourão et al., 2018; Sterrett et al., 2019). However, as the findings in this study also imply, the more diverse the pool of online sources, the more likely individuals will experience some level of helplessness when confronted with the vast amount of information available.

The fact that knowledge of search engine is a strong predictor of using search for political information, and a significant mediator of the effects of Internet use, political interest, diverse use of political media, media trust and search ability, whereas no similar effects were observed for search efficacy and perceived search accuracy is also a noteworthy finding. This suggests that, in this study, algorithmic knowledge of search engines, conditioned by educational levels as well as individual attitudes and news habits, plays a key role in explaining the political use of search engines. In line with past research (Cotter & Reisdorf, 2020; Gran et al., 2020), this study shows that education is a key driver of digital divides, including divides in search engine knowledge. The findings here also show that learning about algorithms at work can also be improved via experience, training and consistent exposure, as indicated by the significant effects of general Internet use, ability and multiple media sources. In addition, to the degree that the awareness of how new media technologies work could influence technological adoption, addressing issues of algorithmic accountability and transparency would be important in influencing individuals' attitudes towards technologies. Thus, recent calls from scholars for the need to better understand the processes behind algorithms are highly justified (Diakopoulos, 2015; Diakopoulos & Koliska, 2017; Mustafaraj & Walsh, 2019). Klawitter & Hargittai (2018) showed that artists with "algorithmic skills" could utilize such know-how to improve their visibility and sell products on the online platform Etsy. The similarity of this example to search engine optimization and advertising suggests that knowledge inequities would make some individuals more equipped than others to make use of the systems.

Another important finding is that liberal Democrats and conservative Republicans in the sample felt strongly confident navigating online political information, with Conservative Republicans reporting confidence in the reliability of search results. These findings are in line with findings by a previous study by Wolak (2018), which found that strong partisans felt more politically efficacious in general, but they are also in contrast with recent reports on conservatives' distrust towards major tech platforms. More empirical research exploring this sentiment among political partisans would be needed to corroborate the findings here.

All in all, this analysis shed light on the conditions under which people search for political information, and the results highlight the role of algorithmic knowledge in explaining individual variations in political search. Because of the opacity of algorithms that prevents adequate understanding and critical assessment of online information flows, education and training efforts could perhaps be better directed at improving information literacy and critical thinking skills, both are crucial in the current digital environment.

Finally, the survey findings in this Chapter motivate further examination of the information-seeking tendencies among political partisans. Although political partisans might not differ in the *frequency* of using search engines for political information, *how* they seek information via search engines merits further attention. The next Chapter thus provides more insights into this inquiry by examining the relationship between political predispositions and information-seeking tendencies related to polarizing issues, as well as the role of partisan media consumption in this process.

CHAPTER II: POLITICAL PREDISPOSITIONS, PARTISAN MEDIA USE AND CONFIRMATORY SEARCH TENDENCY AMONG POLITICAL PARTISANS

Past communication research provided evidence that political partisans engage in partisan selective exposure when seeking information through traditional media (newspapers, broadcast TV) (Garrett & Stroud, 2014; Metzger et al., 2015; Stroud, 2017). In the online environment, however, very little support has been found regarding the similar phenomenon of selective exposure in social media or online news consumption habits. In the context of search engines, selective exposure is conceptualized as a type of confirmation bias, in which cognitive biases influence information-seeking behaviors, which in turn influence the type of information people are exposed to. Evidence from survey data shows a tendency towards confirmation bias in how people utilize search engines to seek general information about politics, as well as specific information about political candidates (Arendt & Fawzi, 2019; Mustafaraj et al., 2020; Nechushtai & Lewis, 2019; Whyte, 2016). Some scant evidence from ethnography works also shows that such tendency is particularly prominent among evangelist Conservatives who deeply distrust the news media (Tripodi, 2018). However, no systematic research to date has yet been conducted to examine the magnitude of confirmation bias among political partisans with regard to highly contentious issues.

Similarly, the connection between confirmation bias in Web search and individuals' surrounding information environment remains unclear. Do cues from media sources that individuals consume influence their tendency to search for information? Given what has already been known about partisan selective exposure, it is important to investigate how the use of slanted media sources, and exposure to biased language in such media, relate to the information-seeking tendencies among political partisans. Framing research has shown that media of different slants might have shared agendas but different narratives (e.g. Lee, McLeod, & Shah, 2008), and that people embedded in certain information environments could be more responsive to cues from such environments (Pierce, Redlawsk, & Cohen, 2017; Westerwick, Johnson, Knobloch-Westerwick, 2017). Thus, if it is true that partisans would search, or prefer search terms from their attitude-consistent media sources, the role of partisan media in biased information seeking would need to be understood and addressed. In extension, information system designs that take into account biased search behaviors and promote alternative information seeking would also be needed (Suzuki & Yamamoto, 2020).

The purpose of this Chapter is to tackle two broadly related research questions: 1) to what extent is confirmation bias present in Web search with regard to polarizing issues, and 2) to what extent do partisan media play a role in creating such tendencies. To answer these questions, an original survey was collected with a focus on self-identified political liberals and conservatives in the U.S. Confirmatory search tendencies were operationalized with two measures: biased query selection and ranking (where respondents were asked to rank terms extracted from partisan media as their preferred search queries to seek information about certain topics), and open-ended (where respondents were prompted to list their search terms of choice under accuracy or directional motivated reasoning goals). Altogether, the empirical analysis in this Chapter answered four concrete questions:

RQ1: How do political predispositions (measured as political ideology and issue position) relate to the preference for biased terms as search queries?

RQ2: How does the frequency of using partisan media relate to the preference for biased terms as search queries?

RQ3: To what extent the open-ended search terms provided by respondents reflect preexisting political predispositions?

RQ4: How do different information-seeking goals (accuracy vs directional) affect the open-ended search terms provided by respondents?

Confirmation Bias and Search Behavior

People rely on search engines for a variety of informational purposes, of both personal (e.g. health) and collective importance (e.g. participation in politics or social life). With search engines increasingly become an intermediary between citizens and political information (Newman et al., 2017), the role of search engines in determining access to information has important consequences. Academic scholars and the public alike often stress the necessity that citizens find valid and balanced information online to inform their decision-making process. However, inherent bias embedded in cognitive processing renders this normative expectation difficult to be fulfilled.

Confirmation bias is the human cognitive tendency to unconsciously prioritize information that supports one's opinion or hypothesis (Festinger, 1957). There has been a tradition of communication research that documents evidence of confirmation bias in information selection (Cardenal et al., 2019; Coppini et al., 2017; H. S. Kim et al., 2016; Weeks et al., 2017). Cognitive biases, including confirmation bias, are important topics to understand as they can influence the way people develop beliefs and behaviors based on false premises, such as preferring information consistent with their position regardless of factual correctness (Tversky & Kahneman, 1974). Examples include scenarios where a conclusive answer exists to questions like "Is climate change really happening?"; however, a person who does not believe in climate change might favor a different solution and exhibit a preference for disproving information as a result of their biases.

Confirmation bias with regard to online information seeking has been examined in several issues, including information about elections and political candidates (H. S. Kim et al., 2016; Suzuki & Yamamoto, 2020; Trielli & Diakopoulos, 2020; R. White, 2013). Limited cognitive capacity plays a key role in the decision-making process concerning search engine use. In fact, as online users are subject to time and cognitive constraints, a lot of search activities using online search tools are performed heuristically and without much thought given to the process (Agosto, 2001; Cothey, 2002; Wirth, Bocking, Karnowski, & Pape, 2007).

White (2013) predicts that bias in information retrieval happens when searchers seek or are presented with information that significantly deviates from the truth. The author examined search-related bias using three methods: survey, human labeling of results returned by a search engine, and log analysis of search behavior on the engine in the domain of health searches and yes/no questions. White showed that confirmatory search behaviors were evident in the way searchers retained a prior belief more strongly before searching and browsing activities. The preference for consistent information remained even after an answer was found and other options were presented. The most concerning implications of this study suggest that people would be led to incorrect results when there were a combination of systematic bias and searcher's bias.

Past research on individual tendencies to engage in mindful information seeking for credible and accurate information show that online search behaviors are contingent on several individual factors, including search skills (Hargittai & Dobransky, 2017), domain expertise (White, Dumais, & Teevan, 2009), Web experience (Howard & Massanari, 2007), specific individual characteristics and situational goals (Wirth et al., 2007), cognitive styles (Spink & Cole, 2006) and specific motivation (Edgerly et al., 2014; Yang & Zhuang, 2020). Similar to information seeking, attitudes towards verification strategies in Web search are correlated with education level, search skills, and need for cognition (Yamamoto, T., Yamamoto. Y., & Fujita, 2018).

Confirmatory search behavior is expected to be more pronounced among those with strong attitudes, and in the context of this study, politically motivated individuals, as they tend to have stronger opinions and more coherent views of the political world (Converse, 2007; Zaller, 1992), and rely more often on partisan cues in their attention to the surrounding information environment (Iyengar & Hahn, 2009).

In the context of heightened polarization, the literature on selective exposure posits that political partisans would selectively attend to information in accordance with their political bias. This hypothesis has not been tested extensively and systematically in the context of Web search. If the literature holds with regard to political search behaviors, we should see a direct relationship between political predispositions, issue positions and preference for biased terms as search queries. In other words, with regard to contentious political issues, political partisans in the U.S. will indicate a stronger preference for terms that align with their political predispositions.

Some past research found that biased tendencies might be stronger among political conservatives, especially when this group perpetually perceives the threat of biased mainstream media against conservative viewpoints (Barnidge & Rojas, 2014; Gunther, 1992). Conservatism itself is a form of motivated cognition associated with aversion to novelty (Jost & Amodio, 2012). Some experimental works show that conservatives tend to be less open to messages from the other side and less tolerant of embracing counter-attitudinal thoughts. For example, Lindner & Nosek (2009), in an experiment examining partisans' attitudes post 9/11, found that conservatives showed less tolerance of the counter-attitudinal "Americans are the problem" rhetoric compared to liberals of the "Arabs are the problem" rhetoric. Similarly, Nam, Jost, & van Bavel (2013) found that conservatives who supported Republican presidents in their study refused dissonance-arousing situations where they were made to write a counter-attitudinal essay about who made a "better president".

Based on previous research, I hypothesize that embracing a conservative ideology or advocating so-called "conservative" issue positions would positively correlate with the tendency to prefer conservative terms, and negatively correlate with the tendency to prefer liberal terms. The first hypothesis is posed:

H1: There is a positive correlation between conservative ideology (right-leaning issue position) and preference for conservative terms, and a negative correlation between conservative ideology (right-leaning issue position) and preference for liberal terms.

Partisan Media, Information Cues, and Potential for Biased Search

The need to satisfy the information demands of a niche audience has led to the proliferation of opinionated media sources with distinct partisan perspectives, evidenced by the rise of cable news in the early 2000s and the emergence of online-native hyper-partisan outlets in recent years (Guess, Nagler, & Tucker, 2019; Peacock, Hoewe, Panek, & Willis, 2021; Xu, Sang, & Kim, 2020).

Exposure to biased media framing can significantly alter individual judgments and decision making along with a host of other factors including identities, values and factual information. This has been consistently shown in field and survey experimental framing effects research (Adarves-Yorno et al., 2013; Clifford, 2019; Gallagher & Updegraff, 2012; Liu et al., 2019). In competitive framing environments, increased media choice can make framing effects less durable (Chong & Druckman, 2007); however, partisan polarization can significantly influence a frame's effectiveness as partisan cues are prominent in establishing credibility and appealing to individual values (Druckman, Peterson, & Slothuus, 2013).

Individuals (and voters) account for their knowledge gaps by utilizing heuristics to gain "competence" in important political decisions. These include taking information shortcuts by inferring from the party affiliation of the candidates, their poll standings, or endorsements in voting decisions (Lupia, 2006, 2016). As argued by scholars advocating such critical views, these strategies could correct for the low average of political information (Campbell, Converse, Miller, & Stokes, 1960) or inconsistent issue positions among the mass public (Converse, 2007). Psychological theories on human cognition that depict individuals as cognitive misers also acknowledge the merits of these strategies (Fiske & Taylor, 2013).

On the other hand, news media frames and cues can substantially influence citizens' political preferences (Busby, Flynn, & Druckman, 2018; Lau & Redlawsk, 2001; Shah, Watts, Domke, & Fan, 2002). With the rise of misinformation and polarizing content, taking cues from the surrounding environments also means that individuals are exposed to weaponized language and information with heavy political slant more than ever. This could in turn interfere severely in the information-seeking process. In particular, information searches performed by individuals could be nudged by various types of media (including social media, partisan outlets) which contain deliberately false or misleading narratives from willful political actors (Golebiewski & boyd, 2018).

Thus, based on previous evidence of partisan cue-taking from both observational and experimental studies, partisans' exposure to slanted media sources should be positively related to partisans' preference for biased terms prominently found in such media. In other words, the frequency of consuming right-leaning media is expected to correlate with favoring conservative terms as search queries and vice versa for liberal terms. The second hypothesis is formed:

H2: There is a positive correlation between partisan media consumption and preference for biased terms of the same slant as search queries.

Motivated Reasoning Goals and Search Behavior

Motivations to search can affect two kinds of cognitive processing of information: heuristic and systematic processing. Most past work on motivated reasoning focuses on how initial opinions with regard to a certain attitude object (e.g., an emergent technology, a political issue) shape subsequent attitudes and information-seeking behaviors (Bolsen & Druckman, 2015).

Information seeking is also dependent on the task at hand. Directional motivated reasoning is theorized as the tendency to place more weight in evidence that is consistent with prior opinions and dismiss those inconsistent with such, although the latter might be objectively accurate (Druckman et al., 2013; Lodge & Taber, 2013). With a directional goal, a person is "motivated to arrive at a particular opinion" (Kunda, 1990, p. 236). Directional motivated reasoning prevails in situations where individuals see minimal value in exerting cognitive resources to process convoluted, multi-faceted arguments (Nyhan & Reifler, 2010). This type of reasoning is directly in contrast with accuracy-motivated reasoning, which is the attempt to engage in information processing towards the goal of achieving "accurate" beliefs (Dawson et al., 2002; Epley & Gilovich, 2016; Rudolph, 2006).

An issue with previous research using log analysis of online search queries at the aggregate level is that there is no information about the different motivations behind search activities and the specific search queries used (Ripberger, 2011). People perform searches with concrete goals in mind under both cognitive and time constraints. Such situational goals have been shown to amplify or suppress information selectivity. For instance, Kim (2007) induced a preservation goal and an accuracy goal to examine patterns of online information seeking. The author showed that when put in a scenario where participants were made to justify their choice to select a fictional political candidate over the other (i.e. preservation goal), selective exposure to congruent information was stronger than when the goal was to accurately understand the two candidates (i.e. accuracy goal). Similarly, Edgerly et al. (2014) found that telling subjects they would later discuss with another person who disagreed or agreed with the subjects' stance on an issue significantly changed how these groups navigated and selected online information. Edgerly et al. (2014) further showed that a blank search interface led to less biased searches compared to an interface where competing partisan alternatives were presented. However, it was not clear if the neutralizing effect of the manipulated blank search interface was due to the induced suppression of cognitive dissonance, or the low skills and efforts of information searchers that resulted in less biased content (Hargittai, 2010; Howard & Massanari, 2007).

Returning to the context of this work, partisans, when being asked to seek information about polarizing issues for which they lean towards one side or another, can utilize different search strategies depending on the accuracy or directional motivated reasoning goals. By placing respondents in either goal, the impact of these goals on confirmatory search tendencies among political partisans can be examined. Similar to the ranking of predetermined biased terms as search queries, it is expected that individuals would report search queries of choice that are aligned with their preexisting political bias. In other words, search queries at the group level would indicate some differences when submitted by the two groups on the opposing sides of a polarizing issue. In addition, under the directional motivated reasoning goal condition, partisans on both sides would be more likely to come up with strong and emotion-laden words as queries compared to the accuracy goal condition. The following hypotheses are thus formed:

H3: Open-ended search terms provided by respondents will reflect preexisting issue positions.

H4: Under directional motivated reasoning goals, open-ended search terms provided by respondents would contain more sentiment than under accuracy goals.

Put together, the connections between political predispositions, partisan media exposure, and preference for biased terms are hypothesized with a mediation model specifying the potential mediating role of partisan media in the process. Tripodi (2018) put forward the possibility that biased cues from partisan media, which got picked up by frequent audiences, could reinforce confirmatory search tendencies. By the logic of selective exposure, it is possible that political partisans attentively select themselves into congenial media and adopt biased language from such media as cues to verify information. Thus, the fifth hypothesis is formed:

H5: Partisan media use mediates the relationship between political ideology and preference for biased terms.

Methods

Sample

Quota sampling was conducted to over-sample political liberals and conservatives among the American public. Data was collected from 562 respondents (including 185 students and 377 adults – who were recruited from Qualtrics panel and online participant panel Prolific). After excluding missing and incomplete cases, the analysis was conducted on 526 complete cases. This sample includes 280 self-identified liberals (53.2%), 223 conservatives (42.4%), and 23 moderates (4.4%). The survey was fielded from 7-20 October 2020 with Qualtrics panel and 11-

15 November 2020 on Prolific. The survey was conducted after receiving approval from the IRB at the UW-Madison.

Extraction of biased phrases from partisan media

As one of the research questions of interest is concerning whether political partisans would prefer biased language from partisan media as search queries when being asked to perform searches under different goals, an initial step was to identify these biased phrases from partisan media outlets to sample them as potential search queries. To do this, I collected media coverage on four political issues to extract terms (neutral and partisan) that can be used as search queries. More information on the methodology for this part is included in the Appendix.

Measures

Political ideology: Respondents were asked to self-report their political ideology, from Very liberal (1) to Very conservative (7) (M= 3.55; SD = 2.04).

Issue position:

Abortion: Respondents were asked to indicate on a scale from 1 being "Pro-choice" to 10 being "Pro-life" where they stand on the abortion issue (M = 4.61; SD = 3.92).

Climate change: Respondents were asked to indicate where they stand on the climate debate issue, on a scale from 1 being "Climate change is just part of the natural cycle. Human action has little to do with it" to 10 being "Climate change has become more serious in recent years. Human action causes it and makes it worse" (M = 6.86; SD = 3.63).

Gun control: Respondents were asked how much they agree or disagree with the statement "Gun have no place in a civilized society" on a 7-scale from Strongly Disagree to Strongly Agree. The item was then reverse coded so that a higher value indicates a gun-support POV (M = 3.39; SD = 2.08).

Immigration: Respondents were asked how much they agree or disagree with the statement "Immigration to the US should be limited far more strictly than it currently is" on a 7-scale from Strongly Disagree to Strongly Agree. Higher value indicates a conservative POV on immigration (M = 3.82; SD = 2.16).

Partisan search terms selection:

For each of the four issues, the respondents were presented with 10 phrases associated with the issue taken from partisan media coverage (as described earlier) (4 conservative, 4 liberal, 2 neutral) in a randomized order. These terms were introduced in the question wording as "some phrases and language in the debate over the issue". Respondents were then instructed to select as many of these phrases as they would use to search for information about this topic on a search engine. For the terms they selected, they would rank them in the order from the one they would most likely use to the one least likely use. Using this method, each term was associated with a score from 1 to 10. A higher score indicated a higher likelihood of being selected as a search term. The full descriptive statistics of these terms can be found in the Appendix.

Open-ended search terms:

In order to examine how information-seeking goals might affect the search terms people intend to use to search for information, an experimental manipulation was included at the end of the survey, in which the respondents were assigned to see either a prompt for "Accuracy goal" or "Directional goal" for one of the four issues (abortion, climate change, gun control, or immigration).

The prompt was: "*We are trying to understand how people seek information regarding current news events. If you were to conduct an online search to get information about [issue], what would be the phrases or questions you would use to find accurate and objective* *information?"* (for Accuracy goals), and "*what would be the phrases or questions you would use to find strong and convincing information to support your opinion?"* (for Directional goals). Three search phrases or questions were requested in each prompt.

Control demographic variables: age (M = 36; SD = 17.59), gender (male 38.6%; female 60.3%; self-identified 1.1%), education (M = 4.22; SD = 1.44), income (M = 4.59; SD = 2.36).

Media use and exposure to partisan media:

General media use is the average of the self-reported frequency of using print media (newspapers or magazines), news programs on TV, online news on a 6-point scale, from "Less than once a week" to "Several times a day" (M = 2.51, SD = 1.14).

Conservative media use is the average of the self-reported frequency of using Fox News, conservative talk radio (e.g. Rush Limbaugh), conservative political blogs online (e.g. Hot Air, RedState) on a 6-point scale, from "Less than once a week" to "Several times a day" (M = 2.06; SD = 1.11).

Liberal media use is the average of the self-reported frequency of using MSNBC, CNN news, liberal political blogs online (e.g. Daily Kos, Talking Points) on a 6-point scale, from "Less than once a week" to "Several times a day" (M = 2.00; SD = .96).

Respondents were also asked to self-report their exposure to hyper-partisan online media outlets (i.e. how often they use those sources or have come across news or information from those sources on social media) on a 5-point scale (from Never to Very often). Those include hyper-partisan conservative sources (Breitbart, Daily Caller, Info Wars, Washington Times, Front Page Mag) (M = 1.68; SD = .89), and hyper-partisan liberal sources (Raw Story, Alternet, Truth-Out) (M = 1.38; SD = .89), which were identified in previous research as top-performing "junk news and disinformation domains" online (Taylor et al., 2020).

Political interest: is measured by averaging two questions: asking respondents to indicate how closely they follow what's going on in government and public affairs, and how closely they follow about the presidential election in November 2020 on a 5-scaled range from 1 (not at all closely) to 5 (very closely) (M = 3.63; SD = 1.07).

Analysis strategy

Partial Correlation coefficients were calculated to describe the relationship between conservative ideology (or right-leaning issue position) and preference for biased terms, and the relationship between partisan media use and preference for biased terms. All partial correlation models between conservative ideology (or right-leaning issue position) and preference for biased terms controlled for demographics (age, gender, education, income) and political interest. And all models between partisan media use and preference for biased terms controlled for demographics, political interest, and general media use. It is expected that conservative ideology (or a right-leaning issue position) would be positively correlated with a higher preference for conservative terms (and negatively correlated with a preference for liberal terms). Similarly, the correlations between frequency of using conservative media and higher preference for conservative terms are expected to be positive, and negative for liberal media/liberal terms preference. Partial correlation statistical procedure assumes that each pair of variables is bivariate normal. To control for false discovery rate due to multiple comparisons, p-values were adjusted using BH correction (Benjamini & Hochberg, 1995).

Formal tests of statistical mediation were conducted to examine the relationships between political ideology, the use of partisan media, and the selection of biased terms. Conservative media use and liberal media use were entered as parallel mediators. All mediation analyses were conducted using PROCESS Macro (Hayes, 2018), with percentile bootstrap confidence interval for indirect effects. Finally, text-mining techniques using R were used to examine open-ended search queries reported by survey participants.

Results

Hypothesis 1: There is a positive correlation between conservative ideology (right-leaning issue position) and preference for conservative terms and a negative correlation between conservative ideology (right-leaning issue position) and preference for liberal terms.

Table 5. Partial Correlations between Pro-life position/Conservatism and preference for biased terms - Abortion

		Pro-life position	Conservative ideology
Conservative terms	Abortion on demand	.152	.262
	Live birth abortion	.241	.296
	Pro life	.345***	.357***
	Late term abortion	.127	.163
	Conservative terms (avg)	.860***	.890***
Neutral terms	Unintended pregnancy	.067	.079
	Fetal abnormality	-.056	-.034
	Neutral terms (avg)	-.220	-.063
Liberal terms	Elective abortion	.022	.024
	Pro choice	-.123	-.134
	Women’s rights	-.106	-.085
	Reproductive freedom	.012	-.086
	Liberal terms (avg)	-.196	-.423

Note: Pro-life position was measured by the question “Where do you stand on the issue of “abortion”, with higher value indicating “I am Pro Life” position. Conservative ideology was measured based on self-reported political ideology, with higher value indicating conservative ideology.

Table 6. Partial Correlations between "Natural" climate position/Conservatism and preference for biased terms – Climate change

	Nature-cause climate change position	Conservative ideology
Climate hoax	.289**	.369**

Conservative terms	Climate fraud	.393**	.287*
	Climate hysteria	.062	.418
	Climate change agenda	.201*	.187*
	Conservative terms (avg)	.448	.711
Neutral terms	Climate change consensus	-.090	-.056
	Climate change impact	-.118	-.128*
	Neutral terms (avg)	-.161	-.257*
Liberal terms	Climate crisis	-.209**	-.178*
	Climate deniers	-.045	.138
	Climate change skeptics	.143	.226
	Carbon footprint	-.013	.033
	Liberal terms (avg)	-.340	.049

Note: Nature-cause climate change position was measured by the question "Where do you stand on the issue of "climate change", with higher value indicating placing oneself closer to the statement "Climate change is natural. Human action has little to do with climate change".

Table 7. Partial Correlations between "Gun rights" position/Conservatism and preference for biased terms – Gun control

		Gun-support POV	Conservative ideology
Conservative terms	Armed self defense	.277**	.272**
	Second Amendment	.165*	.223**
	NRA	.088	.235**
	Anti-gun agenda	-.008	.134
	Conservative terms (avg)	.588*	.665**
Neutral terms	Gun accessibility	-.074	-.033
	Terror threats	-.118	-.078
	Neutral terms (avg)	-.289	-.157
Liberal terms	Gun license	.039	.075
	Gun control solutions	-.134	-.144
	Background checks	-.088	-.042
	Gun lobby	-.024	.018
	Liberal terms (avg)	-.015	-.055

Note: Gun-support POV was measured by the reversed score for the statement "Guns have no place in a civilized society" with higher value indicating gun-support POV.

Table 8. Partial Correlations between Anti-immigration position/Conservatism and preference for biased terms - Immigration

		Anti-immigration	Conservative ideology
Conservative terms	Illegal Aliens	.252*	.312***
	Homeland Borders	.117	.202
	Radical Islam	.049	.089
	Alien Invaders	-.202	.213
	Conservative terms (avg)	.094	.041
Neutral terms	Comprehensive Immigration Reform	-.105	-.057
	Refugee Admission Limits	.055	-.059
	Neutral terms (avg)	-.138	-.025
Liberal terms	Refugee Asylum	-.124	.073
	DACA Amnesty	-.105	-.074
	Undocumented Immigrants	.065	.084
	Family Detention	-.137	-.070
	Liberal terms (avg)	-.589***	-.477**

Note: Anti-immigration stance was measured by the agreement level concerning the statement "Immigration to the US should be limited far more strictly than it currently is" with higher value indicating immigration opposing POV.

Correlations between political ideology and term preference showed a pattern that a higher score of conservatism was associated with higher preference for conservative terms, including "pro-life", "climate hoax", "climate fraud", "climate change agenda", "armed self-defense", "second Amendment", "NRA" and "illegal aliens". Higher conservatism was also associated with lower preference for only a few liberal and neutral terms, including "climate change impact", and "climate crisis". Climate change had the most terms with significant results, followed by gun control, abortion, and immigration.

Similarly, issue positions that aligned with conservatism were positively correlated with selecting conservative-leaning terms, mirroring patterns of conservatism. However, once controlling for the effect of ideology, such relationships between issue positions and biased preferences were weakened or in some cases eliminated. H1 found varying levels of support for specific terms and specific issues.

Hypothesis 2: There is a positive correlation between partisan media consumption and preference for biased terms of the same slant as search queries.

Table 9. Partial Correlations between Partisan media use and preference for biased terms - Abortion

		Conservative media use	Liberal media use
Conservative terms	Abortion on demand	.127	-.130
	Live birth abortion	.180	.079
	Pro life	.204*	-.179
	Late term abortion	.158	-.120
	Conservative terms (avg)	.699*	-.785**
Neutral terms	Unintended pregnancy	.056	.085
	Fetal abnormality	.005	.075
	Neutral terms (avg)	.041	.156
Liberal terms	Elective abortion	.121	.050
	Pro choice	-.062	-.027
	Women's rights	-.071	-.083
	Reproductive freedom	.088	.135
	Liberal terms (avg)	-.144	.188

.

Table 10. Partial Correlations between Partisan media use and preference for biased terms - Climate change

		Conservative media use	Liberal media use
Conservative terms	Climate hoax	.336	-.055
	Climate fraud	.276	-.109
	Climate hysteria	.421	-.007
	Climate change agenda	.163	-.035
	Conservative terms (avg)	.601	-.413
Neutral terms	Climate change consensus	-.119	.076
	Climate change impact	-.098	.053
	Neutral terms (avg)	-.340*	.084
	Climate crisis	-.119	.079
	Climate deniers	.297	.114
	Climate change skeptics	.253	-.026

Liberal terms	Carbon footprint	-.002	.114
	Liberal terms (avg)	.264	.047

Table 11. Partial Correlations between Partisan media use and preference for biased terms - Gun control

		Conservative media use	Liberal media use
Conservative terms	Armed self defense	.210	-.105
	Second Amendment	.172	-.106
	NRA	.169	-.012
	Anti-gun agenda	.061	-.018
	Conservative terms (avg)	.480	-.407
Neutral terms	Gun accessibility	.025	.020
	Terror threats	.144	.112
	Neutral terms (avg)	.199	.166
Liberal terms	Gun license	.086	-.002
	Gun control solutions	-.110	.012
	Background checks	-.119*	.053
	Gun lobby	.069	.026
	Liberal terms (avg)	-.008	.151

Table 12. Partial Correlations between Partisan media use and preference for biased terms - Immigration

		Conservative media use	Liberal media use
Conservative terms	Illegal Aliens	.140	-.213
	Homeland Borders	.027	-.146
	Radical Islam	.161	-.063
	Alien Invaders	-.553*	-.242
	Conservative terms (avg)	.007	-.080
Neutral terms	Comprehensive Immigration Reform	-.120	.095
	Refugee Admission Limits	.062	.043
	Neutral terms (avg)	-.026	.148
	Refugee Asylum	-.024	.031
	DACA Amnesty	-.182	.039
	Undocumented Immigrants	-.001	-.065

Liberal terms	Family Detention	-.116	-.093
	Liberal terms (avg)	-.449*	.019

To examine H2, partial correlation models were run between the frequency of consuming conservative/liberal media and ranking of biased terms, with all models controlled for demographics, political interest, general media use. The results show that the frequency of using conservative media was associated with higher preference for only a few terms, including "pro-life" (abortion issue), and lower preference for "background checks" (gun issue) and "alien invaders" (immigration issue). The relationships between liberal media use and all biased terms were not significant. Thus, there was not enough evidence to support H2 which hypothesized a positive correlation between partisan media consumption and preference for biased terms of the same slant as search queries.

Mediation results

Figure 4. Graphical presentation of the hypothesized mediation model

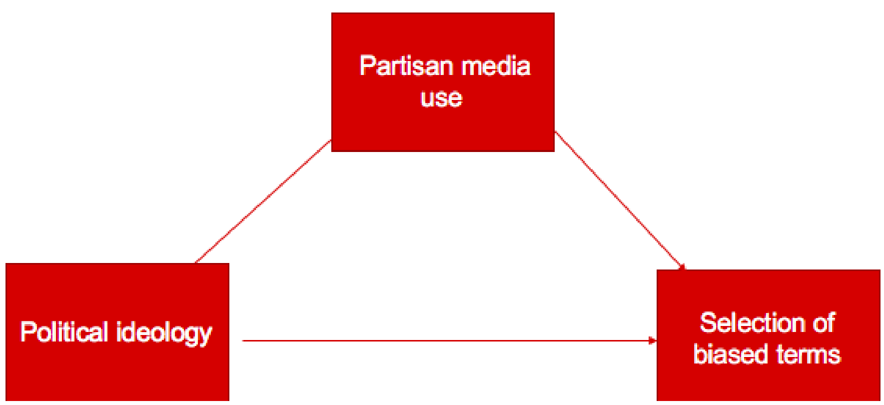


Table 13. Direct and indirect effects of Conservatism on preference for biased terms - Abortion

	Conservative terms selection		Liberal terms selection	
	Point estimate	Bias-corrected Bootstrapped 95% C.I.	Point estimate	Bias-corrected Bootstrapped 95% C.I.

	b	Lower Limit	Upper Limit	b	Lower Limit	Upper Limit
Direct effects of conservative ideology	.654**	.285	1.023	-.185	-.569	.199
Indirect effects of conservative ideology through:						
Conservative media	.051	-.197	.324	-.069	-.272	.167
Liberal media	.157	-.333	.545	-.071	-.084	.232

Table 14. Direct and indirect effects of Conservatism on preference for biased terms - Climate change

	Conservative terms selection			Liberal terms selection		
	Point estimate	Bias-corrected Bootstrapped 95% C.I.		Point estimate	Bias-corrected Bootstrapped 95% C.I.	
	b	Lower Limit	Upper Limit	b	Lower Limit	Upper Limit
Direct effects of conservative ideology	.955*	.039	1.872	-.446	-1.041	.149
Indirect effects of conservative ideology through:						
Conservative media	.029	-.480	.542	.173	-.570	.864
Liberal media	-.050	-.690	450	.050	-.474	.409

Table 15. Direct and indirect effects of Conservatism on preference for biased terms - Gun control

	Conservative terms selection			Liberal terms selection		
	Point estimate	Bias-corrected Bootstrapped 95% C.I.		Point estimate	Bias-corrected Bootstrapped 95% C.I.	
	b	Lower Limit	Upper Limit	b	Lower Limit	Upper Limit
Direct effects of conservative ideology	.600*	.161	1.039	.032	-.243	.308
Indirect effects of conservative ideology through:						
Conservative media	.066	-.287	.438	-.001	-.149	.135
Liberal media	-.110	-.542	.226	-.052	-.205	.047

Table 16. Direct and indirect effects of Conservatism on preference for biased terms - Immigration

	Conservative terms selection			Liberal terms selection		
	Point estimate	Bias-corrected Bootstrapped 95% C.I.		Point estimate	Bias-corrected Bootstrapped 95% C.I.	
	b	Lower Limit	Upper Limit	b	Lower Limit	Upper Limit
Direct effects of conservative ideology	.459	-.820	1.739	-.235	-.472	-.002
Indirect effects of conservative ideology through:						
Conservative media	.266	-.375	.225	-.140	-.369	.033
Liberal media	-.005	-.625	.150	.001	-.051	.068

Mediation analysis was run to examine how ideology affects conservative and liberal media use, which in turn affects biased term preference. The results in Table 13-16 indicate that – except for the issue of Immigration - ideology was significantly associated with biased term preference. However, there was no evidence of indirect effects through conservative and liberal media use. In other words, the use of partisan media was not the route through which ideology influenced the preference for biased terms.

Hypothesis 3: Open-ended search terms provided by respondents would reflect preexisting issue positions.

Confirmatory search tendencies, measured as open-ended self-reported search queries from the survey participants, suggest that at the group level, search queries submitted by the two groups on the opposing sides of polarized issues would be different to a certain degree.

Respondents' self-reported search terms were merged into groups of queries based on their issue positions. Text mining techniques were then used to compare the text corpus comprised of search queries from the two groups. The corpus was tokenized into unigrams (one-words) and bigrams (two-word phrases) and the frequency for each n-gram was calculated.

The following graphical presentations illustrate the words/phrases with a higher probability of being used by one group more than the other (log-likelihood ratio tests were used to determine how an n-gram was more or less likely to come from either group). The results showed some evidence that open-ended search queries reflected preexisting issue positions. For example, Figure 5 indicated that words like “heartbeat”, “fetus”, “procedure”, “killing” were relatively more overused in the pro-life group (n = 121 search phrases), whereas words like “rights”, “laws”, “reproductive”, and “roe v wade” were found more in the pro-choice group (n = 118 search phrases). Regarding climate change, Figure 6 showed that “hoax”, “real”, “fraud”, “agenda” were more prominent in the climate skeptics group (n = 90 search phrases), whereas “global warming”, “effects”, “science”, “environment” in the climate believers group (n = 110 search phrases). Turning to gun issues, Figure 7 showed that the pro-gun groups (n = 175 phrases) had proportionally more words like “NRA”, “gun laws”, “firearms”, “gun restrictions” compared to the anti-gun groups (n = 98 phrases) with words like “gun violence”, “gun reform”, “anti-gun”, “gun lobby”. Finally, regarding immigration issue, words like “illegal”, “immigration laws”, “Fox News”, “Trump” were more proportionally overused in the anti-immigration group (n = 145 phrases) compared to words like “camps”, “DACA”, “ICE” and “detention centers” in the pro-immigration group (n = 115 phrases) (Figure 8).

Figure 5. N-gram likelihood differences in the open-ended search queries reported by Pro-life & Pro-choice groups

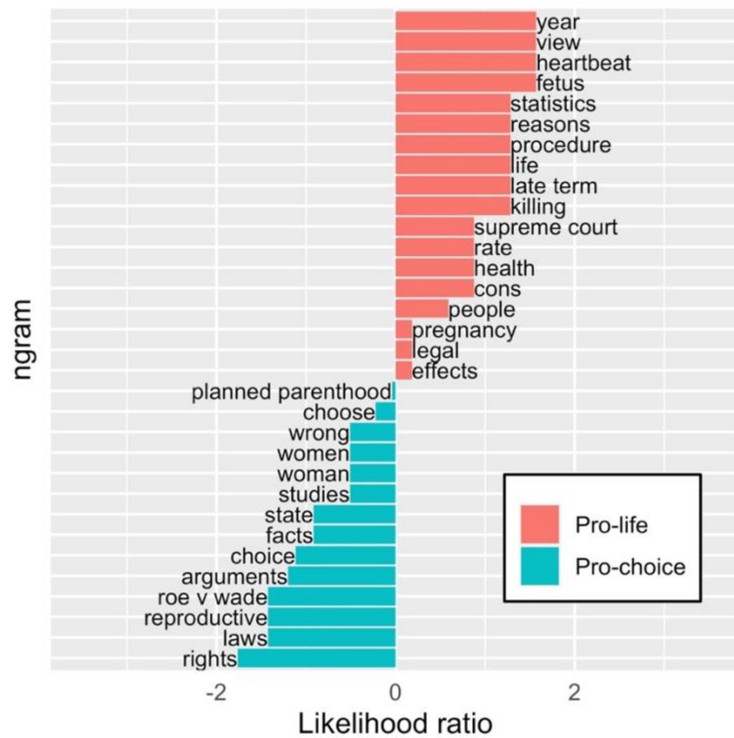


Figure 6. N-gram likelihood differences in the open-ended search queries reported by Climate change skeptics & Climate change believers' groups

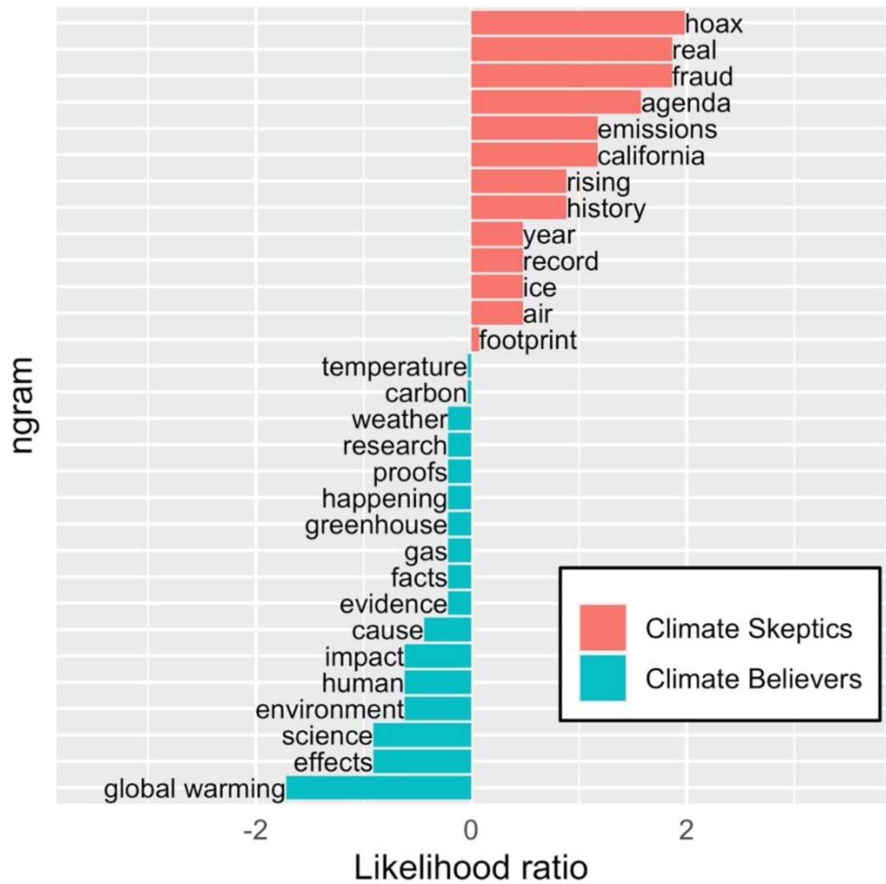


Figure 7. N-gram likelihood differences in the open-ended search queries reported by Pro- & Anti-Gun rights groups

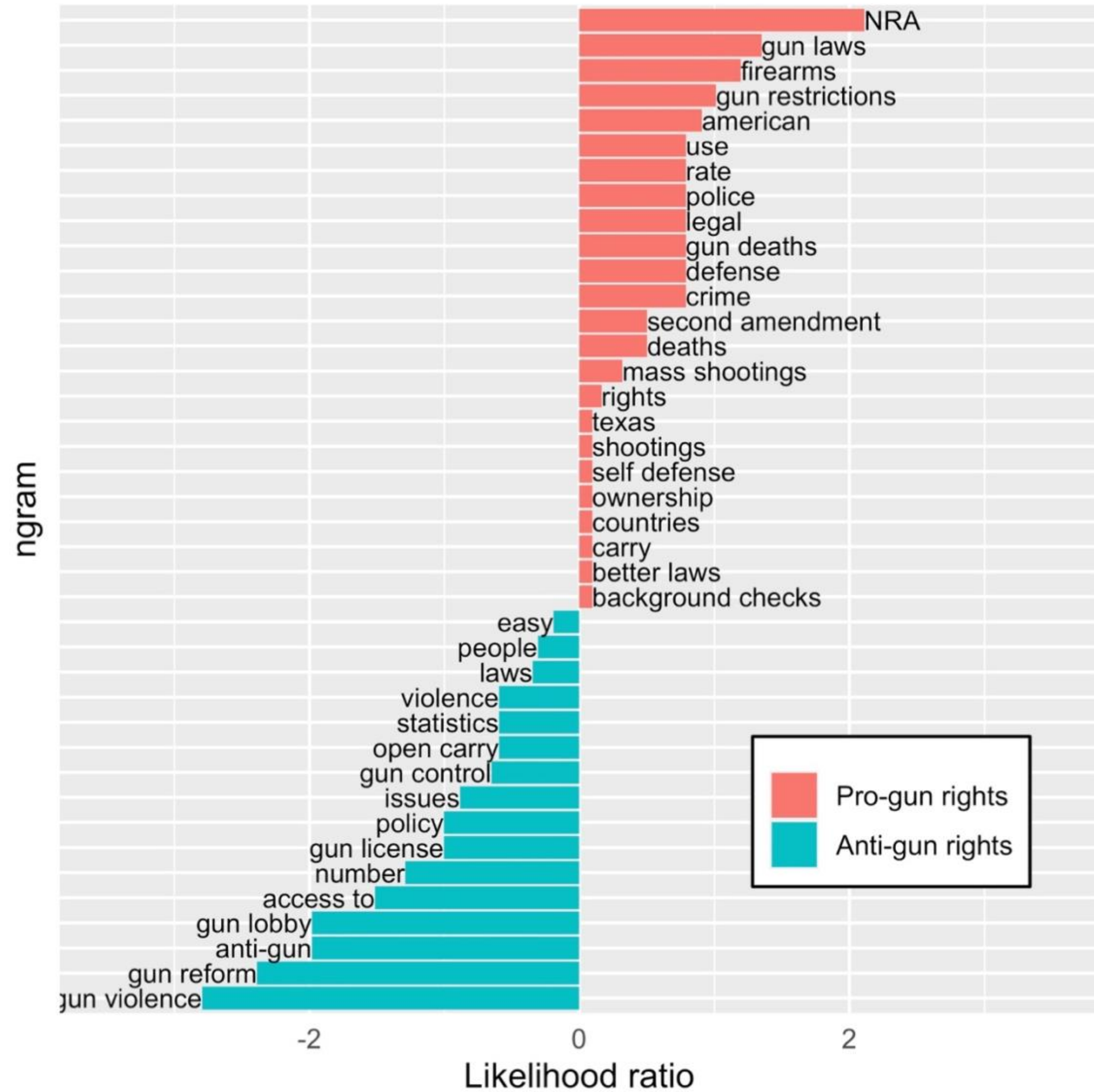
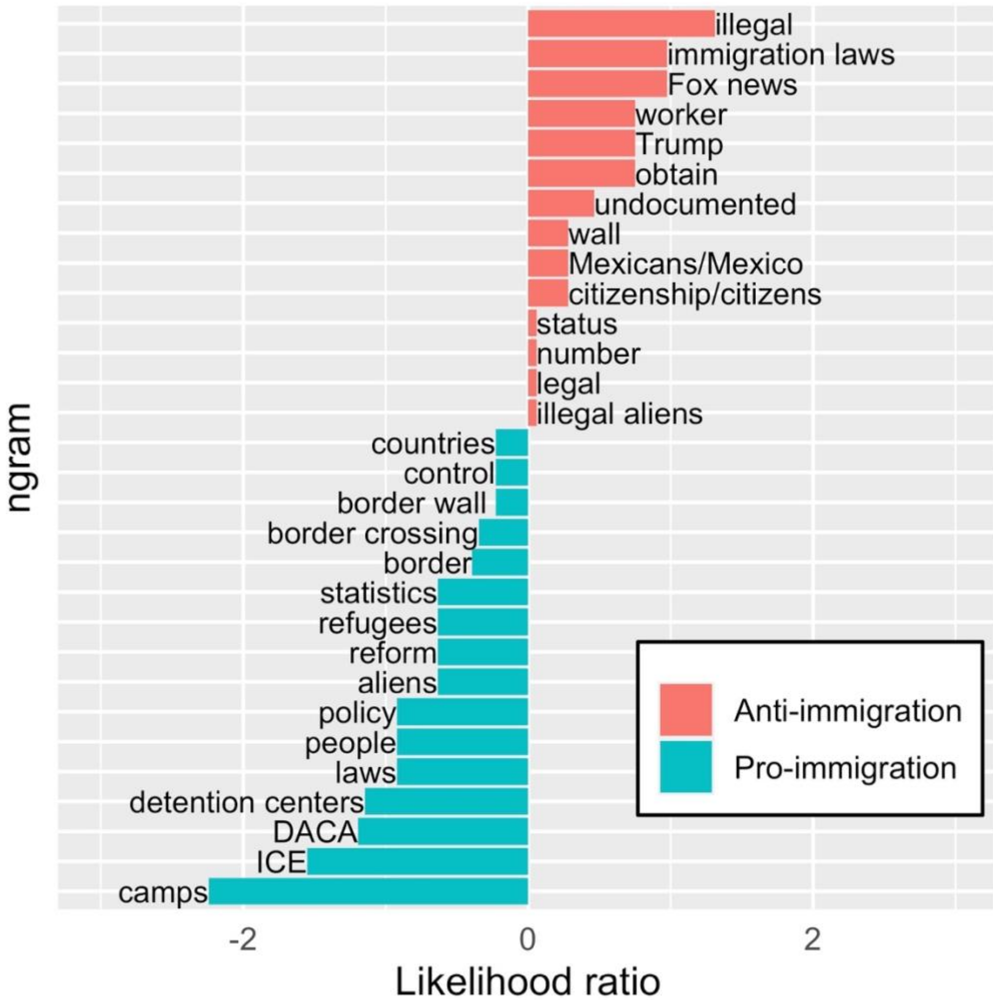


Figure 8. N-gram likelihood differences in the open-ended search queries reported by Anti & Pro-immigration groups



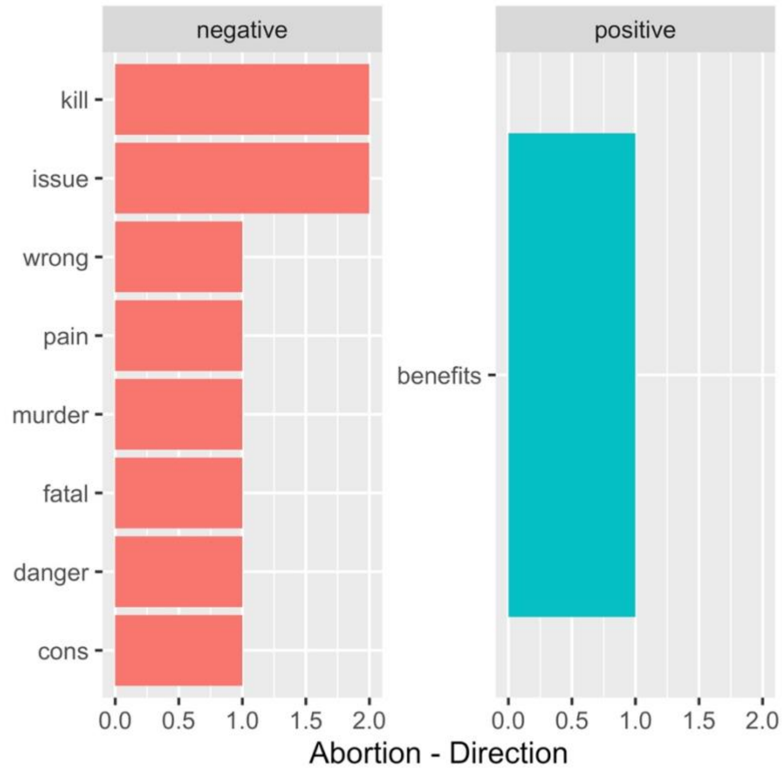
Hypothesis 4: Under directional motivated reasoning goals, open-ended search terms provided by respondents would contain more sentiment than under accuracy goals.

The respondents were also asked to provide their own search terms regarding each issue using two different prompts. The first prompt asked the respondents to list search terms that they think would help them find *accurate* information (this prompt represented the "Accuracy" goal), and the second prompt asked for search terms that respondents think would help them find *strong and convincing arguments to support their position* (this prompt represented the "Direction" goal).

For each issue, sentiment analysis was conducted on the two text corpora from the two goals (Direction versus Accuracy). The tidy-text R package provided access to three general-purpose lexicons for opinion and sentiment analysis, which contain English words with assigned scores for positive/negative sentiment and emotions. Here, the "Bing" lexicon was used. This lexicon, constructed by mining customer reviews of e-commerce products, identifying positive or negative opinions through summarization task (M. Hu & Liu, 2004), contained 6,786 words classified as either "positive" or "negative". An important note is that since the lexicon is based on unigrams, or single words, qualifiers before a word such as "not" or "no" were not taken into consideration.

The number of positive and negative words in Directional goals versus Accuracy goals for each issue was counted to compare how sentiments detected from the text were different by Goals. It is expected that when respondents were primed towards the Directional goals, the open-ended answers induced from this condition would contain more sentiment compared to the Accuracy goals.

Figure 9. Words indicating negative and positive sentiments found in Direction vs Accuracy goal condition - Abortion


negative
positive
kill
issue
wrong
pain
murder
fatal
danger
cons
benefits
0.0
0.5
1.0
1.5
2.0
Abortion - Direction

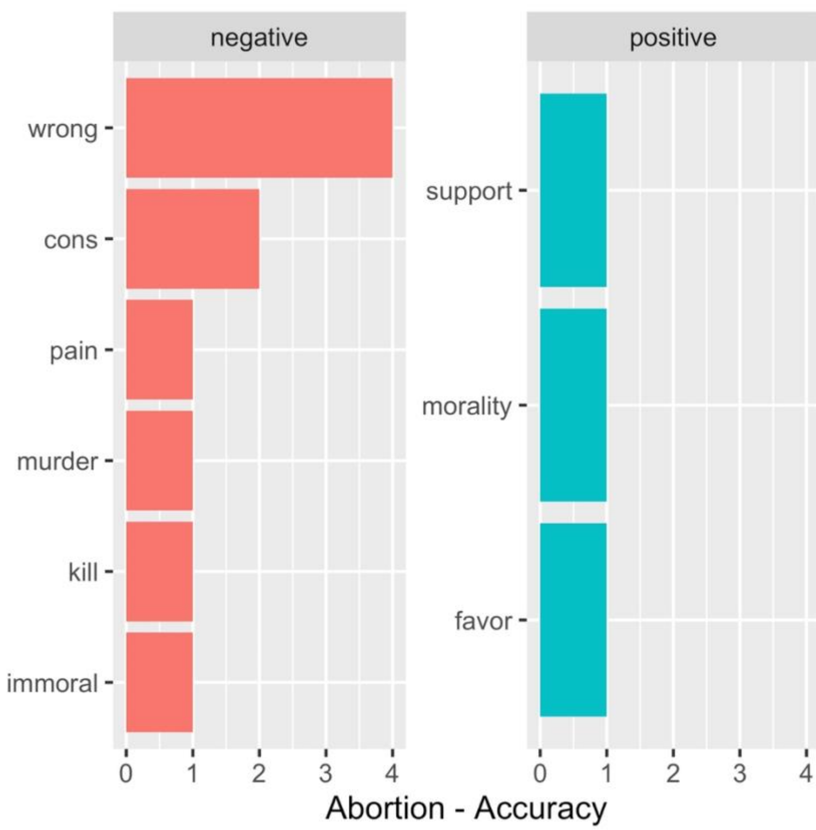
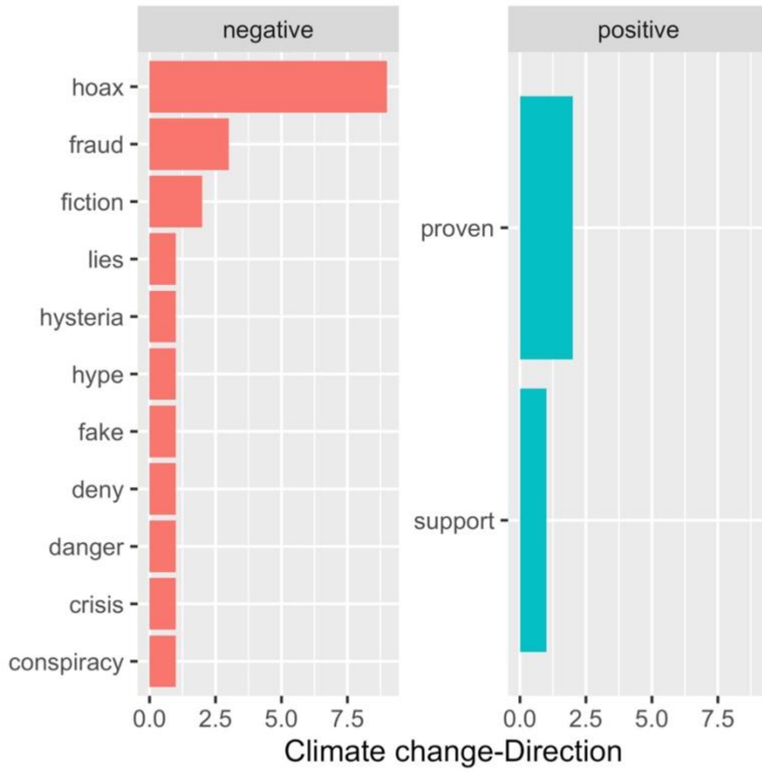

negative
positive
wrong
cons
pain
murder
kill
immoral
support
morality
favor
0
1
2
3
4
Abortion - Accuracy

Figure 9 illustrate the top words that contributed the most sentiment (by their frequency of occurrence) in each goal condition for the Abortion issue. More specifically, in the open-ended search queries regarding Abortion in the Directional condition, eight words were identified as “negative” and one word (“benefits”) was identified as “positive”. The word “kill” and “issue” appeared two times in the corpus, and other negative words (“wrong”, “pain”, “murder”, “fatal”, “danger” and “cons”) appeared once in the corpus. Turning to the Accuracy condition of abortion-related search queries, six negative words were detected (“wrong” with four counts, “cons” with two counts; “pain”, “murder”, “kill” and “immoral” each with one count), and three positive words were detected (“support”, “morality”, “favor”). The sentiment score for Directional vs Accuracy condition could be calculated as the sum of negative words (with a minus sign) and positive words (with a positive sign). Thus, the score of Directional condition was -9 compared to the score of Accuracy goal of -7. This indicates that search queries related to abortion generated in the Directional condition were more negative.

Figure 10. Words indicating negative and positive sentiments found in Direction vs Accuracy goal condition - Climate change


negative
positive
hoax
fraud
fiction
lies
hysteria
hype
fake
deny
danger
crisis
conspiracy
proven
support
0.0 2.5 5.0 7.5
0.0 2.5 5.0 7.5
Climate change-Direction

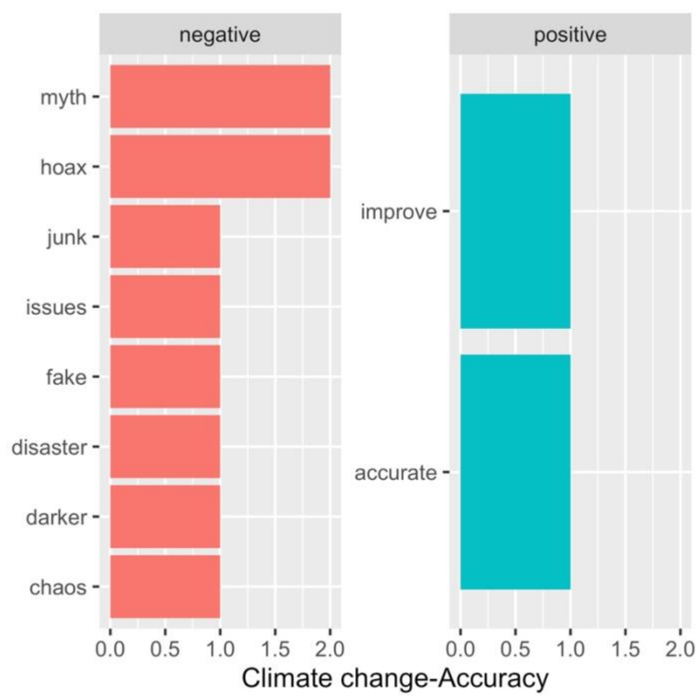
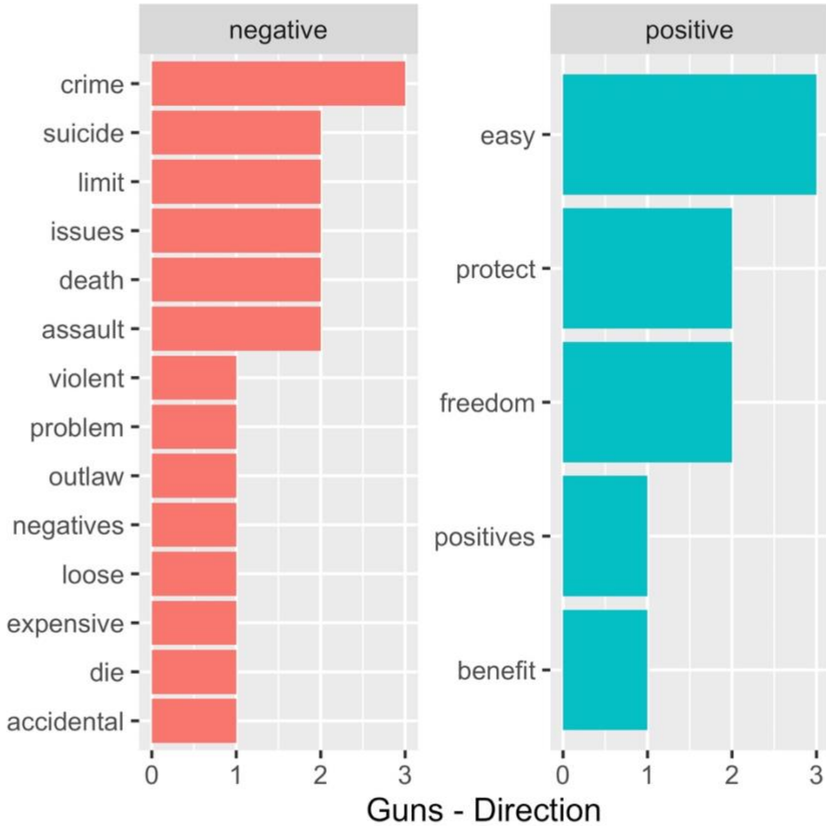

negative
positive
myth
hoax
junk
issues
fake
disaster
darker
chaos
improve
accurate
0.0 0.5 1.0 1.5 2.0
0.0 0.5 1.0 1.5 2.0
Climate change-Accuracy

Figure 10 illustrate the top words that contributed the most sentiment (by their frequency of occurrence) in each goal condition for the Climate change issue. More specifically, in the open-ended search queries regarding climate change in the Directional condition, 11 words were identified as “negative” and two word (“proven”, “support”) was identified as “positive”. Of the negative words, the word “hoax” appeared 9 times in the corpus, followed by “fraud” (n = 3), “fiction” (n = 2), “conspiracy”, “crisis”, “danger”, “deny”, “fake”, “hype”, “hysteria”, “lies” each with n = 1. Of the positive words, “proven” showed up twice and “support” showed up once. The sentiment score in the Directional goal thus was -19. Turning to the Accuracy condition of climate change-related search queries, eight negative words were detected: “myth” (n = 2), “hoax” (n = 2), “junk”, “issues”, “chaos”, “darker”, “disaster”, “fake” each with n =1. Positive words included “accurate” and “improve” with n = 1 each. Thus, the sentiment score for Accuracy condition was -8. Compared to the sentiment score for Directional condition, it could be seen that search queries related to climate change generated in the Directional condition were more negative.

Figure 11. Words indicating negative and positive sentiments found in Direction vs Accuracy goal condition – Gun


negative
positive
crime
suicide
limit
issues
death
assault
violent
problem
outlaw
negatives
loose
expensive
die
accidental
easy
protect
freedom
positives
benefit
0
1
2
3
Guns - Direction

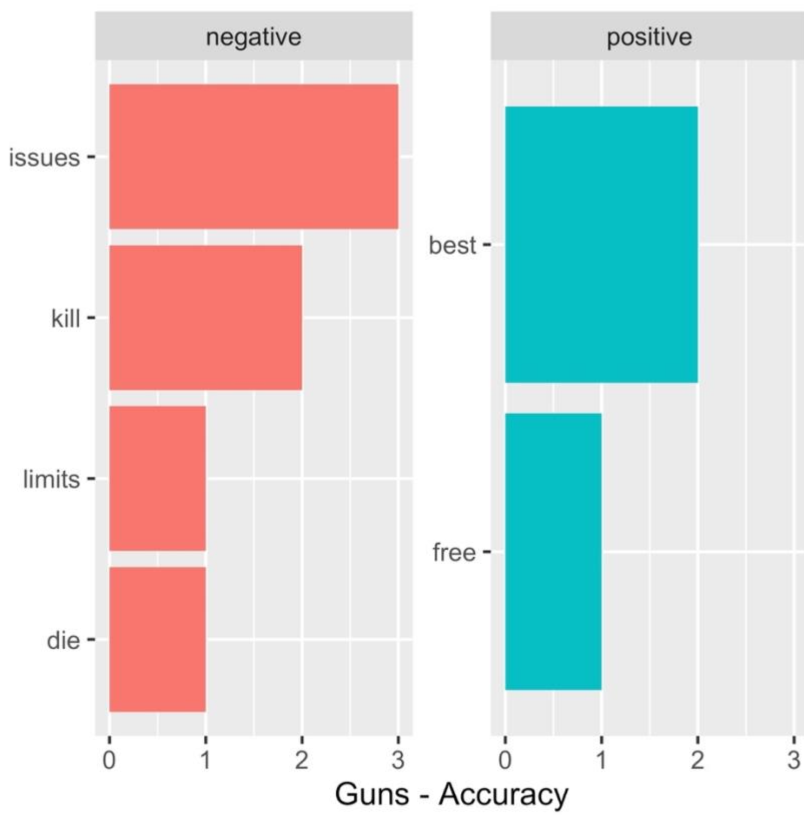
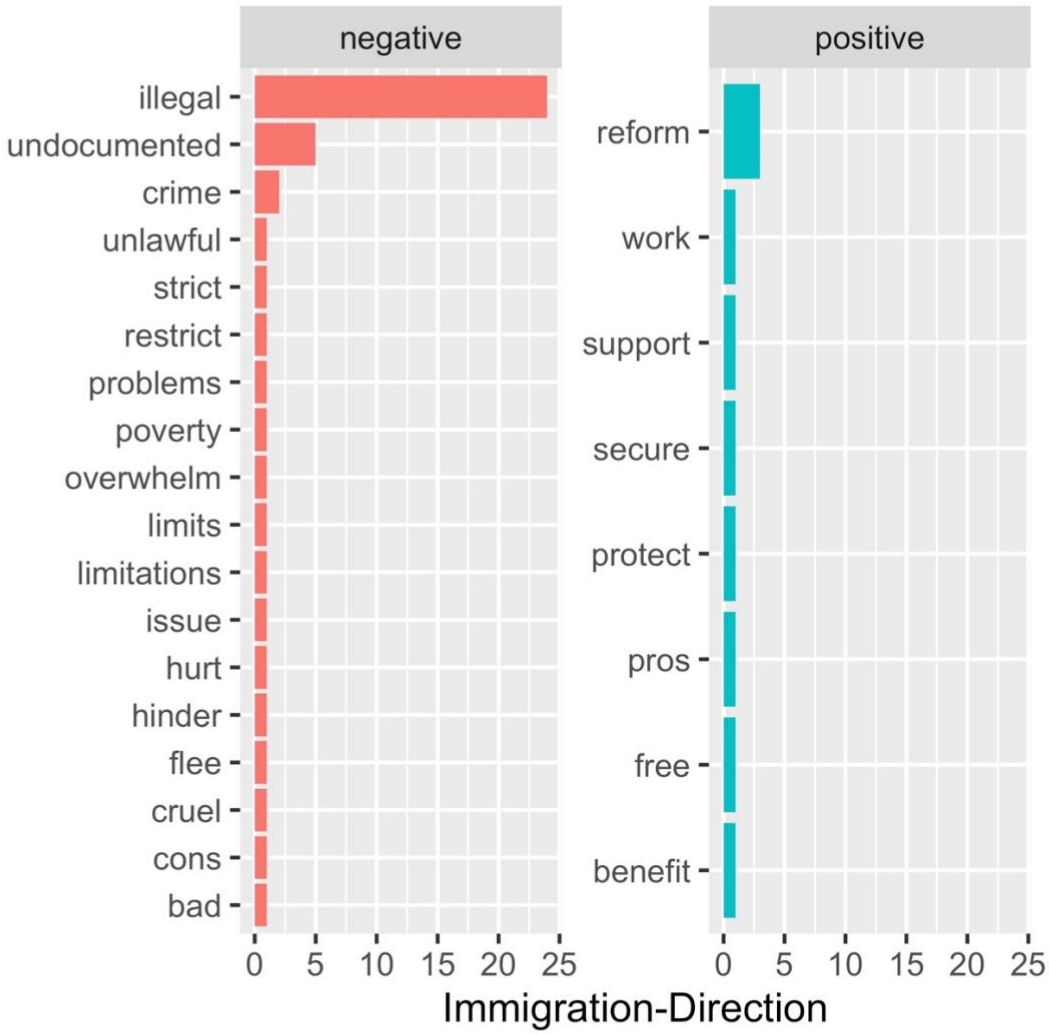

negative
positive
issues
kill
limits
die
best
free
0
1
2
3
Guns - Accuracy

Figure 11 illustrate the top words that contributed the most sentiment (by their frequency of occurrence) in each goal condition for the Gun control issue. More specifically, in the open-ended search queries regarding the issue of “guns” in the Directional condition, 14 words were identified as “negative” and 5 words were identified as “positive”. Of the negative words, the word “crime” appeared 3 times in the corpus, followed by “suicide”, “limit”, “issue”, “death” and “assault” (each with n = 2), “violent”, “problem”, “outlaw”, “negative”, “ loose”, expensive”, “die”, “accidental” each with n = 1. Of the positive words were “easy” (n = 3), “protect” (n = 2), “freedom” (n = 2), “positives” and “benefits” each with n = 1. The sentiment score in the Directional goal thus was -12. Turning to the Accuracy condition of guns-related search queries, four negative words were detected: “issue” (n = 3), “kill” (n = 2), “limits” and “die” each with n =1. Positive words included “best” (n = 2) and “free” with n = 1. Thus, the sentiment score for Accuracy condition was -4. Compared to the sentiment score for Directional condition, it could be seen that search queries related to gun control generated in the Directional condition were more negative.

Figure 12. Words indicating negative and positive sentiments found in Direction vs Accuracy goal condition – Immigration


negative
positive
illegal
undocumented
crime
unlawful
strict
restrict
problems
poverty
overwhelm
limits
limitations
issue
hurt
hinder
flee
cruel
cons
bad
reform
work
support
secure
protect
pros
free
benefit
0 5 10 15 20 25
0 5 10 15 20 25
Immigration-Direction

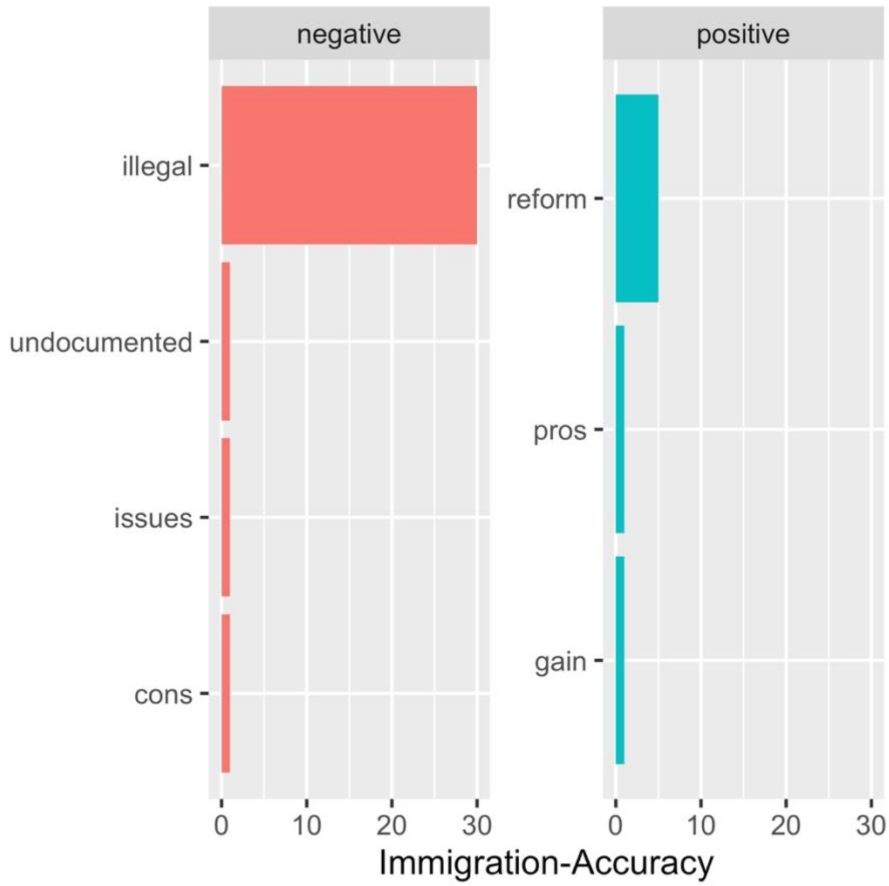


Figure 12 illustrate the top words that contributed the most sentiment (by their frequency of occurrence) in each goal condition for the Immigration issue. More specifically, in the open-ended search queries regarding the issue of "immigration" in the Directional condition, 18 words were identified as "negative" and 8 words were identified as "positive". Of the negative words, the word "illegal" appeared 24 times in the corpus, followed by "undocumented" (n = 5), "crime" (n = 2), "bad", "cons", "cruel", "flee", "hinder", "hurt", "issue", "limitations", "limits", "overwhelm", "poverty", "problems", "restrict", "strict", "unlawful" (each with n = 1). Of the positive words were "reform" (n = 3), "benefits", "free", "pros", "protect", "secure", "support", "work" each with n = 1. The sentiment score in the Directional goal thus was -36. Turning to the Accuracy condition of immigration-related search queries, four negative words were detected: "illegal" (n = 30), "undocumented", "issues", and "cons" each with n =1. Positive words included “reform” (n = 5), “gain” (n = 1), and “pros” with n = 1. Thus, the sentiment score for Accuracy condition was -26. Compared to the sentiment score for Directional condition, it could be seen that search queries related to immigration generated in the Directional condition were more negative.

To sum up, Figure 9-12 illustrate the top words that contributed the most sentiment (by their frequency of occurrence) in each goal condition for each issue. The results indicate that across all four issues and regardless of issue position, the Directional goal condition generated more negativity in respondent-supplied queries than the Accuracy goal condition.

Conclusion

The overarching goal of this Chapter is to examine the confirmatory search tendencies among political partisans in the context of American politics with regard to four polarizing political issues (abortion, climate change, gun control and immigration). Operationalizing confirmatory search tendencies as the preference for biased terms as search queries, this study found that the hypothesis regarding confirmatory search tendencies among partisans was only partially supported. Survey participants indicated preference for a number of biased terms that were congruent with both their broad political ideology and their position regarding the issues; however, this pattern varied across different specific terms and different issues. Partial correlations showed preference for biased terms had marginal association with the frequency of partisan media consumption. Mediation analysis further showed that the hypothesis that partisans adopted cues (i.e. biased language) from partisan media to search was not supported.

The patterns from open-ended search queries obtained from respondents showed that to some extent, people on the two sides of a political issue reported search terms that somewhat reflected their issue positions. This tendency was found to persist in both conditions, regardless of the information-seeking goals (Accuracy/Direction) that the individuals were primed to. Direction goals generated more negativity in open-ended search terms compared to Accuracy goals.

Implications

Implications of confirmatory search tendencies: A consequence of such tendencies is the potential that in the context of search engines, users of liberal and conservative ideologies will be exposed to two different worlds of information depending on the slant and the specificity of the search queries they use, a hypothesis widely coined as the "filter bubble" effect (Pariser, 2011). The next chapter of the dissertation, which explores this question, shows that varying degrees of differences in search results for conservatives and liberals suggest some support for the "filter bubble" effect; however, specific queries (rather than personalization) might be the key trigger that leads to discrepant search results.

Another implication is with regard to the role of partisan media. Here, no evidence was found concerning individuals adopting biased language from congenial media to seek information. The correlations between frequency of using conservative media use and selecting biased terms as search queries were found only for a handful of terms, and no significant association was found between liberal media use and biased term preference. The mediation analysis found no indirect effects through partisan media use.

These results suggest that confirmatory search tendencies were related to preexisting political orientations, rather than from media use habits. The observation of partisan media use having no significant association with term preference were somewhat perplexing, given the fact that these terms were commonly found in public discourse over polarizing issues. Some caution needs to be heeded in interpreting these results. First, partial correlation analysis was conducted only on complete cases and allowed no missing data. This left only a small number of cases with full data to be included and not enough power to detect significant relationships. P-value adjustment introduced further correction and made the results more conservative. Second, when hyper-partisan media outlets were included in the measure of partisan media, the relationships were also not found to be statistically significant. These outlets were not familiar names to the majority of the respondents in this survey (with the reported average of coming across those sources when browsing social media or online websites at about 1-2 in the 5-point scale), despite the fact that these outlets were identified by previous research as hyper-partisan sites with broad online circulation and traffic (Benkler, Faris, & Roberts, 2018; Heft, Mayerhoffer, Reinhardt, Knupfer, 2021). This could result in underestimating the effects of online partisan media. Future studies could challenge the findings reported here by increasing the number of cases, using pairwise deletion instead of listwise to use as much of the data as possible, and including a greater number of partisan media outlets which might be more familiar to the general public.

Implications with regard to political predispositions: Humans are innately biased creatures and such cognitive bias of being attentive to and more favorable of information consistent with previous beliefs has concrete manifestations in behavior. The fact that bias occurs even with very neutral question wording and even when individuals were asked to achieve accuracy and objectivity in search activities speaks to the unsettling effect of such cognitive bias. Past research, however, also shows that the need for cognition, education background, and search expertise could increase individuals’ attitudes towards using verification strategies in Web search. These characteristics could be included as moderators in future research to test the boundary conditions of confirmatory search.

On the other hand, this study and some previous works show that those with strong attitudes often use terms like “evidence”, “statistics” and “data” in search queries, which has the potential positive effects of directing users to official, legitimate and neutral sources (Yamamoto et al., 2018). Recent studies that tested different ways to induce “alternative” search behavior and promote balanced search showed that providing explicit feedback regarding search skills or past search histories could be beneficial. Similarly, displaying examples of high-quality queries could promote more effective formulation of search queries (Harvey, Crestani, Carman (2013). Other interventions including manipulating auto-suggested terms or inserting more attitude-incongruent information into search interfaces (Rele & Duchowski, 2005; Resnick, Maldonado, Santos, & Lergier, 2001) could also be considered in tandem with user behavior to design search systems that improve balanced information seeking.

Implications regarding selectivity in communication contexts: Past research on selectivity focused on different modes of communication, and the effects of such tendencies on individuals’ political outcomes. For example, the works of Stroud (2011) focus on the effects of mediated news exposure through partisan television on political participation, Mutz (2006) focused on the mechanisms of disagreement rooted from exposure to homogenous or heterogenous interpersonal networks on participation, or Garrett, Carnahan & Lynch (2013) which specified the phenomenon of “selective avoidance” vs “exposure” in the online environment. This study focuses the inquiry on the tendency of confirmatory search which draws from the same theoretical foundation of cognitive dissonance (Festinger 1957) - the underlying psychological mechanism behind selective exposure.

The findings reported here relate to the works on confirmation bias and informational utility and their impact on exposure (Knobloch-Westerwick & Kleinman, 2011). Fundamentally, in certain conditions, media users will be more motivated to attend to information that will best serve their goals and purposes (i.e. heeding "information utility" promotes more diverse news exposure) than giving in to confirmation bias (which would most likely motivate selective exposure). For example, before elections, information aligned with the stances of the party that might take office soon might carry more informational utility. The findings in this study did not support the idea that under Accuracy goals, information utility would override confirmation bias and motivate the use of more diverse search queries as respondents are instructed to search to find balanced and accurate information. Here, evidence of confirmation bias was still found to be stronger than evidence of information utility. Future research could draw on these findings to examine the conditions under which confirmation bias can be suppressed.

Limitations

The study suffers from several limitations. The first limitation is with regard to how respondents were "forced" to select biased terms as search queries and were put in hypothetical scenarios to supply search queries related to politics even if they might not search on their own. Patterns of aggregated search volume show that popular search terms are rarely related to politics, and search interests for political queries are much lower than those for celebrity, sports and entertainment (Dutton et al., 2017). Analysis of search engine query log also shows that different demographic groups in the U.S. are significantly different in their search behavior, in particular the topics they search for and how they search. Even for likely voters of specific parties, popular query topics are largely non-political. For example, the biggest single query topic among females who voted Democrat in the 2008 elections was shopping, and such among conservative males who voted Republicans were business, home & gardening or automobiles (Weber & Jaimes, 2011).

In short, criticism regarding the non-political nature of search behavior and search queries is also a common weakness (and criticism point) applied to selective exposure research. Future research could increase the ecological validity of online selectivity examinations by giving participants different information-seeking scenarios or goals, and used software that could capture and record how participants interact with web contents at real time (e.g. how they formulate queries, what kind of results they click on, how determined they are to find information that support their view, etc.).

The second limitation is with regard to the order of the two operationalized measures of confirmatory search tendencies in the survey questionnaire, where the introduction to biased search queries and ranking was at the top of the survey, and the examination of the impact of information seeking goals (Accuracy or Direction) on open-ended search queries was at the very end of the survey. This decision was made due to the understanding that search queries obtained from the partisan respondents might differ depending on motivated reasoning goals. However, presenting this experiment from the beginning of the survey could mistakenly put respondents in the mindset of either Accuracy or Direction goal conditions and pollute the responses to subsequent survey items, especially the ranking of biased queries part.

Placing and isolating the experimental manipulation of goals as the final survey items achieved two purposes, i.e. it allows testing the confirmatory search as a *general* tendency as well as *situational* behavior in the same sectional study. Open-ended queries from the situational manipulation showed a small number of queries that were the same as the biased queries in the set of queries presented to the people right at the beginning of the survey questionnaire, indicating that some respondents remembered these terms from the previous part of the survey, and were inadvertently "primed" in the open-ended part. However, despite a few instances of such similarity, the majority of open-ended search queries acquired from the respondents were organic and did not affect the conclusion drawn from this part. Still, future research would do well to conduct a separate, full-blown experimental design to establish the causal connection between information-seeking goals and confirmatory search behavior.

CHAPTER III: BIASED SEARCH QUERIES AND GOOGLE SEARCH RESULTS FOR LIBERALS AND CONSERVATIVES

Chapter II provided some evidence about the confirmatory search tendencies among political partisans with regard to contentious issues. A follow-up question that a reader might ask regarding such findings is "with what implications?". An emerging line of research in recent years that seeks to "audit" or reverse engineer search and recommendation algorithms on online platforms including YouTube, Amazon, Google News, Google Search among others has reported observations of search results promoting partisan bias (Hu, Jiang, Robertson & Wilson, 2019; Robertson et al., 2018), gender and race bias (Noble, 2018; Diakopoulos, Trielli, Stark, & Mussenden, 2018), misinformation and problematic contents (Hussein, Juneja, & Mitra, 2020; Juneja & Mitra, 2021). In particular, the effects of search engines on citizens' exposure to news diversity versus personalization - which have important ramifications for a well-informed citizenry and pluralist society – are of major scholarly concerns.

Motivated by previous research on the diversity of content delivered by search engines' algorithmic curation (Granka, 2010), this study aims to seek answers to the main question: How does Google Search results differ for users who self-identify as political liberals and conservatives using politically biased search queries? To understand the extent to which search engines provide users having different information needs (in this case, political partisans with demands for specific polarizing contents) with homogeneous or diverging search results, I recruited political partisans across the country to report search results given to them when conducting identical searches simultaneously. In particular, the empirical analysis in this Chapter seeks to answer three specific research questions:

RQ1. *Group effect:* When using the same queries, to what extent do conservatives and liberals see different results?

RQ2. *Query effect:* To what extent does query bias affect search result differences? In other words, do conservatives and liberals receive different results depending on the queries?
RQ3. *"Filter bubble" effect:* When conservatives use conservative queries, to what extent are search results different from liberals using liberal queries?

Before going into details about the methodology, the literature review section below briefly summarizes related works and key findings from extant research.

The "Google Effect": Information Diversity or Fragmentation

Search engines, in particular Google, hold enormous power in shaping user attention towards news outlets and online contents (Trielli & Diakopoulos, 2019) as they help reduce complexity for users and provide them the needed orientation to navigate the convoluted online environment (Newman et al., 2020; Steiner et al., 2020). In fact, the "core" function of search engines, as some scholars would argue, is to induce some sort of bias - in the forms of automated ranking and filtering mechanisms - that would prioritize relevant and reliable sources to tailor to specific needs of information seekers (Goldman, 2008; Lewandowski, 2017). As such, these mechanisms can profoundly shape individuals' understandings, judgments (Epstein & Robertson, 2015) and even reinforce existing bias (Noble, 2018; White & Horvitz, 2015). More importantly, algorithmic curation can create information inequities for individuals using the same search engine due to the mere effects of randomization (Makhortykh et al., 2020).

To evaluate the "gatekeeping" function of search engines, it is crucial to compare their outputs, in the form of search results, to certain normative standards (Nechushtai & Lewis, 2019). With regard to political search queries, in addition to quantitative measures of differences, these standards could include qualitative dimensions, such as "content and source diversity as well as the representation of different viewpoints" (Unkel & Haim, 2019). The extant empirical evidence suggests that apart from measurable discrepancies, the degree of source diversity (i.e. different types of sources included in the Search Engine Result Pages (SERPs)) and content diversity (i.e. informational thematic aspects) in search results vary by topics and contexts.

Regarding source diversity, past research found a high concentration of well-established and mainstream news organizations (e.g. The New York Times, Wikipedia, CNN, NPR, etc.) as the top domains returned in search results for neutral queries related to politics, such as names of political candidates. For example, Magin et al. (2015) examined the results of five search engines for political queries in German politics and found a range of diverse media in more than half of the results from Google Search, which included well-known news brands, government and public websites. Several studies came to similar conclusions with how the source diversity in Google Search and Google News results is somewhat limited due to the focus on journalism sites over alternative media, e.g. weblogs (Metaxas & Pruksachatkun, 2017; Puschmann, 2018; Trielli & Diakopoulos, 2019; Unkel & Haim, 2019). In addition, during periods of elections and campaigns, candidate-affiliated websites were found to rank high in search outputs (Kulshrestha et al., 2019), suggesting that search engines allow political candidates some level of control over their public image by prioritizing campaign sites.

With regard to non-political topics, Makhorthykh, Urman & Ulloa (2020) showed substantial differences between six search engines in how they prioritized information from official government agencies and alternative media to respond to the COVID-19 pandemic. In light of the dominance of national news across a wide range of information requests, some scholars have expressed concerns about the low traffic to local and community outlets directed from these news gatekeepers (Fischer, Jaidka & Lelkes, 2020).

Regarding content diversity, only a few studies to date have delved into exploring content diversity provided by search engines. These works indicate that diversity tends to increase in the *lower*-ranked search results (Li et al., 2014; Steiner et al., 2020) and that the teaser information on the SERPs, i.e. the search snippets, might contain stronger partisan bias than the contents embedded in the articles (D. Hu et al., 2019).

Finally, with regard to search results differences, the results are somewhat mixed. Studies with neutral search queries found little empirical support for the ideological “filter bubble” hypothesis (Haim et al., 2018; Nechushtai & Lewis, 2019; Puschmann, 2018), although search result differences of various sizes have been reported across geolocations (Hannák et al., 2017; Kliman-Silver et al., 2015), browser modes (Makhortykh et al., 2020; Robertson, Lazer, et al., 2018) or political preferences (Le et al., 2019; Robertson, Jiang, et al., 2018). In one of the latest studies examining the 2020 U.S. primary elections, Urman, Makhortykh & Ulloa (2021), examining queries “US elections”, “Donald Trump”, “Joe Biden”, “Bernie Sanders” in six engines (Google, Baidu, Bing, DuckDuckGo, Yahoo, and Yandex) found “substantial” differences in search results both *within* and *across* search engines by query, such that exposure to certain information seemed like “a matter of chance”, potentially leading to information discrepancies between users of both the same and different platforms. The authors also reported some evidence of Yahoo Search prioritizing pro-Sanders results compared to the ratio of supportive sources in results for Biden and Trump, resurfacing concerns about the political left-leaning bias found in search results in previous investigations (Diakopoulos et al., 2015, 2018; Epstein, Robertson, Shepherd, et al., 2017).

Since the questions of interest in this Chapter are with regard to the differences in the information sources, i.e., sets of information, returned to individuals of different political identifications in the SERPs, the following research question is formed:

RQ1. *Group effect:* When using the same queries, to what extent do conservatives and liberals see different results?

Biased Queries and Interactions with Search Algorithms

Contemporary Internet users have taken on new roles as producers and transmitters in addition to being consumers of online content. By producing original content and interacting with content produced by others (e.g. commenting, liking, sharing), human inputs play a role in creating the supply of "information priors" which machine algorithms learn from and operate upon. The interaction between machine learning and social learning is seen in how interactions with information in the digital world have led to trillions of trace data left behind that are then applied to machine learning technologies (Metaxa et al., 2019). With the increase of "smart" infrastructure that continually depends on user input to grow, users are 'complicit' and "performatively involved in shaping their own conditions of information access" (Gran, Booth & Bucher, 2020).

Sun, Nasraoui, Shafto (2020) argues that algorithmic bias interacts with humans in an iterative manner, which has a long-term effect on algorithms' performance. These authors argue for the close interaction of two sources of bias, including human action (i.e. the process by which people select information to label), and the process by which algorithm selects the subsets of information to present to people (which consists of three forms: personalization filter, active learning, and randomization). Conducting several controlled experiments, the authors showed that these bias modes, in addition to the initial training data class imbalance and human action, greatly affected the models learned by machine learning algorithms. Similarly, Kulshrestha et al. (2019) found that the interactions of input bias and ranking bias led to pronounced output bias in social media and Web search.

These ideas are further supported in two empirical studies in the domain of political search. First, Dutton et al. (2017) found that training algorithms by focusing on political search queries increased the proportion of politically related results compared to pre-training, suggesting the potential for “topical filter bubbles”. In other words, if political terms were entered into search engines multiple times within a short period, the results returned would increasingly become more political over time. Second, Le et al. (2019) found how certain search terms led to personalized search results that further reinforced the partisanship of pre-trained information personas. For example, “carbon footprint”, “Paris climate agreement”, and “uninsured Americans” amplified liberal bias for pro-immigration profiles, compared to terms like “flat tax”, “Medicare for all”, and “national debt” which led to more conservative-leaning results.

Past studies also consistently found evidence of individuals performing biased searches (D. Hu et al., 2019; Mustafaraj et al., 2020; Suzuki & Yamamoto, 2020; R. White, 2013). Mustafaraj et al. (2020) conducted their collection of search queries performed by Amazon MTurk workers in the U.S. identified two sources of bias in these terms: queries indicating bias regardless of context include words like “lose”, “win”, “good”, “bad”, compared to queries with bias placed within the context of a broader narrative, such as “blue wave”, “Diane Feinstein’s age”. This study found that two-thirds of the participants formulated queries (75% out of total) that conformed to the idea of “information cues”. In addition to biased phrases (which accounted for about 60% of the collection of 657 phrases), certain phrases were argued to potentially lead to data voids – which refer to how hostile and ill-natured terms or phrases could be optimized to lead people to problematic information, such as “MAGA bomber”, or “25th amendment”, “voter purge” or “immigration-based crime”. These results indicate the need for auditing research to elicit search phrases from users and use these phrases to understand how search results are returned based on different requests for information. The patterns found in this study are also in line with other research, which illustrate that simple search queries do not reveal well-known controversies but rather overrepresenting the “sunny side” of topics (Gerhart, 2004), and problematic queries could be strategically optimized to mislead people into distorted and misleading contents (Golebiewski & boyd, 2018).

As search engines algorithmically curate information by filtering and sorting web contents with high relevance to search queries, it is expected that specific keywords in search queries will largely determine search results. However, the magnitude of such personalization for political partisans regarding polarizing issues remains unknown. Thus, two research questions are formed regarding the query effect and filter bubble effect in search results.

RQ2. *Query effect:* To what extent do conservatives and liberals receive different results depending on the queries?

RQ3. *“Filter bubble” effect:* When conservatives use conservative queries, to what extent are the results different from liberals using liberal queries?

Search Engine Audits – Methodologies and Approaches

Two major challenges in studying search results are the fact that they are ephemeral (i.e. appearing in real-time responding to user queries), and that they are sensitive to changes in the media environment (in response to user feedback, times, locations, and a host of other factors) (Hannák et al., 2017; Kliman-Silver et al., 2015; Makhortykh et al., 2020). In their longitudinal, systematic investigation of search results in the six months leading up to the 2018 U.S. midterm elections, Metaxa, Park, Landay & Hancock (2019) argued for the need to conceptualize search results as a form of media - which deserves scrutinizing attention like other forms of new media - given their particular importance in elections and political contexts. In discussing the mechanisms that play into the volatile nature of search results, the authors wrote:

> *The production of search media can be decomposed into two main dimensions: endogenous factors and exogenous ones. Endogenous factors are those internal to the algorithm itself, such as strategic policy decisions made by Google about what content to surface or bury, user behavior which may feed back into the algorithm, and technical limitations like the rate at which a search engine can crawl and update its indices. Exogenous factors are attributes of "the real world"; in the context of political search media, this includes the behavior of political candidates, changes in current events, and decisions by news media of what to cover and how. (Metaxa et al., 2019)*

To reveal insights into the workings of search engines and their algorithms, "auditing" methods have been used extensively in the field of computer science. This methodology involves querying search engines repeatedly and recording the results under a range of conditions for the purpose of comparison (Haim, 2020; Mittelstadt, 2016).

The extant research that conduct search engine audits can be categorized according to their methodological approaches. First, researchers in the past developed browser plug-ins for different types of browsers (e.g. Chrome, Firefox), recruited volunteer participants to install these applications in their devices to automatically collect search results in addition to other metadata, including the user's language, geolocation, time and date of different searches (McMahon et al., 2017; Puschmann, 2018; Robertson, Jiang, et al., 2018; Robertson, Lazer, et al., 2018). While this method provides some ecological validity by scraping search results from real users, it is often expensive, susceptible to privacy concerns and representativeness due to the self-selection of participants. As an example, Robertson and colleagues found that search personalization was driven most strongly by the account log-in status, and varied as a function of the root query, multiple uses of Google services (e.g. Google Drive, Google Plus, etc.), political preferences. During the month following previous POTUS Donald Trump's inauguration, Robertson, Lazer, et al. (2018) found that participants who provided low-strength ratings of Trump received significantly more personalization in the ranking compositions of the SERPs. This finding was considered important given previous experimental evidence that those with low level of preferences were the most vulnerable to the effect of biased search rankings (Epstein & Robertson, 2015).

The second approach is automated scraping methods, which allow the capacity to scale up and control investigations with large quantities of search queries, multiple search engines, devices and locations. These methods often use a single (Kulshrestha et al., 2019; Trielli & Diakopoulos, 2019) or several virtual agents (H. Le et al., 2019; Makhortykh et al., 2020; Urman et al., 2021) to simulate browsing activities of different personas, isolate external factors (e.g. time, location), and control for the effects of inherent randomization and continuous updates of search results. Different from the previous browser plug-in approach, these methodologies examine how search engines rank and filter information in relation to different queries under the default, non-personalized conditions (Urman et al., 2021). The subjects of study, therefore, are the *algorithms* rather than actual human information-seeking behavior or user-based personalization factors. For example, Unkel & Haim (2019) found that in the context of the 2017 federal elections in Germany, news organizations and websites controlled by political parties were similarly prioritized in Google Search results for five different types of information personas.

Finally, Nechushtai & Lewis (2019) recommended a different approach that uses real-world settings to examine the extent of algorithmic personalization on news recommendation for individuals across different locations in real-time. With this method, the goal is “not to reverse-engineer algorithms, but to estimate their impact on public life by testing how they serve real users who seek political information”. As argued by the authors, the method takes into account individuals’ interactions with black-boxed algorithms and can be extended to understand the differential effects of search engines on diverse groups of participants. In detail, the authors conducted two experiments via Amazon M-Turk in the leading months to the 2016 presidential election, asking participants to conduct searches on Google News for “Hillary Clinton” and “Donald Trump” and submit the first five links they were recommended on each candidate. The analysis revealed quite similar news recommendations from a narrow set of news publishers for different searchers of various political leanings across the U.S. The authors concluded by deliberating the implications of algorithms as gatekeepers of digital news, as well as the challenges in assessing their performance from a normative perspective.

Considering the shared goals and purposes, I use a similar method to Nechushtai & Lewis (2019) to conduct the empirical inquiry in this Chapter, in addition to employing a more human-centered approach which takes into account actual biased search queries supplied by political partisans in the previous survey part. The methodology is described in more detail below.

Methods

From the list of open-ended search queries provided by self-identified liberals and conservatives in the survey dataset in Chapter II, I selected several specific search queries to create a lexicon of liberal and conservative search queries for 3 issues of interest (election integrity, abortion, and climate change). To systematically examine differences in search results as a function of the political ideology of users at the group level (i.e. group effect), a function of the political slant of queries at the group level (i.e. query effect), and a function of the match between users and queries (i.e. "filter bubble" effect), I conducted three crowdsourced experiments. In particular, I recruited individuals who identified as political liberals and conservatives through an online research subject pool (i.e. Prolific), collected their demographic information via a short survey, and instructed them to perform and submit the top 10 results from Google searches with six queries (three liberal and three conservative) regarding the issue of election integrity on December 3, 2020; abortion on January 28, 2021; and climate change on March 3, 2021. Queries related to "election integrity" were taken from an open-ended item in the survey in Chapter II in which the respondents recalled and described a recent instance in which they conducted an online search and discovered important information related to politics and governments. Queries related to "abortion" and "climate change" were taken from another open-ended item in which they reported three search queries or questions to look for information about each issue.

To ensure search results are personalized (i.e. tailored to each user based on their web data), restrictions were enforced to include only participants who had not deleted their Internet browser cookies in the past one month, and participants who would complete the study using a personal laptop, tablet, or desktop device. Participants were further instructed to log in to their Google accounts, open Google search result pages by clicking on pre-curated URLs with biased queries (e.g. www.google.com/search?q=climate+crisis+danger&num=10 to return the top 10 results for the left-leaning query "climate crisis danger"), and copy the result links onto the survey form. None of the participants' private information related to their Google account was visible to the researcher, their reported demographic information remained confidential, and the study followed protocols previously approved by the IRB (at UW-Madison). In total, 160 participants were asked to copy 60 result links associated with six queries per issue and given instructions to clear their search/browsing history and delete activities saved to their Google account at the end of the task.

To minimize the confounding effect of time variation and changes in news cycles, each experiment iteration (for each issue) was open to accepting submissions for only 2 hours, and the public information environment was monitored during this period to ensure there was no major disruption to the supply of information such as the occurrence of breaking news events. To have a roughly balanced set of political liberals and conservatives to submit the results within the same time window, a quota was set for each group and progress was monitored throughout the duration of each iteration.

Datasets from three experiments

Search results were collected from 50 liberals and 50 conservatives for the issue of "election integrity" on December 3, 2020. The analysis was conducted on 3,060 links which were based on the combination of the participants' ideology and the slant of queries, as described in the next part. For the second experiment, 15 liberals and 15 conservatives participated in the "abortion issue" on January 28, 2021. A total of 2,069 links were analyzed from this iteration. Lastly, for the issue of "climate change", 13 liberals and 17 conservatives participated on March 3, 2021, and 1,870 links collected from this iteration were included in the analysis.

Analysis strategy

From the lists of URLs collected for each issue, I merged all the URLs for 3 left-leaning queries together, and all the URLs for 3 right-leaning queries together to form two separate collections of left-leaning and right-leaning query URLs. The reason for this is to generate aggregate results for two slants of queries (conservative and liberal), rather than results for each individual query. These lists of URLs were further grouped according to the self-reported political ideology of the participant performing searches. As a result, four aggregate lists of URLs were created based on the 2 ideologies *2 query slants matrix. From the long strings of URLs (e.g. *cnn.com/2021/11/07/us/name-of-article/index.html*), the top-level domains (e.g. *cnn.com*) were extracted and the number of times the domains occurred in each of the four lists was counted. For example, *11alive.com* accounted for 11.8% (92 out of 783) of the links collected for conservative queries in the Conservatives' group.

Let *C/Qc*, *C/Ql*, *L/Qc*, and *L/Ql* respectively denote the lists of web domains found in "conservatives using conservative queries", "conservatives using liberal queries", "liberals using conservative queries", and "liberals using liberal queries" results. Differences were quantified by comparing the four lists in a pairwise manner. In other words, to analyze search result differences between any two lists, I calculated the proportion of unique results. Differences ranged from 0%-100%, where 0% indicates no difference between the two lists and 100% indicates a complete difference. In the Results section, I report the top media sources in search results returned for each of the three issues, and the search result differences (in percentage) by group effect, query effect, and "filter bubble" effect.

Results

The top domains showing up in search results for conservative queries and liberal queries performed by self-identified conservatives and liberals in the sample regarding the three topics

(Election, Abortion and Climate change) were presented in Table 17-19, and Table 20 presents the search result differences (in percentage). In the Appendix, the full lists of domains and occurrence rates in the search results regarding each of the three issues were reported.

Table 17. Top domains (sources) in search results returned for Conservatives versus Liberals – Election Integrity

Conservative queries – Conservatives (C/Qc)	Conservative queries – Liberals (L/Qc)	Liberal queries – Conservatives (C/Ql)	Liberal queries – Liberals (L/Ql)
11alive	11alive	cnn	history
clickondetroit	foxnews	usatoday	nbcnews
bridgemi	clickondetroit	nbcnews	washingtonpost
forbes	forbes	nytimes	cnn
wtoc	bridgemi	history	usatoday
npr	politifact	clickondetroit	npr
foxnews	northwestgeorgianews	ntdaily	ntdaily
wabe	factcheck	npr	thehill
12news	fox2detroit	washingtonpost	malibutimes
factcheck	washingtonpost	thehill	businessinsider
pbs	wsbtv	ramaponews	ramaponews
burlingtonfreepress	pbs	cbsnews	nytimes
abc7ny	9and10news	malibutimes	clickondetroit
clickorlando	nytimes	ft.com	jonesborosun
usatoday	courier-journal	wikipedia	wikipedia
nbcnews	thehill	ndtv	cbsnews
newsweek	wtoc	wjhl	magicvalley
nytimes	texastribune	10tv	nationalreview
tribuneindia	cnn	businessinsider	amazon
fox5atlanta	apnews	wreg	loc.gov

Note: Conservative queries are “fake ballots”, “voter fraud”, “ballot recounts”. Liberal queries are “Biden transition”, “transfer of power”, “peaceful transfer”. 50 liberals and 50 conservatives performed searches with these queries on December 3, 2020. Highlighted rows are shared domains across all four categories.

Search results regarding “Election integrity”

Table 17 demonstrates the most recommended publishers returned from election-related search queries on Google Search. Only two outlets, *nytimes.com* and *clickondetroit.com* (a local

Michigan outlet), were consistently recommended for conservatives and liberals regardless of query slant. Other legacy news organizations such as NPR, Washington Post, USA Today or NBC News were among the top-recommended publishers despite not consistently reaching both liberals and conservatives. The presence of local outlets from Georgia (*wabe.org, wotc.com, fox5atlanta.com*) in addition to other local media such as *bridgemi.com* (Michigan), *12news.com* (Arizona), *burlingtonfreepress.com* (Vermont), and *jonesborosun.com* (Arkansas) indicated that local media were also highly active in the production of post-election information.

Table 20 shows that when using conservative queries ("fake ballots", "voter fraud", and "ballot recounts"), conservatives and liberals in the sample saw approximately 46% difference in the total set of results, compared to a 53% difference when these two groups used liberal queries ("Biden transition", "transfer of power", "peaceful transfer"). More specifically, in the C/Qc vs L/Qc comparison ("conservatives using conservative queries" vs "liberals using conservative queries"), there were 33 shared domains between the two lists, 25 unique domains that showed up in the C/Qc including *npr.org, wabe.org, burlingtonfreepress.com, abc7ny.com, clickorlando.com, newsweek.com, fox5atlanta.com*, etc., and 3 unique domains that showed up in the L/Qc (*startribune.com, bbc.uk, and reuters.com*), hence 46% difference. In the C/Ql *vs* L/Ql comparison ("conservatives using liberal queries" vs "liberals using liberal queries"), search result differences rose to 53%. Together, these two sets of comparisons (i.e. group effect) show that conservatives and liberals in the sample saw a roughly equal set of both similarities and divergences in search results.

Table 20 further shows that search result differences were significantly larger by *query slant*. The conservative participants in the sample saw 88% different results if they used right-leaning as opposed to left-leaning queries (C/Qc vs C/Ql), and the liberals saw approximately

86% different results in the same scenario (L/Qc vs L/Ql). Finally, in the most extreme scenario, the "filter bubble" effect, when conservatives using right-leaning queries compared to liberals using left-leaning queries (C/Qc vs L/Ql), the magnitude of search result differences were at 88.3%.

Search results regarding "Abortion"

The top domains found in search results regarding "Abortion" could be found in Table 18, and quantified search result differences could be found in Table 20. Here, the two sites *wikipedia.org* and *bbc.com* showed up as top results across all four lists. Due to the nature of the issue as long as the search queries, in addition to ".com" sites, other domains ".gov", ".edu", ".org" from government agencies, advocacy groups, and educational institutions were among the top recommendations.

When using conservative queries ("why abortion is wrong", "abortion mental health", "life at conception"), conservatives and liberals in the sample saw about a 30% difference in search results. Results uniquely returned from right-leaning queries included medical journal articles (*jme.bmj.com; jstor.org, bmcmedicine.biomedcentral.com*), educational site (*innovating-education.org*), advocacy (*marchforlife.org*), and commercial site (*harpersbazaar.com*) among others. When using left-leaning queries ("women's rights to choose", "reproductive rights", "my body my rights"), the two groups saw a 46% difference in the aggregate results. These queries turned up domains like *prochoiceamerica.org, civilrights.org, statusofwomendata.org, reproductiverights.org* in addition to news (*nbcnews.com, bbc.com*), and local outlets (*shondaland.com, azcentral.com, shreveporttimes.com*). Altogether, the differences driven by ideological group profiles (i.e. group effect) (C/Qc vs L/Qc or C/Ql vs L/Ql) were at the 30-40%

range, indicating that conservatives and liberals saw largely similar results when using the same type of queries.

The query effect was much higher compared to the group effect. Specifically, conservatives saw an 80.2% difference when using different query slants (C/Qc vs C/Ql), and liberals saw a 93.5% difference when using different query slants (L/Qc vs L/Ql). In the case of matched ideologies, the "filter bubble" effect, there was a clear gap (85% difference) between the two worlds. In particular, it can be seen that right- and left-leaning web sources showed up in accordance with the political slant of the queries/searchers; for example, *prochoiceamerica.org, amnesty.org, aclu.org, wrj.org (Women of Reform Judaism)* versus *plannedparenthood.com, naapc.org (The National Association for the Advancement of Preborn Children), masscitizensforlife.org (Massachusetts Citizens For Life)*.

Table 18. Top domains (sources) in search results returned for Conservatives versus Liberals – Abortion

Conservative queries - Conservatives	Conservative queries – Liberals	Liberal queries - Conservatives	Liberal queries – Liberals
guttmacher.org	guttmacher.org	amazon.com	ohchr.org
princeton.edu	wikipedia.org	wikipedia.org	amazon.com
congress.gov	congress.gov	ohchr.org	wikipedia.org
wikipedia.org	princeton.edu	prochoiceamerica.org	prochoiceamerica.org
csulb.edu	medicine.missouri.edu	aclu.org	aclu.org
medicine.missouri.edu	csulb.edu	amnesty.org	beta.reproductiverights.org
cdlex.org	theatlantic.com	bbc.com	nbcnews.com
bc.edu	cdlex.org	wrj.org	amnesty.org
bbc.com	apa.org	beta.reproductiverights.org	civilrights.org
ncbi.nlm.nih.gov	ncbi.nlm.nih.gov	nbcnews.com	pubmed.ncbi.nlm.nih.gov
acpeds.org	plannedparenthood.com	pubmed.ncbi.nlm.nih.gov	azcentral.com
jme.bmj.com	psychiatryadvisor.com	yang2020.com	yang2020.com
apa.org	innovating-education.org	shondaland.com	bbc.com
psychiatryadvisor.com	bmcmedicine.biomedcentral.com	findlaw.com	wrj.org
naapc.org	acpeds.org	statusofwomendata.org	shreveporttimes.com
jstor.org	bc.edu	civilrights.org	brookings.edu
harpersbazaar.com	bbc.com	azcentral.com	shondaland.com
plannedparenthood.com	masscitizensforlife.org	shreveporttimes.com	link.springer.com

masscitizensforlife.org	naapc.org	brookings.edu	statusofwomendata.org
theatlantic.com	marchforlife.org	link.springer.com	findlaw.com

Note: Conservative queries are "why abortion is wrong", "abortion mental health", "life at conception". Liberal queries are "women's rights to choose", "reproductive rights", "my body my choice". 15 liberals and 15 conservatives performed searches with these queries on January 28, 2021. Highlighted rows are shared domains across all four categories.

Search results regarding "Climate change"

Table 19 presents the top domains regarding the issue of climate change. The three websites *climate.nasa.gov, edf.org (Environmental Defense Fund organization)* and *carbonbrief.org (a UK-based site covering climate science)* were the most frequently found information sources in search results regarding climate change. Unsurprisingly, websites belonging to environmental and international organizations (WWF, Green Peace, WHO), and non-US sites (*europa.eu, bundestag.de*) were top sources returned from the queries.

Conservatives and liberals saw about 20% different results for right-leaning queries ("climate change is junk", "lies about climate change", "is man-made climate change real"), and approximately 28% different results for left-leaning queries ("climate crisis danger", "human influence on climate", "mass extinction due to climate change"). These differences were driven by the presence of sites such as *lavoisier.com.au* (an Australian curated site), *junkscience.com, greenpeace.org* returned from conservative queries, and sites like *independent.org* (of the Independent Institute), *who.int* (World Health Organization), *cgd.ucar.edu* (from the National Center for Atmospheric Research NCAR) from liberal queries. The query effect (i.e. search result differences when comparing the top results if the conservatives or liberals in the sample performed one type of biased search versus the other) were at about 90% and the filter bubble effect tipped at nearly 85%.

Table 19. Top domains (sources) in search results returned for Conservatives versus Liberals - Climate change

Conservative queries - Conservatives	Conservative queries - Liberals	Liberal queries - Conservatives	Liberal queries - Liberals
wwf.org.uk	wwf.org.uk	climate.gov	climate.gov
edf.org	edf.org	climate.nasa.gov	climate.nasa.gov
theguardian.com	cei.org	edf.org	edf.org
lavoisier.com.au	climate.nasa.gov	europa.eu	ucsusa.org
theconversation.com	apnews.com	ucsusa.org	europa.eu
blogs.ei.columbia.edu	theguardian.com	scientificamerican.com	scientificamerican.com
farmprogress.com	abcnews.go.com	beforetheflood.com	independent.org
cei.org	junkscience.com	cordis.europa.eu	beforetheflood.com
huffpost.com	whistleblower.org	worldwildlife.org	climatecentral.org
forbes.com	forbes.com	carbonbrief.org	nationalgeographic.com
georgiapolicy.org	youtube.com	cleanet.org	worldwildlife.org
climate.nasa.gov	georgiapolicy.org	wikipedia.org	carbonbrief.org
northsidesun.com	heartland.org	news.mongabay.com	cordis.europa.eu
abcnews.go.com	terrapass.com	theconversation.com	wikipedia.org
edberry.com	theconversation.com	royalsocietypublishing.org	news.mongabay.com
greenpeace.org	carbonbrief.org	nationalgeographic.com	iberdrola.com
terrapass.com	channelnewsasia.com	climatecentral.org	royalsocietypublishing.org
carbonbrief.org	bundestag.de	independent.co.uk	who.int
channelnewsasia.com	farmprogress.com	iberdrola.com	cleanet.org
apnews.com	edberry.com	usatoday.com	cgd.ucar.edu

Note: Conservative queries are “climate change is junk”, “lies about climate change”, “is man-made climate change real”. Liberal queries are “climate crisis danger”, “human influence on climate”, “mass extinction due to climate change”. 13 liberals and 17 conservatives performed searches with these queries on March 3, 2021. Highlighted rows are shared domains across all four categories.

Altogether, the results suggest that compared to the Election and Abortion issue, conservatives and liberals shared more similarities in their search results returned from the same type of queries. However, the query effect and filter bubble effect were consistently high across all three issues under examination. This seems to suggest that what truly drove differences in search results were the queries submitted to the search engines. Rather than the “ideological confinement" that the "filter bubble" suggests, the questions submitted to search engines might actually what led individuals to see different results.

Table 20. Search results differences (%) for all 3 issues (Election integrity, Abortion, Climate Change)

Issue	Differences (%)					
	Group effect		Query effect		"Filter bubble" effect	
Election integrity	C/Qc *vs* L/Qc	46	C/Qc *vs* C/Ql	88	C/Qc *vs* L/Ql	88.3
	C/Ql *vs* L/Ql	53	L/Qc *vs* L/Ql	85.7		
Abortion	C/Qc *vs* L/Qc	30.7	C/Qc *vs* C/Ql	80.2	C/Qc *vs* L/Ql	85
	C/Ql *vs* L/Ql	46	L/Qc *vs* L/Ql	93.5		
Climate change	C/Qc *vs* L/Qc	20.9	C/Qc *vs* C/Ql	89.6	C/Qc *vs* L/Ql	84.6
	C/Ql *vs* L/Ql	28.2	L/Qc *vs* L/Ql	90.6		

Note: C = "Conservatives", L = "Liberals", Qc = "Queries that are conservative", Ql = "Queries that are liberal". C/Qc = "Conservatives use Queries that are conservative", etc.

To sum up, the analyses of search results returned from queries conducted by real-world participants on polarizing issues revealed the following insights to the RQs posed from the beginning:

Group effect: When using the same type of queries, conservatives and liberals received varying amounts of different results.

Query effect: Different types of queries led to significantly different results.

Filter bubble effect: Under the most extreme scenario of ideological confinement, conservatives using right-leaning queries and liberals using left-leaning queries saw fundamentally different sets of information, but such differences were not significantly larger than those by query effects.

Conclusion

This Chapter is conducted with the motivation to understand the implications of the confirmatory search tendencies identified among political partisans. In other words, how the choice of search queries and search algorithms shape exposure to online political information. Open-ended search queries supplied by survey participants in Chapter II indicate that people have broad information needs that might indicate prior preferences. Thus, the implications of such tendencies examined in this Chapter are: if two groups of conservatives and liberals use search terms that reflect or contradict their bias, to what degree do search result differences vary depending on political identification, query slant and the “filter bubble” scenario?

The comparison of the composition of search results for 160 self-identified political partisans in the U.S. at three different timepoints consistently showed that search personalization was driven more strongly by query slant than by the ideology of the searchers. The most extreme scenario in which the ideology of the users matched with the ideology of the queries also revealed substantial differences in search results; however, this magnitude was as large as such produced by the query effect. The size of search result differences also varied for the three issues under investigation (election integrity, abortion and climate change), with climate change seeing the smallest differences by group effect and largest differences by query effect. That is, for the same types of queries, liberals and conservatives saw only about 30% different results. The set of results was dominated by government-related sources and science-oriented pages, which could have the depolarizing effect for partisans. However, slanted queries regarding climate change resulted in fundamentally different sets of information. That is, for the conservative participants in the experiment, when they used conservative queries related to climate change, 90% of the results were different from when they used liberal queries. For the liberal participants, this number was 91%.

Altogether, the findings largely echo those of past research, which identified search queries as the key determinants of search results, due to the "relevance" criteria in search algorithms (Gerhart, 2004; Van Couvering, 2007). Also, the observation that political groups received both similar and different results from identical queries is in line with previous observations of Google Search exhibiting a mainstreaming effect that counterbalanced heterogenous search behaviors and dissimilar searches (Trielli & Diakopoulos, 2020). Compared to previous investigations of the political filter bubbles (e.g. Dutton et al., 2017; Nechushtai & Lewis, 2019), the magnitude of discrepancies was found to be larger in this study, which could be due to several factors, including the inclusion of all ten results in the first SERP in the analysis, and the manual collection of data which did not account for user-based personalization factors such as geolocation (Hannak et al., 2013). As search engines tend to display more pronounced differences in the long tails of the results (Li et al., 2014; Steiner et al., 2020) and discrepancies in search results often grow in proportion to the increase in physical distances (Kliman-Silver et al., 2015), these factors were among those that could significantly augment the size of differences observed in this study.

The findings here extend the literature by providing more insights into search results in the context of general Google web search, instead of results in news aggregation platforms like Google News. Previous studies commonly reported the prioritization of mainstream news sources in Google's Top Stories and Google News, but little is known about the sources of information showing up in the organic results (i.e. the blue links) of the general Web search. Since the issues examined here are established issues in American politics and have been widely discussed in the online public discourse, this study found that the information sources in search results extended beyond legacy news organizations to include alternative media, advocacy organizations, and weblog contents. Investigating whether these sources contain legitimate and credible information is beyond the scope of this study; however, this is an important research question to examine in the future.

Limitations

The study suffers from several limitations. First, I focused only on a limited set of biased queries regarding polarizing issues. The inclusion of a set of biased instead of generic queries in this study provides some external validity to search behavior in the real world (i.e. the observation that people do perform biased search activities) but also introduces several points of concern. First, the magnitude of differences in search results was somewhat overestimated due to the bias of the search queries, which accounted for the reason why search result differences found in this study were comparatively larger than in some earlier studies. Second, compared to previous works which studied personalization but did so using artificially created and controlled profiles (Haim et al., 2018; Le et al., 2019), in this study, external confounding factors (including A/B testing, geolocation, noise) were not controlled for and thus might play a part in augmenting the differences in search results.

Similarly, since political profiles were not strictly controlled in this study, causal inferences could not be made about the effect of political identification on search results. In the real world, a ‘liberal’ might not have a strictly ‘liberal online profile’ and a ‘conservative’ similarly might not consume only conservative media sources. In other words, it is likely that the participants' web activities and digital profiles are not truly reflective of their ideological leaning. However, since the goal was to detect search result discrepancies by group identification and search queries, this did not fundamentally alter the implications of the results.

Third, due to the group-level analysis, i.e. the collapse of queries of the same slant into one query group instead of discrete specific queries and the grouping of individuals into two ideological groups, individual-level variables including gender, age, location, political characteristics were not controlled for. The next step would be to include this level of analysis to re-examine the potential changes in the magnitude of search results differences.

Last, there was no baseline or control group in the design of these experiments to compare the magnitude of differences. This control condition could be the so-called non-personalized search results, i.e. search results without personalization based on user web activities such as browsing or search histories. Unfortunately, there is no way to go back and comprehensively collect these non-personalized results from the past. Future research could address this limitation by creating a baseline condition for comparison.

Implications

Implications for information-seeking habits: The results regarding search differences induced by query variation suggest the vulnerability and information inequalities potentially encountered by online users with specific demands for skewed and one-sided information. A phenomenon increasingly documented among individuals with low trust in the news media and government institutions (e.g. the anti-vaxxers or COVID-19 skeptics) showed that these particular groups and individuals approach data and information in problematic manners. For example, they would use public health data reported in the news media but focus on different metrics to come to their own conclusions about what the data means irrespective of official guidelines (e.g. COVID deaths rather than cases) (C. Le et al., 2021) These groups, who actively juxtapose official data with their lived experience, demonstrate information behaviors similar to those of the conservative evangelical Trump supporters in Tripodi's (2018) study, who prefer to do their own research to fact-check information reported in the media. People seek information to gain an understanding of the world, and "critical thinking", as well as motivated skepticism, among certain communities might entail using search engines to demand one-sided information. This introduces a moral dilemma of how search engines can best serve information seekers whose queries and surfing go against the normative values of open-mindedness and pluralism (Gerhart, 2004).

Implications for the epistemic responsibility of search engines as gatekeepers: The fact that there is evidence of dissimilarities in Google search results provided to users under different conditions calls into the role of search engines as modern gatekeepers in digital news consumption. Some scholars argued that by diversifying information presented to individuals, search engines encourage new knowledge discovery (Helberger, 2011). Others suggest that in times of emergency, such as a public health crisis, information needs to be uniform and accurate (Makhortykh et al., 2020). The tension between information diversity and the risks of misinformation remains a gray area considering how even for human gatekeepers, there are still debates over their roles in the news curation process (Nechushtai & Lewis, 2019). Despite contrasting perspectives on the normative function of search engines on information heterogeneity, it seems reasonable to hold these entities responsible for performing their epistemological duty in offering the needed transparency and reliability of search results to the public (Mustafaraj & Walsh, 2019; Pasquale, 2015). Finally, the "relevance" aspect of search algorithms could create the potential for “data voids”. This calls for the need to frequently and systematically audit search engines results to detect malicious content.

CONCLUSION AND FUTURE DIRECTIONS

In combination, the three empirical investigations in this dissertation provide new insights into the role of Internet search engines in the information acquisition and potential information inequalities in the current high-choice information environment.

Chapter I of the dissertation highlights that although the baseline frequency of using search engines for political purposes varied based on a range of demographic and informational characteristics, algorithmic knowledge, or the understanding of how search engines as gatekeepers operate in real life, was crucial both as the independent determinant of political search behaviors and the mediating path for the effects of various exogenous variables.

Such knowledge could have important political implications if the tendency to search to confirm preexisting beliefs is the information-seeking tendency among the most politically active individuals, i.e. the political partisans, under various information-seeking conditions. As illustrated in Chapter II, the preference for specific biased terms as search queries was observed among these individuals regarding a number of polarizing issues. The supply of biased queries among these respondents was also observed under both accuracy-motivated and directional motivated reasoning goals. These biased queries could send political partisans to different perspectives about the issues and further away from the other side depending on the specific queries used to access information (Chapter III).

Altogether, these findings tie back to the previous discussion on the two conceptions of technology (technological determinism and social shaping approach) and provide empirical evidence to assess the relative merits of these competing perspectives concerning the use and impact of search media in the process of political information acquisition.

The social shaping of technology perspective explores the ways in which "social, institutional, economic and cultural factors shape the *direction* of technological innovation, the *form* of technology, and the *outcomes* of technological change for different groups in society" (R. Williams & Edge, 1996, p. 868). In Chapters I and II, the impact of such social factors was demonstrated in how individuals and partisans made choices about search versus other media and adopted search tools for their information-seeking purposes. Individual factors, such as media trust, political interest and political predispositions, were shown to determine, for instance, how much confidence people placed in search results, how much understanding they had about search engines, and the degree to which such confidence and understanding were translated into political search behaviors. This "interpretive flexibility" demonstrated in the understanding and application of search media suggests that the social shaping perspective was useful in contextualizing the effects of search technology.

However, this is not to say that search technology does not have an important shaping impact. One dominant theoretical conception of search algorithms posits that the "filter bubbles" generated by personalization will lead to fragmentation and social disintegration. Chapter III indicates that there remained a possibility of such divergence, although it might be triggered by the "relevance" aspect of search queries rather than the "personalization" aspect of algorithms. Similarly, the popularity and high ranking of legacy mainstream media outlets in search results returned for several queries observed in Chapter III illustrated the "rich-get-richer" dynamics inherent in the link structure of the Web. These findings illustrate the influence of search technology on society, particularly the potential information inequalities resulted from discrepant search results due to built-in randomization.

Directions for future research

Expanding online selective exposure research to the context of online search: It becomes clear that information exposure in the digital age entails the interactions of individuals and platforms, especially algorithms, in the information selection process. New advances in selective exposure research take into account algorithmic filtering and ranking mechanisms (Cinelli et al., 2020) together with user behaviors (i.e., initial query formation, selection/non-selection of results, query abandonment/revision) in the shaping of selective exposure (Slechten et al., 2021).

Experimental evidence shows that the highest search rankings have a significant impact on individuals' beliefs and information selection regardless of the information being consistent with prior beliefs or not (Epstein & Robertson, 2015; Slechten et al., 2021). Also, search results that people select to read on were found to significantly impact priorly held beliefs (Knobloch-Westerwick et al., 2015; Westerwick et al., 2017). Future research would need to further explore confirmation bias in relation to Web search and search result selection process, as well as how to suppress confirmation bias and promote critical information seeking.

Understanding the impact of partisan personalization on news behaviors and implications: One of the findings in this dissertation suggests that if political partisans use search engines to select attitude-congruent content, then these platforms might have considerable economic incentives to tailor specific types of content to meet the demands of these users. Experimental studies exploring the same question but concerning digital news aggregators found that when being shown more politically congruent news in a made-up news aggregator, users perceived the site to be more credible, returned to it more for news, and more worrisomely, read less mainstream news on the site (Bryanov et al., 2020). This evidence further demonstrates how a subtle change in the supply of information, solely determined by the alteration of algorithms at the discretion of digital platforms, could go a long way in directing user attention and consumption of political information. A promising area for future research would be to explore how news engagement and alternative information-seeking could be induced by manipulating different degrees of personalization and partisan content. This could help shed more light on the specific ways in which digital platforms can increase their social accountability and maximize their potential as information gatekeepers.

Understanding search media, in conjunction with social media in the current information flows and misinformation: A growing line of research has shifted focus to examine how search results across digital platforms including Amazon, Bing, Facebook, Twitter, and YouTube contain false and deliberately misleading information. For example, Juneja and Mitra (2021) found the presence of a filter bubble effect in Amazon recommendations for vaccine-related queries, where users' interactions with misinformative products on the e-commerce platform led to even more misinformation in product recommendations. Similarly, on YouTube, personalization based on watch history was also found to further emphasize misinformation; in other words, consuming videos that feature misinformation related to a variety of science and health topics triggered the algorithms to recommend more misinformative content (Hussein et al., 2020). Finally, regarding Bing search engines, Bush & Zaheer (2019) found that Bing returns more problematic content in their results and at a higher rate than Google does. This included conspiracy, white-supremacist content, Russian propaganda, and low-quality sites, which were placed in high-ranked results in response to unrelated queries and regardless of users' explicit intent to look for such information in several cases.

These findings suggest similar auditing methods can be applied to different topics of public importance on multiple digital platforms to scrutinize the potential for misinformation. Perhaps more importantly, these call for the need to examine search and social media platforms in tandem for a nuanced examination of the interconnected roles of these digital intermediaries in surfacing and promoting misinformation, particularly during breaking news events.

REFERENCES

Adarves-Yorno, I., Jetten, J., Postmes, T., & Haslam, S. A. (2013). What Are We Fighting For?: The Effects of Framing on Ingroup Identification and Allegiance. *The Journal of Social Psychology*, *153*(1), 25–37. https://doi.org/10.1080/00224545.2012.701673

Afifi, W. A. (2017). Motivated Information Management and Other Approaches to Information Seeking. In *Oxford Research Encyclopedia of Communication*. https://doi.org/10.1093/acrefore/9780190228613.013.512

Agosto, D., E. (2001). Bounded rationality and satisficing in young people's Web-based decision making. *Journal of the American Society for Information Science and Technology*. https://doi-org.ezproxy.library.wisc.edu/10.1002/asi.10024

Anspach, N. M., & Carlson, T. N. (2018). What to Believe? Social Media Commentary and Belief in Misinformation. *Political Behavior*. https://doi.org/10.1007/s11109-018-9515-z

Arendt, F., & Fawzi, N. (2019). Googling for Trump: Investigating online information seeking during the 2016 US presidential election. *Information, Communication & Society*, *22*(13), 1945–1955. aph.

Bandura, A. (2001). Social Cognitive Theory: An Agentic Perspective. *Annual Review of Psychology*, *52*, 1–26.

Barnidge, M., & Rojas, H. (2014). Hostile Media Perceptions, Presumed Media Influence, and Political Talk: Expanding the Corrective Action Hypothesis. *International Journal of Public Opinion Research*, *26*(2), 135–156. https://doi.org/10.1093/ijpor/edt032

Bawden, D., & Robinson, L. (2009). The dark side of information: Overload, anxiety and other paradoxes and pathologies. *Journal of Information Science*, *35*(2), 180–191. https://doi.org/10.1177/0165551508095781

Bawden, D., & Robinson, L. (2020). Information Overload: An Introduction. In *Oxford Research Encyclopedia of Politics*. https://oxfordre-com.ezproxy.library.wisc.edu/politics/view/10.1093/acrefore/9780190228637.001.0001/acrefore-9780190228637-e-1360

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press.

Blank, G., & Groselj, D. (2014). Dimensions of Internet use: Amount, variety, and types. *Information, Communication & Society*, *17*(4), 417–435. https://doi.org/10.1080/1369118X.2014.889189

Bodó, B., Helberger, N., Eskens, S., & Möller, J. (2019). Interested in Diversity. *Digital Journalism*, *7*(2), 206–229. https://doi.org/10.1080/21670811.2018.1521292

Bolsen, T., & Druckman, J. N. (2015). Counteracting the Politicization of Science. *Journal of Communication*, *65*(5), 745–769.

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

Bradshaw, S. (2019). Disinformation optimised: Gaming search engine algorithms to amplify junk news. *Internet Policy Review*, *8*(4). https://policyreview.info/articles/analysis/disinformation-optimised-gaming-search-engine-algorithms-amplify-junk-news

Bryanov, K., Watson, B. K., Pingree, R. J., & Santia, M. (2020). Effects of Partisan Personalization in a News Portal Experiment. *Public Opinion Quarterly*, *84*(S1), 216–235. https://doi.org/10.1093/poq/nfaa011

Buente, W., & Robbin, A. (2008). Trends in Internet information behavior, 2000–2004. *Journal of the American Society for Information Science and Technology*, *59*(11), 1743–1760. https://doi.org/10.1002/asi.20883

Busby, E., Flynn, D. J., & Druckman, J. N. (2018). Studying Framing Effects on Political Preferences: Existing Research and Lingering Questions. In *Doing News Framing Analysis II: Empirical and Theoretical Perspectives*. Routledge.

Bush, D., & Zaheer, A. (2019). *Bing's Top Search Results Contain an Alarming Amount of Disinformation*. Stanford Freeman Spogli Institute for International Studies. https://fsi.stanford.edu/news/bing-search-disinformation

Byrnes, J. (2020, July 29). *Jordan confronts tech CEOs over claims of anti-conservative bias* [Text]. TheHill. https://thehill.com/homenews/house/509619-jordan-confronts-tech-ceos-with-claims-of-anti-conservative-bias

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American Voter*. University of Chicago Press.

Cardenal, A. S., Aguilar-Paredes, C., Galais, C., & Pérez-Montoro, M. (2019). Digital Technologies and Selective Exposure: How Choice and Filter Bubbles Shape News Media Exposure. *The International Journal of Press/Politics*, *24*(4), 465–486. https://doi.org/10.1177/1940161219862988

Chadwick, A. (2017). *The Hybrid Media System: Politics and Power*. Oxford University Press.

Chong, D., & Druckman, J. N. (2007). Spanish Abstract. *Journal of Communication*, *57*(1), 99–118. https://doi.org/10.1111/j.1460-2466.2006.00331_3.x

Cinelli, M., Brugnoli, E., Schmidt, A. L., Zollo, F., Quattrociocchi, W., & Scala, A. (2020). Selective exposure shapes the Facebook news diet. *PLOS ONE*, *15*(3), e0229129. https://doi.org/10.1371/journal.pone.0229129

Clifford, S. (2019). How Emotional Frames Moralize and Polarize Political Attitudes: Moralizing Effects of Persuasive Frames. *Political Psychology*, *40*(1), 75–91. https://doi.org/10.1111/pops.12507

Converse, P. E. (2007). Perspectives on Mass Belief Systems and Communication. In *The Oxford Handbook of Political Behavior*. https://doi.org/10.1093/oxfordhb/9780199270125.003.0008

Coppini, D., Duncan, M. A., McLeod, D. M., Wise, D. A., Bialik, K. E., & Wu, Y. (2017). When the whole world is watching: A motivations-based account of selective expression and exposure. *Computers in Human Behavior*, *75*, 766–774. https://doi.org/10.1016/j.chb.2017.04.020

Cothey, V. (2002). A longitudinal study of World Wide Web users' information-searching behavior. *Journal of the American Society for Information Science and Technology*, *53*(2), 67–78. https://doi.org/10.1002/asi.10011

Cotter, K., & Reisdorf, B. C. (2020). *Algorithmic Knowledge Gaps: A New Dimension of (Digital) Inequality*. 21.

Craig, S. C., Niemi, R. G., & Silver, G. E. (1990). Political efficacy and trust: A report on the NES pilot study items. *Political Behavior, 12*(3), 289–314. https://doi.org/10.1007/BF00992337

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Dawson, E., Gilovich, T., & Regan, D. T. (2002). Motivated Reasoning and Performance on the was on Selection Task. *Personality and Social Psychology Bulletin*, *28*(10), 1379–1387. https://doi.org/10.1177/014616702236869

Diakopoulos, N. (2015). Algorithmic Accountability: Journalistic investigation of computational power structures. *Digital Journalism*, *3*(3), 398–415. https://doi.org/10.1080/21670811.2014.976411

Diakopoulos, N., & Koliska, M. (2017). Algorithmic Transparency in the News Media. *Digital Journalism*, *5*(7), 809–828. https://doi.org/10.1080/21670811.2016.1208053

Diakopoulos, N., Trielli, D., & Mussenden, S. (2015, December 7). *Why Google Search Results Favor Democrats*. Slate Magazine. https://slate.com/technology/2015/12/why-google-search-results-favor-democrats.html

Diakopoulos, N., Trielli, D., Stark, J., & Mussenden, S. (2018). I Vote For—How Search Informs Our Choice of Candidate. In *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple*. M. Moore and D. Tambini (Eds.). https://www.academia.edu/37432634/I_Vote_For_How_Search_Informs_Our_Choice_of_Candidate

Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How Elite Partisan Polarization Affects Public Opinion Formation. *American Political Science Review*, *107*(1). https://www-jstor-org.ezproxy.library.wisc.edu/stable/23357757?seq=1#metadata_info_tab_contents

Dutton, W. H., & Reisdorf, B. C. (2019). Cultural divides and digital inequalities: Attitudes shaping Internet and social media divides. *Information, Communication & Society*, *22*(1), 18–38. https://doi.org/10.1080/1369118X.2017.1353640

Dutton, W. H., Reisdorf, B. C., Dubois, E., & Blank, G. (2017). Search and Politics: The Uses and Impacts of Search in Britain, France, Germany, Italy, Poland, Spain, and the United States. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2960697

Edgerly, S., Vraga, E. K., McLaughlin, B., Alvarez, G., Yang, J., & Kim, Y. M. (2014). Navigational Structures and Information Selection Goals: A Closer Look at Online Selectivity. *Journal of Broadcasting & Electronic Media*, *58*(4), 542–561. https://doi.org/10.1080/08838151.2014.966360

Epley, N., & Gilovich, T. (2016). The Mechanics of Motivated Reasoning. *Journal of Economic Perspectives*, *30*(3), 133–140. https://doi.org/10.1257/jep.30.3.133

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*, *112*(33), E4512–E4521. https://doi.org/10.1073/pnas.1419828112

Epstein, R., Robertson, R. E., Lazer, D., & Wilson, C. (2017). Suppressing the Search Engine Manipulation Effect (SEME). *Proceedings of the ACM on Human-Computer Interaction*, *1*(CSCW), 42:1-42:22. https://doi.org/10.1145/3134677

Epstein, R., Robertson, R. E., Shepherd, S. J., & Zhang, S. (2017). *A Method for Detecting Bias in Search Rankings, with Evidence of Systematic Bias Related to the 2016 Presidential Election*. 5.

Eslami, M., & Karahalios, K. (2017). *Investigating Users ' Understanding of Invisible Algorithms and Designing around It*.

Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K., & Kirlik, A. (2016). First I “like” it, then I hide it: Folk Theories of Social Feeds. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI ’16*, 2371–2382. https://doi.org/10.1145/2858036.2858494

Fallows, D. (2005). Internet searchers are confident, satisfied and trusting – but they are also unaware and naïve. [Pew Internet & American Life Project]. *Pew Internet & American Life Project: Search Engine Users*. https://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2005/PIP_Searchengine_users.pdf.pdf

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

Fischer, S., Jaidka, K., & Lelkes, Y. (2020). Auditing local news presence on Google News. *Nature Human Behaviour*, *4*(12), 1236–1244. https://doi.org/10.1038/s41562-020-00954-0

Fiske, S. T., & Taylor, S. E. (2013). *Social Cognition: From Brains to Culture*. SAGE.

Fletcher, R., & Nielsen, R. K. (2019). Generalised scepticism: How people navigate news on social media. *Information, Communication & Society*, *22*(12), 1751–1769. https://doi.org/10.1080/1369118X.2018.1450887

Fletcher, R., & Park, S. (2017). The Impact of Trust in the News Media on Online News Consumption and Participation. *Digital Journalism*, *5*(10), 1281–1299. https://doi.org/10.1080/21670811.2017.1279979

Gallagher, K. M., & Updegraff, J. A. (2012). Health Message Framing Effects on Attitudes, Intentions, and Behavior: A Meta-analytic Review. *Annals of Behavioral Medicine*, *43*(1), 101–116. https://doi.org/10.1007/s12160-011-9308-7

Garrett, R. K., Carnahan, D., & Lynch, E. K. (2013). A Turn Toward Avoidance? Selective Exposure to Online Political Information, 2004–2008. *Political Behavior*, *35*(1), 113–134. https://doi.org/10.1007/s11109-011-9185-6

Garrett, R. K., & Stroud, N. J. (2014). Partisan Paths to Exposure Diversity: Differences in pro- and Counterattitudinal News Consumption. *Journal of Communication*, *64*(4), 680–701. https://doi.org/10.1111/jcom.12105

Gerhart, S. (2004). Do Web search engines suppress controversy? *First Monday*. https://doi.org/10.5210/fm.v9i1.1111

Gil de Zúñiga, H., Diehl, T., & Ardèvol-Abreu, A. (2017). Internal, External, and Government Political Efficacy: Effects on News Use, Discussion, and Political Participation. *Journal of Broadcasting & Electronic Media*, *61*(3). https://www-tandfonline-com.ezproxy.library.wisc.edu/doi/full/10.1080/08838151.2017.1344672

Giomelakis, D., & Veglis, A. (2015). Investigating Search Engine Optimization Factors in Media Websites: The case of Greece. *Digital Journalism*, *4*, 379–400.

Goldman, E. (2008). Search Engine Bias and the Demise of Search Engine Utopianism. In A. Spink & M. Zimmer (Eds.), *Web Search* (Vol. 14, pp. 121–133). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-75829-7_8

Golebiewski, M., & boyd, danah. (2018). Where Missing Data Can Easily Be Exploited. *Data & Society*.

Gran, A.-B., Booth, P., & Bucher, T. (2020a). To be or not to be algorithm aware: A question of a new digital divide? *Information, Communication & Society*, 1–18. https://doi.org/10.1080/1369118X.2020.1736124

Gran, A.-B., Booth, P., & Bucher, T. (2020b). To be or not to be algorithm aware: A question of a new digital divide? *Information, Communication & Society*, 1–18. https://doi.org/10.1080/1369118X.2020.1736124

Granka, L. A. (2010). The Politics of Search: A Decade Retrospective. *The Information Society*, *26*(5), 364–374. https://doi.org/10.1080/01972243.2010.511560

Gruszczynski, M., & Wagner, M. W. (2017). Information Flow in the 21st Century: The Dynamics of Agenda-Uptake. *Mass Communication and Society*, *20*(3), 378–402. https://doi.org/10.1080/15205436.2016.1255757

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

Gunter, B., Rowlands, I., & Nicholas, D. (2009). *The Google Generation: Are ICT innovations Changing information Seeking Behaviour?* Elsevier.

Gunther, A. C. (1992). Biased Press or Biased Public? Attitudes towards Media Coverage of Social Groups. *Public Opinion Quarterly*, *56*(2), 147–167. https://doi.org/10.1086/269308

Guo, Y., Lu, Z., Kuang, H., & Wang, C. (2020). Information avoidance behavior on social network sites: Information irrelevance, overload, and the moderating role of time pressure. *International Journal of Information Management*, *52,* 102067. https://doi.org/10.1016/j.ijinfomgt.2020.102067

Haim, M. (2020). Agent-based Testing: An Automated Approach toward Artificial Reactions to Human Behavior. *Journalism Studies*, *21*(7), 895–911. https://doi.org/10.1080/1461670X.2019.1702892

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*, 330.

Hannák, A., Sapieżyński, P., Khaki, A. M., Lazer, D., Mislove, A., & Wilson, C. (2017). Measuring Personalization of Web Search. *ArXiv:1706.05011 [Cs]*. http://arxiv.org/abs/1706.05011

Hannak, A., Sapiezynski, P., Molavi Kakhki, A., Krishnamurthy, B., Lazer, D., Mislove, A., & Wilson, C. (2013). Measuring personalization of web search. *Proceedings of the 22nd International Conference on World Wide Web - WWW ’13*, 527–538. https://doi.org/10.1145/2488388.2488435

Hargittai, E. (2010). Digital Na(t)ives? Variation in Internet Skills and Uses among Members of the “Net Generation”*. *Sociological Inquiry*, *80*(1), 92–113. https://doi.org/10.1111/j.1475-682X.2009.00317.x

Hargittai, E., & Dobransky, K. (2017). Old Dogs, New Clicks: Digital Inequality in Skills and Uses among Older Adults. *Canadian Journal of Communication*, *42*(2). https://doi.org/10.22230/cjc.2017v42n2a3176

Harvey, M., Crestani, F., & Carman, M., J. (2013). Building user profiles from topic models for personalised search. *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management*, 2309–2314.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (Second edition). Guilford Press.

Heft, A., Knüpfer, C., Reinhardt, S., & Mayerhöffer, E. (2021). Toward a Transnational Information Ecology on the Right? Hyperlink Networking among Right-Wing Digital

News Sites in Europe and the United States. *The International Journal of Press/Politics*, *26*(2), 484–504. https://doi.org/10.1177/1940161220963670

Helberger, N. (2011). Diversity by Design. *Journal of Information Policy, 1*, 441–469. https://doi.org/10.5325/jinfopoli.1.2011.0441

Helsper, E. J., & Deursen, A. J. A. M. van. (2017). Do the rich get digitally richer? Quantity and quality of support for digital engagement. *Information, Communication & Society*, *20*(5), 700–714. https://doi.org/10.1080/1369118X.2016.1203454

Hoff, K. A., & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Howard, P. N., & Massanari, A. (2007). Learning to Search and Searching to Learn: Income, Education, and Experience Online. *Journal of Computer-Mediated Communication*, *12*(3), 846–865. https://doi.org/10.1111/j.1083-6101.2007.00353.x

Hu, D., Jiang, S., Robertson, R. E., & Wilson, C. (2019). Auditing the Partisanship of Google Search Snippets. *Proceedings of the 2019 World Wide Web Conference (WWW'19), May 2019, San Francisco, CA, USA*, 12. https://dl-acm-org.ezproxy.library.wisc.edu/doi/10.1145/3308558.3313654

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 168–177. https://doi.org/10.1145/1014052.1014073

Hussein, E., Juneja, P., & Mitra, T. (2020). Measuring Misinformation in Video Search Platforms: An Audit Study on YouTube. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW1), 1–27. https://doi.org/10.1145/3392854

Iyengar, S., & Hahn, Kyu. S. (2009). Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use | Journal of Communication | Oxford Academic. *Journal of Communication*, *59*(1), 19–39.

Jakesch, M., Koren, M., Evtushenko, A., & Naaman, M. (n.d.). *How Partisan Crowds Affect News Evaluation*. 11.

Jost, J. T., & Amodio, D. M. (2012). Political ideology as motivated social cognition: Behavioral and neuroscientific evidence. *Motivation and Emotion*, *36*(1), 55–64. https://doi.org/10.1007/s11031-011-9260-7

Juneja, P., & Mitra, T. (2021). Auditing E-Commerce Platforms for Algorithmically Curated Vaccine Misinformation. *CHI Conference on Human Factors in Computing Systems (CHI'21), May 13-18, 2021, Yokohama, Japan*. https://doi.org/10.1145/3411764.3445250

Just, N., & Latzer, M. (2017). Governance by algorithms: Reality construction by algorithmic selection on the Internet. *Media, Culture & Society*, *39*(2), 238–258. https://doi.org/10.1177/0163443716643157

Kalogeropoulos, A., Suiter, J., Udris, L., & Eisenegger, M. (2019). News media trust and news consumption: Factors related to trust in news in 35 countries. *International Journal of Communication*, *13*.

Kim, H. S., Forquer, H., Rusko, J., Hornik, R. C., & Cappella, J. N. (2016). Selective Exposure to Health Information: The Role of Headline Features in the Choice of Health Newsletter Articles. *Media Psychology*, *19*(4), 614–637. https://doi.org/10.1080/15213269.2015.1090907

Kim, Y. M. (2007). How Intrinsic and Extrinsic Motivations Interact in Selectivity: Investigating the Moderating Effects of Situational Information Processing Goals in Issue Publics'

Web Behavior. *Communication Research*, *34*(2), 185–211. https://doi.org/10.1177/0093650206298069

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society*, *20*(1), 14–29. https://doi.org/10.1080/1369118X.2016.1154087

Klawitter, E., & Hargittai, E. (2018). "It's Like Learning a Whole Other Language": The Role of Algorithmic Skills in the Curation of Creative Goods. *International Journal of Communication*, *12*(0), 21.

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 2015 Internet Measurement Conference*, 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., & Kleinman, S. B. (2011). Preelection Selective Exposure: Confirmation Bias Versus Informational Utility. *Communication Research*. https://journals-sagepub-com.ezproxy.library.wisc.edu/doi/full/10.1177/0093650211400597

Kulshrestha, J., Eslami, M., Messias, J., Zafar, M. B., Ghosh, S., Gummadi, K. P., & Karahalios, K. (2019). Search bias quantification: Investigating political bias in social media and web search. *Information Retrieval Journal*, *22*(1–2), 188–227. https://doi.org/10.1007/s10791-018-9341-2

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Latzer, N. J., Michael. (2016). Governance by algorithms: Reality construction by algorithmic selection on the Internet - Natascha Just, Michael Latzer, 2017. *Media, Culture & Society*. http://journals.sagepub.com/doi/10.1177/0163443716643157

Lau, R. R., & Redlawsk, D. P. (2001). Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making. *American Journal of Political Science*, *45*(4), 951–971. https://doi.org/10.2307/2669334

Le, C., Yang, T., Inchoco, G., & Jones, G. (2021). *The Data Visualizations Behind COVID-19 Skepticism*. The Data Visualizations Behind COVID-19 Skepticism. http://vis.csail.mit.edu/covid-story/

Le, H., Maragh, R., Ekdale, B., High, A., Havens, T., & Shafiq, Z. (2019). Measuring Political Personalization of Google News Search. *The World Wide Web Conference on - WWW '19*, 2957–2963. https://doi.org/10.1145/3308558.3313682

Lee, N.-J., McLeod, D. M., & Dhavan, S. (2008). Framing Policy Debates: Issue Dualism, Journalistic Frames, and Opinions on Controversial Policy Issues—Lee, Douglas M. , Dhavan V. Shah, 2008. *Communication Research, 35*(5), 695–718.

Lewandowski, D. (2017). Is Google Responsible for Providing Fair and Unbiased Results? In M. Taddeo & L. Floridi (Eds.), *The Responsibilities of Online Service Providers* (pp. 61–77). Springer International Publishing. https://doi.org/10.1007/978-3-319-47852-4_4

Li, N., Anderson, A. A., Brossard, D., & Scheufele, D. A. (2014). Channeling Science Information Seekers' Attention? A Content Analysis of Top-Ranked vs. Lower-Ranked Sites in Google. *Journal of Computer-Mediated Communication*, *19*(3), 562–575. https://doi.org/10.1111/jcc4.12043

Liang, H., & Fu, K. (2017). Information Overload, Similarity, and Redundancy: Unsubscribing Information Sources on Twitter. *Journal of Computer-Mediated Communication*, *22*(1), 1–17. https://doi.org/10.1111/jcc4.12178

Lindner, N. M., & Nosek, B. A. (2009). Alienable Speech: Ideological Variations in the Application of Free-Speech Principles. *Political Psychology*, *30*(1), 67–92. https://doi.org/10.1111/j.1467-9221.2008.00681.x

Liu, J., Lee, B., McLeod, D. M., & Choung, H. (2019). Effects of Frame Repetition Through Cues in the Online Environment. *Mass Communication and Society*, *22*(4), 447–465. https://doi.org/10.1080/15205436.2018.1560475

Livingstone, S. (2019). Audiences in an Age of Datafication: Critical Questions for Media Research. *Television & New Media*, *20*(2), 170–183. https://doi.org/10.1177/1527476418811118

Lodge, M., & Taber, C. S. (2013). *The Rationalizing Voter*. Cambridge University Press.

Logg, J. (2017). Theory of Machine: When Do People Rely on Algorithms? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2941774

Lupia, A. (2006). How elitism undermines the study of voter competence. *Critical Review*, *18*(1–3), 217–232. https://doi.org/10.1080/08913810608443658

Lupia, A. (2016). *Uninformed: Why People Know So Little about Politics and what We Can Do about it*. Oxford University Press.

Magin, M., Steiner, M., Heinbach, D., Bosold, S., Pieper, A., Felka, E.-M., & Stark, B. (2015). Suchmaschinen auf dem Prüfstand – eine vergleichende Inhaltsanalyse der Qualität von Trefferlisten [Search engines put to the test – a comparative content analysis of the quality of result pages]. *Medien & Kommunikationswissenschaft*, *63*(4), 495–516. https://doi.org/10.5771/1615-634X-2015-4-495

Makhortykh, M., Urman, A., & Ulloa, R. (2020). How search engines disseminate information about COVID-19 and why they should do better. *Harvard Kennedy School Misinformation Review*, *1*(COVID-19 and Misinformation). https://doi.org/10.37016/mr-2020-017

McMahon, C., Johnson, I., & Hecht, B. (2017). The Substantial Interdependence of Wikipedia and Google: A Case Study on the Relationship Between Peer Production Communities and Information Technologies. *Proceedings of the International AAAI Conference on Web and Social Media*, *11*(1), Article 1. https://ojs.aaai.org/index.php/ICWSM/article/view/14883

Mesch, G. S., & Talmud, I. (2011). *Ethnic Differences in Internet Access: The Role of Occupation and Exposure*. 445–471. https://doi-org.ezproxy.library.wisc.edu/10.1080/1369118X.2011.562218

Metaxa, D., Park, J. S., Landay, J. A., & Hancock, J. (2019). Search Media and Elections: A Longitudinal Investigation of Political Search Results. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–17. https://doi.org/10.1145/3359231

Metaxas, P. T., & Pruksachatkun, Y. (2017). *Manipulation of Search Engine Results during the 2016 US Congressional Elections*. https://repository.wellesley.edu/islandora/object/ir%3A264/

Metzger, M. J., Hartsell, E. H., & Flanagin, A. J. (2015). Cognitive Dissonance or Credibility?: A Comparison of Two Theoretical Explanations for Selective Exposure to Partisan News. *Communication Research*, 009365021561313. https://doi.org/10.1177/0093650215613136

Mittelstadt, B. (2016). Automation, Algorithms, and Politics| Auditing for Transparency in Content Personalization Systems. *International Journal of Communication*, *10*(0), 12.

Moeller, J., Kühne, R., & Vreese, C. D. (2018). Mobilizing Youth in the 21st Century: How Digital Media Use Fosters Civic Duty, Information Efficacy, and Political Participation. *Journal of Broadcasting & Electronic Media*, *62*(3), 445–460. https://doi.org/10.1080/08838151.2018.1451866

Monzer, C., Moeller, J., Helberger, N., & Eskens, S. (2020). User Perspectives on the News Personalisation Process: Agency, Trust and Utility as Building Blocks. *Digital Journalism*, *0*(0), 1–21. https://doi.org/10.1080/21670811.2020.1773291

Mourão, R. R., Thorson, E., Chen, W., & Tham, S. M. (2018). Media Repertoires and News Trust During the Early Trump Administration. *Journalism Studies*, *19*(13), 1945–1956. https://doi.org/10.1080/1461670X.2018.1500492

Mustafaraj, E., Lurie, E., & Devine, C. (2020). *The Case for Voter-Centered Audits of Search Engines During Political Elections*. 11.

Mustafaraj, E., & Walsh, J. (2019). *On Epistemic Responsibility: Web Search and Reputation*. 4.

Mutz, D. C. (2006). *Hearing the Other Side: Deliberative Versus Participatory Democracy*. Cambridge University Press.

Nam, H. H., Jost, J. T., & Bavel, J. J. V. (2013). “Not for All the Tea in China!” Political Ideology and the Avoidance of Dissonance-Arousing Situations. *PLOS ONE*, *8*(4), e59837. https://doi.org/10.1371/journal.pone.0059837

Nechushtai, E., & Lewis, S. C. (2019a). What kind of news gatekeepers do we want machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. *Computers in Human Behavior*, *90*, 298–307. https://doi.org/10.1016/j.chb.2018.07.043

Nechushtai, E., & Lewis, S. C. (2019b). What kind of news gatekeepers do we want machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. *Computers in Human Behavior*, *90*, 298–307. https://doi.org/10.1016/j.chb.2018.07.043

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D., & Nielsen, R. K. (2017). *Reuters Institute Digital News Report 2017* (SSRN Scholarly Paper ID 3026082). Social Science Research Network. https://papers.ssrn.com/abstract=3026082

Newman, N., Fletcher, R., Schulz, A., Andi, S., & Nielsen, R. K. (2020). *Reuters Institute Digital News Report 2020*. Reuters Institute, University of Oxford.

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.

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

Olsson, T., Samuelsson, U., & Viscovi, D. (2019). At risk of exclusion? Degrees of ICT access and literacy among senior citizens. *Information, Communication & Society*, *22*(1), 55–72. https://doi.org/10.1080/1369118X.2017.1355007

Ono, H., & Zavodny, M. (2003). Gender and the Internet*. *Social Science Quarterly*, *84*(1), 111–121. https://doi.org/10.1111/1540-6237.t01-1-8401007

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

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Pariser, E. (2011). *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. Penguin.

Parissi, M., Komis, V., Lavidas, K., Dumouchel, G., & Karsenti, T. (2019). A pre-post study to assess the impact of an information-problem solving intervention on university students' perceptions and self-efficacy towards search engines. *International Journal of Technologies in Higher Education*, *16*(1), 68–87.

Park, C. S. (2019). Does Too Much News on Social Media Discourage News Seeking? Mediating Role of News Efficacy Between Perceived News Overload and News Avoidance on Social Media. *Social Media + Society*, *5*(3), 2056305119872956. https://doi.org/10.1177/2056305119872956

Pasquale, F. (2015). *The Black Box Society*. Harvard University Press.

Peacock, C., Hoewe, J., Panek, E., & Willis, G. P. (2021). Hyperpartisan News Use: Relationships with Partisanship and Cognitive and Affective Involvement. *Mass Communication and Society*, *24*(2), 210–232. https://doi.org/10.1080/15205436.2020.1844902

Pew Research Center. (2017, September 11). The elements of the information-engagement typology. *Pew Research Center: Internet, Science & Tech*. https://www.pewresearch.org/internet/2017/09/11/the-elements-of-the-information-engagement-typology/

Pew Research Center. (2019, October 9). Americans and Digital Knowledge. *Pew Research Center: Internet, Science & Tech*. https://www.pewresearch.org/internet/2019/10/09/americans-and-digital-knowledge/

Pew Research Center. (2020, August 19). Most Americans Think Social Media Sites Censor Political Viewpoints. *Pew Research Center: Internet, Science & Tech*. https://www.pewresearch.org/internet/2020/08/19/most-americans-think-social-media-sites-censor-political-viewpoints/

Pierce, D. R., Redlawsk, D. P., & Cohen, W. W. (2017). Social Influences on Online Political Information Search and Evaluation. *Political Behavior*, *39*(3), 651–673. https://doi.org/10.1007/s11109-016-9374-4

Puschmann, C. (2018). Beyond the Bubble: Assessing the Diversity of Political Search Results. *Digital Journalism*, 1–20. https://doi.org/10.1080/21670811.2018.1539626

Qiu, X., Oliveira, D. F. M., Shirazi, A. S., Flammini, A., & Menczer, F. (2017). Limited individual attention and online virality of low-quality information. *ArXiv:1701.02694 [Physics]*. http://arxiv.org/abs/1701.02694

Rader, E., & Gray, R. (2015). Understanding User Beliefs About Algorithmic Curation in the Facebook News Feed. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 173–182. https://doi.org/10.1145/2702123.2702174

Rele, R. S., & Duchowski, A. T. (2005). Using Eye Tracking to Evaluate Alternative Search Results Interfaces. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *49*(15), 1459–1463. https://doi.org/10.1177/154193120504901508

Resnick, M. L., Maldonado, C., Santos, J. M., & Lergier, R. (2001). Modeling On-Line Search Behavior Using Alternative Output Structures. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *45*(15), 1166–1170. https://doi.org/10.1177/154193120104501503

Ripberger, J. T. (2011). Capturing Curiosity: Using Internet Search Trends to Measure Public Attentiveness. *Policy Studies Journal*, *39*(2), 239–259. https://doi.org/10.1111/j.1541-0072.2011.00406.x

Robertson, R. E., Jiang, S., Joseph, K., Friedland, L., Lazer, D., & Wilson, C. (2018). Auditing Partisan Audience Bias within Google Search. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 1–22. https://doi.org/10.1145/3274417

Robertson, R. E., Lazer, D., & Wilson, C. (2018). Auditing the Personalization and Composition of Politically-Related Search Engine Results Pages. *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*, 955–965. https://doi.org/10.1145/3178876.3186143

Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., Schulz, J., Hale, T. M., & Stern, M. J. (2015). Digital inequalities and why they matter. *Information,* *Communication & Society*, *18*(5), 569–582. https://doi.org/10.1080/1369118X.2015.1012532

Rudolph, T. J. (2006). Triangulating Political Responsibility: The Motivated Formation of Responsibility Judgments. *Political Psychology*, *27*(1), 99–122. https://doi.org/10.1111/j.1467-9221.2006.00451.x

Scheufele, D. A., Nisbet, M. C., Brossard, D., & Nisbet, E. C. (2004). Social Structure and Citizenship: Examining the Impacts of Social Setting, Network Heterogeneity, and Informational Variables on Political Participation. *Political Communication*, *21*(3), 315–338. https://doi.org/10.1080/10584600490481389

Shah, D. V., Watts, M., D., Domke, D., & Fan, D., P. (2002). News Framing and Cueing of Issue Regimes: Explaining Clinton's Public Approval in Spite of Scandal*. *Public Opinion Quarterly*, *66*(3), 339–370. https://doi.org/10.1086/341396

Singer, J., B. (2013). *User-generated visibility: Secondary gatekeeping in a shared media space*. https://journals-sagepub-com.ezproxy.library.wisc.edu/doi/full/10.1177/1461444813477833

Skovsgaard, M., & Andersen, K. (2020). Conceptualizing News Avoidance: Towards a Shared Understanding of Different Causes and Potential Solutions. *Journalism Studies*, *21*(4), 459–476. https://doi.org/10.1080/1461670X.2019.1686410

Slechten, L., Courtois, C., Coenen, L., & Zaman, B. (2021). Adapting the Selective Exposure Perspective to Algorithmically Governed Platforms: The Case of Google Search. *Communication Research*. https://journals-sagepub-com.ezproxy.library.wisc.edu/doi/full/10.1177/00936502211012154

Spink, A., & Cole, C. (2006). Human information behavior: Integrating diverse approaches and information use. *Journal of the American Society for Information Science and Technology*, *57*(1), 25–35. https://doi.org/10.1002/asi.20249

Statcounter. (2021). *Search Engine Market Share Worldwide*. StatCounter Global Stats. https://gs.statcounter.com/search-engine-market-share

Steiner, M., Magin, M., Stark, B., & Geiß, S. (2020). Seek and you shall find? A content analysis on the diversity of five search engines' results on political queries. *Information, Communication & Society*, *0*(0), 1–25. https://doi.org/10.1080/1369118X.2020.1776367

Sterrett, D., Malato, D., Benz, J., Kantor, L., Tompson, T., Rosenstiel, T., Sonderman, J., & Loker, K. (2019). Who Shared It?: Deciding What News to Trust on Social Media. *Digital Journalism*, *7*(6), 783–801. https://doi.org/10.1080/21670811.2019.1623702

Strömbäck, J., Tsfati, Y., Boomgaarden, H., Damstra, A., Lindgren, E., Vliegenthart, R., & Lindholm, T. (2020). News media trust and its impact on media use: Toward a framework for future research. *Annals of the International Communication Association*, *44*(2), 139–156. https://doi.org/10.1080/23808985.2020.1755338

Stroud, N. J. (2011). *Niche News: The Politics of News Choice*. Oxford University Press, USA.

Stroud, N. J. (2017). Selective Exposure Theories. In *The Oxford Handbook of Political Communication*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199793471.013.009_update_001

Sullivan, D. (2021, March 24). *Google Search sends more traffic to the open web every year*. Google. https://blog.google/products/search/google-search-sends-more-traffic-open-web-every-year/

Sun, W., Nasraoui, O., & Shafto, P. (2020). Evolution and impact of bias in human and machine learning algorithm interaction. *PLOS ONE*, *15*(8), e0235502. https://doi.org/10.1371/journal.pone.0235502

Suzuki, M., & Yamamoto, Y. (2020). *Analysis of Relationship between Confirmation Bias and Web Search Behavior*. 8.

Taylor, E., Neudert, L.-M., Hoffmann, S., & Howard, P. N. (2020). Follow the Money: How the Online Advertising Ecosystem Funds COVID-19 Junk News and Disinformation. *Computational Propaganda Project (COMPROP) Working Paper 2020.1., Oxford Internet Institute, University of Oxford*, 8.

Thompson, L. F., Meriac, J., P., & Cope, J., G. (2002). Motivating Online Performance: The Influences of Goal Setting and Internet Self-Efficacy. *Social Science Computer Review*. https://journals-sagepub-com.ezproxy.library.wisc.edu/doi/abs/10.1177/089443930202000205

Trevisan, F., Hoskins, A., Oates, S., & Mahlouly, D. (2018). The Google Voter: Search Engines and Elections in the New Media Ecology.”. *Information, Communication & Society*, *21*, 111.

Trielli, D., & Diakopoulos, N. (2020). Partisan search behavior and Google results in the 2018 U.S. midterm elections. *Information, Communication & Society*, 1–17. https://doi.org/10.1080/1369118X.2020.1764605

Trielli, D., & Diakopoulos, N. (2019). Search as News Curator: The Role of Google in Shaping Attention to News Information. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI ’19*, 1–15. https://doi.org/10.1145/3290605.3300683

Tripodi, F. (2018). Searching for Alternative Facts. *2018*, 64. https://datasociety.net/wp-content/uploads/2018/05/Data_Society_Searching-for-Alternative-Facts.pdf

Turcotte, J., York, C., Irving, J., Scholl, R. M., & Pingree, R. J. (2015). News Recommendations from Social Media Opinion Leaders: Effects on Media Trust and Information Seeking. *Journal of Computer-Mediated Communication*, *20*(5), 520–535. https://doi.org/10.1111/jcc4.12127

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Unkel, J., & Haim, M. (2019). Googling Politics: Parties, Sources, and Issue Ownerships on Google in the 2017 German Federal Election Campaign. *Social Science Computer Review*, 0894439319881634. https://doi.org/10.1177/0894439319881634

Urman, A., Makhortykh, M., & Ulloa, R. (2021). The Matter of Chance: Auditing Web Search Results Related to the 2020 U.S. Presidential Primary Elections Across Six Search Engines. *Social Science Computer Review*, 08944393211006863. https://doi.org/10.1177/08944393211006863

Van Couvering, E. (2007). Is Relevance Relevant? Market, Science, and War: Discourses of Search Engine Quality. *Journal of Computer-Mediated Communication*, *12*(3), 866–887. https://doi.org/10.1111/j.1083-6101.2007.00354.x

van, D. A. J. A. M., & Helsper, E. J. (2015). The Third-Level Digital Divide: Who Benefits Most from Being Online? In *Communication and Information Technologies Annual* (Vol. 10, pp. 29–52). Emerald Group Publishing Limited. https://doi.org/10.1108/S2050-206020150000010002

Wallace, J. (2018). Modelling Contemporary Gatekeeping. *Digital Journalism*, *6*(3), 274–293. https://doi.org/10.1080/21670811.2017.1343648

Weber, I., & Jaimes, A. (2011). Who uses web search for what: And how. *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, 15–24. https://doi.org/10.1145/1935826.1935839

Weeks, B. E., Lane, D. S., Kim, D. H., Lee, S. S., & Kwak, N. (2017). Incidental Exposure, Selective Exposure, and Political Information Sharing: Integrating Online Exposure Patterns and Expression on Social Media. *Journal of Computer-Mediated Communication*, *22*(6), 363–379. https://doi.org/10.1111/jcc4.12199

Westerwick, A., Johnson, B. K., & Knobloch-Westerwick, S. (2017a). Confirmation biases in selective exposure to political online information: Source bias vs. content bias. *Communication Monographs*, *84*(3), 343–364. https://doi.org/10.1080/03637751.2016.1272761

Westerwick, A., Johnson, B. K., & Knobloch-Westerwick, S. (2017b). Confirmation biases in selective exposure to political online information: Source bias vs. content bias. *Communication Monographs*, *84*(3), 343–364. https://doi.org/10.1080/03637751.2016.1272761

White, R. (2013). Beliefs and biases in web search. *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 3–12. https://doi.org/10.1145/2484028.2484053

White, R. W., Dumais, S. T., & Teevan, J. (2009). Characterizing the influence of domain expertise on web search behavior. *Proceedings of the Second ACM International*

Conference on Web Search and Data Mining, 132–141. https://doi.org/10.1145/1498759.1498819

White, R. W., & Horvitz, E. (2015). Belief Dynamics and Biases in Web Search. *ACM Transactions on Information Systems*, *33*(4), 18:1-18:46. https://doi.org/10.1145/2746229

Whyte, C. E. (2016). Thinking inside the (black) box: Agenda setting, information seeking, and the marketplace of ideas in the 2012 presidential election. *New Media & Society*, *18*(8), 1680–1697. aph.

Wickens, C. M., & Miller, T. (2020). Gender, Digital Literacies, and Higher Education. In *The Wiley Handbook of Gender Equity in Higher Education* (pp. 53–67). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119257639.ch3

Williams, A. E. (2012). Trust or Bust?: Questioning the Relationship Between Media Trust and News Attention. *Journal of Broadcasting & Electronic Media*, *56*(1), 116–131. https://doi.org/10.1080/08838151.2011.651186

Williams, R., & Edge, D. (1996). The social shaping of technology. *Research Policy*, *25*(6), 865–899. https://doi.org/10.1016/0048-7333(96)00885-2

Wirth, W., Böcking, T., Karnowski, V., & Pape, T. V. (2007). Heuristic and Systematic Use of Search Engines. *Journal of Computer-Mediated Communication*, *12*(3), 778–800. https://doi.org/10.1111/j.1083-6101.2007.00350.x

Witte, J. C., & Mannon, S. E. (2010). *The Internet and Social Inequalitites*. Routledge.

Wolak, J. (2018). Feelings of Political Efficacy in the Fifty States. *Political Behavior*, *40*(3), 763–784. https://doi.org/10.1007/s11109-017-9421-9

Xiao, X., Su, Y., & Lee, D. K. L. (2021). Who Consumes New Media Content More Wisely? Examining Personality Factors, SNS Use, and New Media Literacy in the Era of

Misinformation. *Social Media + Society*, *7*(1), 2056305121990635. https://doi.org/10.1177/2056305121990635

Xu, W. W., Sang, Y., & Kim, C. (2020). What Drives Hyper-Partisan News Sharing: Exploring the Role of Source, Style, and Content. *Digital Journalism*, *8*(4), 486–505. https://doi.org/10.1080/21670811.2020.1761264

Yamamoto, T., Yamamoto, Y., & Fujita, S. (2018). Exploring People's Attitudes and Behaviors Toward Careful Information Seeking in Web Search. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 963–972. https://doi.org/10.1145/3269206.3271799

Yang, J. Z., & Zhuang, J. (2020). Information Seeking and Information Sharing Related to Hurricane Harvey. *Journalism & Mass Communication Quarterly*, *97*(4), 1054–1079. https://doi.org/10.1177/1077699019887675

Zaller, J. R., & R, Z. J. (1992). *The Nature and Origins of Mass Opinion*. Cambridge University Press.

Zúñiga, A. A.-A., Homero Gil de. (2016). Effects of Editorial Media Bias Perception and Media Trust on the Use of Traditional, Citizen, and Social Media News—Alberto Ardèvol-Abreu, Homero Gil de Zúñiga, 2017. *Journalism & Mass Communication Quarterly*. http://journals.sagepub.com/doi/10.1177/1077699016654684

APPENDIX

A1. Methodology: Extraction of biased phrases from partisan media

MediaCloud was used to crawl two collections of right-leaning (including 471 outlets) and left-leaning (including 175 outlets) online media outlets compiled by Buzzfeed, using queries related to each topic, e.g. “climate change”, “abortion”, “immigration” in the last five years. MediaCloud then returned the number of articles associated with the queries. MediaCloud also returned the top words, bigrams (two-word phrases) or trigrams (three-word phrases) used with each query, i.e. the words or word combinations that showed up more often in a random subset of the entire corpus. For example, MediaCloud would return 432,323 articles associated with the query “gun control” in the Buzzfeed right-leaning media sources, then randomly selected 10,000 articles and extracted the n-grams from this smaller pool of 10,000 articles.

The data collection pipeline can be summarized as follows: Step 1) the query/topic was specified, e.g., “climate change”, Step 2) MediaCloud was queried to return the number of articles including “climate change” in the right-leaning and left-leaning media sources (here, right/left-leaning media sources are based on Buzzfeed categorization), Step 3) MediaCloud randomly sampled 10,000 articles from each collection and returned the top n-grams in each. Then, based on these subsets, the relative frequency to determine the right-leaning or left-leaning bias of each n-gram was calculated, i.e. a term was labelled “right-leaning” if it was more frequently found in the randomized 10,000 articles in the right-leaning corpus than in the randomized 10,000 articles in the left-leaning corpus.

The final terms were later validated by crowdsourcing 100 American users on Prolific (an online research platform) to rate whether these terms had a left-leaning, neutral, or right-leaning slant. The results indicate that there were largely overlaps between human ratings of the political bias of the terms and their relative frequencies in the two corpora. However, some n-grams were rated differently by crowdsourced workers compared to their proportional appearance in the left- and right-leaning media language. For the parsimony of the study, only n-grams that were rated uniformly by both criteria as right-leaning or left-leaning were included in the final survey.

A2. Descriptive statistics of biased term preference

	N	Minimum	Maximum	Mean	SD
ArmedSelfDefense	179	2	10	7.64	1.99
GunLicense	257	2	10	7.9	1.68
BackgroundChecks	306	1	10	8.26	1.58
GunLobby	111	1	10	7.23	2.08
GunAccessibility	234	1	10	7.83	1.75
NRA	230	1	10	7.59	2.05
SecondAmendment	310	1	10	8.55	1.73
GunControlSolutions	278	1	10	8.47	1.83
TerrorThreats	86	1	10	6.67	2.51
AntiGunAgenda	113	1	10	7.47	2.26
RefugeeAsylum	213	2	10	8.28	1.61
CompreImmigrationReform	242	1	10	8.5	1.67
IllegalAliens	171	2	10	8.61	1.8
AlienInvaders	47	1	10	7.34	2.73
DACAAmnesty	198	1	10	7.84	1.77
RadicalIslam	53	1	10	7.38	2.45
UndocumentedImmigrants	340	4	10	8.77	1.38
HomelandBorders	179	3	10	8.25	1.69
FamilyDetention	197	1	10	8.21	1.61
RefugeeAdmissionLimits	136	4	10	7.84	1.51
C_impact	363	1	10	8.77	1.65
C_consensus	133	2	10	7.83	1.86
C_skeptics	116	2	10	7.32	2.02
C_agenda	187	3	10	8.26	1.51
C_footprint	267	2	10	8.33	1.43
C_fraud	120	1	10	7.78	2.27
C_hysteria	96	1	10	7.7	2.34
C_crisis	296	1	10	8.69	1.51
C_deniers	88	1	10	7.26	1.85
C_hoax	135	1	10	8.3	2.23
A_ondemand	93	1	10	7.34	2.38
LiveBirthAbortion	93	1	10	7.22	2.56
UnintendedPregnancy	163	2	10	7.77	1.65
ElectiveAbortion	115	1	10	7.47	2.21
ProChoice	303	1	10	8.86	1.47
ProLife	237	1	10	8.89	1.62
WomensRights	289	3	10	8.45	1.53

LateTermAbortion	201	1	10	7.94	1.85
FetalAbnormality	81	2	10	7.06	2.08
ReproductiveFreedom	235	1	10	8.24	1.45
Liberal_abortion	57	2.75	8.5	7.4	1.31
Conser_abortion	33	3.25	8.5	6.45	2
Neutral_abortion	53	3	9.5	6.81	1.69
Liberal_climate	34	4.5	8.5	6.66	1.14
Conser_climate	35	3	8.5	6.55	2.16
Neutral_climate	113	1.5	9.5	7.9	1.76
Liberal_gun	44	3.25	8.5	6.9	1.13
Conser_gun	34	2.5	8.5	6.16	1.8
Neutral_gun	48	2.5	9.5	6.02	1.78
Liberal_immi	63	3.5	8.5	7.36	1.17
Conser_immi	14	3.25	8.5	6.21	1.82
Neutral_immi	79	5	9.5	7.7	1.17

A3. List of domains and occurrence rate (OR) in search results – Election Integrity

C/Qc	OR	L/Qc	OR	C/Ql	OR	L/Ql	OR
11alive	11.7	11alive	12.2	cnn	9.5	history	12.0
clickondetroit	6.4	foxnews	6.9	usatoday	8.5	nbcnews	11.6
bridgemi	6.0	clickondetroit	6.5	nbcnews	8.4	washingtonpost	11.0
forbes	5.9	forbes	6.4	nytimes	8.0	cnn	8.7
wtoc	5.2	bridgemi	6.2	history	7.8	usatoday	8.0
npr	5.1	politifact	6.1	clickondetroit	6.4	npr	7.8
foxnews	4.9	northwestgeorgianews	6.1	ntdaily	6.3	ntdaily	6.1
wabe	4.9	factcheck	5.8	npr	6.1	thehill	5.7
12news	4.5	fox2detroit	5.7	washingtonpost	5.6	malibutimes	5.0
factcheck	4.0	washingtonpost	5.5	thehill	4.9	businessinsider	4.1
pbs	3.3	wsbtv	3.9	ramaponews	4.8	ramaponews	3.5
burlingtonfreepress	2.8	pbs	3.7	cbsnews	2.3	nytimes	3.4
abc7ny	2.3	9and10news	3.4	malibutimes	2.3	clickondetroit	3.0
clickorlando	2.2	nytimes	3.1	ft.com	2.0	jonesborosun	2.7
usatoday	1.8	courier-journal	2.4	wikipedia	1.5	wikipedia	1.8
nbcnews	1.7	thehill	2.3	ndtv	1.5	cbsnews	0.7
newsweek	1.7	wtoc	2.2	wjhl	1.3	magicvalley	0.7
nytimes	1.5	texastribune	2.0	10tv	1.1	nationalreview	0.7
tribuneindia	1.4	cnn	1.2	businessinsider	1.0	amazon	0.5
fox5atlanta	1.4	apnews	1.1	wreg	1.0	loc.gov	0.5
politifact	1.3	12news	0.8	tass	0.8	billofrightsinstitute	0.5
washingtonpost	1.3	startribune	0.7	observer.case.edu	0.8	twitter	0.4
texastribune	1.3	theguardian	0.7	wspa	0.8	channel3000	0.4
fox2detroit	1.1	ballotpedia	0.7	katc	0.6	bostonglobe	0.4
apnews	1.1	brennancenter	0.5	wric	0.6	abcnews	0.3
wisconsinexaminer.com	1.0	heritage	0.5	loc.gov	0.6	orlandosentinel	0.3
cnn	1.0	ncsl	0.5	billofrightsinstitute	0.6	tass	0.1
courier-journal	1.0	whitehouse.gov	0.4	vox	0.5	journalistsresource.org	0.1
northwestgeorgianews	1.0	bbc	0.4	jonesborosun	0.5		
ballotpedia	1.0	politico	0.4	nymag	0.4		
9and10news	0.8	wikipedia	0.4	wabe.org	0.4		
wsbtv	0.8	usatoday	0.3	magicvalley	0.4		
gpb.org	0.8	reuters	0.3	springfieldnewssun	0.4		
jsonline.com	0.8	nbcnews	0.3	wkyc	0.4		
theguardian	0.5	jsonline	0.3	localmemphis	0.4		
brennancenter	0.5	tribuneindia	0.1	orlandosentinel	0.3		

heritage	0.5	amazon	0.3
wtvm	0.5	youtube	0.3
ncsl	0.5	twitter	0.1
thehill	0.4	channel3000	0.1
politico	0.4	cnbc	0.1
sos.wa.gov	0.3	axios	0.1
gallopade.com	0.3	politico	0.1
scholastic.com	0.3	myfox8	0.1
teacherspayteachers.com	0.3	wfla	0.1
burlingtonfreepress.com	0.3	historians.org	0.1
whitehouse.gov	0.3	wthr	0.1
wxyz.com	0.3		
qz.com	0.3		
eac.gov	0.3		
wikipedia	0.3		
news4jax	0.3		
ajc.com	0.3		
sos.ca.gov	0.3		
nymag	0.1		
woodtv	0.1		
browardsoe.org	0.1		
wtsp.com	0.1		

Notes: OR (i.e. occurrence rate) was calculated as the rate of domain occurrence by the total number of unique URLs.

A4. List of domains and occurrence rate (OR) in search results – Abortion

C/Qc	OR	L/Qc	OR	C/Ql	OR	L/Ql	OR
guttmacher.org	6.8	guttmacher.org	6.5	amazon.com	4.7	ohchr.org	6.5
princeton.edu	6.6	wikipedia.org	6.1	wikipedia.org	4.5	amazon.com	6.5
congress.gov	5.5	congress.gov	6.1	ohchr.org	4.0	wikipedia.org	6.1
wikipedia.com	5.3	princeton.edu	5.9	prochoiceamerica.org	3.0	prochoiceamerica.org	4.6
csulb.edu	3.6	medicine.missouri.edu	3.4	aclu.org	3.0	aclu.org	4.4
medicine.missouri.edu	3.6	csulb.edu	3.2	amnesty.org	2.9	beta.reproductiverights.org	4.0
cdlex.org	3.4	theatlantic.com	3.2	bbc.com	2.9	nbcnews.com	3.8
bc.edu	3.4	cdlex.org	3.2	wrj.org	2.7	amnesty.org	3.6
bbc.com	3.4	apa.org	3.2	beta.reproductiverights.org	2.7	civilrights.org	3.6
ncbi.nlm.nih.gov	3.4	ncbi.nlm.nih.gov	3.2	nbcnews.com	2.7	pubmed.ncbi.nlm.nih.gov	3.6
acpeds.org	3.4	plannedparenthood.com	3.2	pubmed.ncbi.nlm.nih.gov	2.7	azcentral.com	3.6
jme.bmj.com	3.2	psychiatryadvisor.com	3.2	yang2020.com	2.7	yang2020.com	3.3
apa.org	3.2	innovating-education.org	3.2	shondaland.com	2.7	bbc.com	3.3
psychiatryadvisor.com	3.2	bmcmedicine.biomedcentral.com	3.2	findlaw.com	2.7	wrj.org	3.3
naapc.org	3.2	acpeds.org	3.2	statusofwomendata.org	2.7	shreveporttimes.com	3.3
jstor.org	3.0	bc.edu	3.0	civilrights.org	2.7	brookings.edu	3.3
harpersbazaar.com	3.0	bbc.com	3.0	azcentral.com	2.7	shondaland.com	3.3
plannedparenthood.com	3.0	masscitizensforlife.org	3.0	shreveporttimes.com	2.5	link.springer.com	3.3
masscitizensforlife.org	3.0	naapc.org	3.0	brookings.edu	2.5	statusofwomendata.org	3.3
theatlantic.com	2.8	marchforlife.org	3.0	link.springer.com	2.5	findlaw.com	3.3
innovating-education.org	2.8	jme.bmj.com	2.8	thirdway.org	2.5	thirdway.org	3.1
bmcmedicine.biomedcentral.com	2.6	lozierinstitute.org	2.8	echopress.com	2.5	echopress.com	3.1
quillette.com	2.3	csus.edu	2.5	hli.org	2.5	ibisreproductivehealth.org	2.9
csus.edu	2.1	lagunatreatment.com	2.5	ibisreproductivehealth.org	2.4	newsweek.com	2.1
marchforlife.org	2.1	harpersbazaar.com	2.3	newsweek.com	2.4	acluaz.org	1.5
lagunatreatment.com	1.9	quillette.com	2.3	jstor.org	2.0	nwlc.org	1.3
hli.org	1.5	focusonthefamily.com	2.1	acluaz.org	2.0	pbs.org	0.8
lozierinstitute.org	1.3	drjamesdobson.org	1.7	rememberingactivism.eu	2.0	rememberingactivism.eu	0.8
h3helpline.org	1.1	adflegal.org	1.3	hrw.org	2.0	voanews.com	0.6

focusonthefamily.com	0.6	academy4sc.org	1.3	pbs.org	1.5	opensocietyfoundations.org	0.6
amnesty.org	0.6	hli.org	0.4	usatoday.com	1.5	kcbd.com	0.4
all.org	0.6	ibisreproductivehealth.org	0.4	voanews.com	1.2	jstor.org	0.2
drjamesdobson.org	0.4	uffl.org	0.4	nwlc.org	1.0	elizabethwarren.com	0.2
hurtafterabortion.com	0.4	whatisessential.org	0.4	theglobeandmail.com	0.8	guttmacher.org	0.2
academy4sc.org	0.4	cdohope.org	0.2	politico.com	0.8	commondreams.org	0.2
reproductiverights.org	0.4	ansirh.org	0.2	opensocietyfoundations.org	0.8	kake.com	0.2
adflegal.org	0.2	h3helpline.org	0.2	donate.pai.org	0.8	donate.pai.org	0.2
whatisessential.org	0.2	all.org	0.2	greenwichfreepress.com	0.7	hrw.org	0.2
educarhoy.org	0.2	rewire.news	0.2	quillette.com	0.7	endangeredspeciescondoms.com	0.2
statusofwomendata.org	0.2			feministsforlife.org	0.5	lifeteen.com	0.2
ohchr.org	0.2			npr.org	0.5	rewire.news	0.2
aclu.org	0.2			theatlantic.com	0.5	vice.com	0.2
nwlc.org	0.2			bostonreview.net	0.5	prolifereplies.liveaction.org	0.2
opensocietyfoundations.org	0.2			lithub.com	0.5		
cdohope.org	0.2			ksnt.com	0.5		
ibisreproductivehealth.org	0.2			cbsnews.com	0.5		
echopress.com	0.2			thecut.com	0.5		
nytimes.com	0.2			plan.international.org	0.3		
prolifeacrossamerica.org	0.2			innovating-education.org	0.3		
				lagunatreatment.com	0.3		
				guttmacher.org	0.3		
				now.org	0.3		
				plannedparenthood.org	0.3		
				wholewomanshealthalliance.org	0.3		

Notes: OR (i.e. occurrence rate) was calculated as the rate of domain occurrence by the total number of unique URLs.

A5. List of domains and occurrence rate (OR) in search results – Climate Change

C/Qc	OR	L/Qc	OR	C/Ql	OR	L/Ql	OR
wwf.org.uk	6.1	wwf.org.uk	6.7	climate.gov	11.4	climate.gov	11.6
edf.org	5.9	edf.org	6.4	climate.nasa.gov	8.6	climate.nasa.gov	8.6
theguardian.com	3.8	cei.org	3.6	edf.org	6.2	edf.org	6.3
lavoisier.com.au	3.6	climate.nasa.gov	3.6	europa.eu	6.2	ucsusa.org	6.0
theconversation.com	3.6	apnews.com	3.4	ucsusa.org	5.7	europa.eu	6.0
blogs.ei.columbia.edu	3.6	theguardian.com	3.4	scientificamerican.com	4.2	scientificamerican.com	4.9
farmprogress.com	3.4	abcnews.go.com	3.4	beforetheflood.com	3.7	independent.org	3.2
cei.org	3.4	junkscience.com	3.4	cordis.europa.eu	3.3	beforetheflood.com	3.0
huffpost.com	3.4	whistleblower.org	3.4	worldwildlife.org	3.1	climatecentral.org	3.0
forbes.com	3.4	forbes.com	3.4	carbonbrief.org	3.1	nationalgeographic.com	3.0
georgiapolicy.org	3.4	youtube.com	3.4	cleanet.org	3.1	worldwildlife.org	3.0
climate.nasa.gov	3.4	georgiapolicy.org	3.4	wikipedia.org	3.1	carbonbrief.org	3.0
northsidesun.com	3.4	heartland.org	3.4	news.mongabay.com	3.1	cordis.europa.eu	3.0
abcnews.go.com	3.2	terrapass.com	3.4	theconversation.com	3.1	wikipedia.org	3.0
edberry.com	3.2	theconversation.com	3.4	royalsocietypublishing.org	3.1	news.mongabay.com	3.0
greenpeace.org	3.2	carbonbrief.org	3.4	nationalgeographic.com	2.9	iberdrola.com	3.0
terrapass.com	3.2	channelnewsasia.com	3.4	climatecentral.org	2.9	royalsocietypublishing.org	3.0
carbonbrief.org	3.2	bundestag.de	3.4	independent.co.uk	2.9	who.int	2.8
channelnewsasia.com	3.2	farmprogress.com	3.1	iberdrola.com	2.9	cleanet.org	2.8
apnews.com	3.0	edberry.com	3.1	usatoday.com	2.7	cgd.ucar.edu	2.8
junkscience.com	3.0	lavoisier.com.au	3.1	cgd.ucar.edu	2.6	usatoday.com	2.8
youtube.com	3.0	iwmc.org	3.1	science.org.au	2.6	theconversation.com	2.8
heartland.org	3.0	northsidesun.com	3.1	who.int	2.0	science.org.au	2.6
bundestag.de	3.0	blogs.ei.columbia.edu	3.1	pnas.org	1.8	pnas.org	2.6
whistleblower.org	2.8	greenpeace.org	2.8	theguardian.com	1.3	scholar.google.com	1.4
vanityfair.com	2.8	huffpost.com	2.6	washingtonpost.com	1.1	theguardian.com	0.7
iwmc.org	2.4	wattsupwiththat.com	2.3	nytimes.com	1.1	nytimes.com	0.5

argusleader.com	2.0	vanityfair.com	2.3	scholar.google.com	0.7	washingtonpost.com	0.2
wattsupwiththat.com	1.4	argusleader.com	1.5	eea.europa.eu	0.5	ei.lehigh.edu	0.2
thenation.com	1.0	fauquiernow.com	0.5	facebook.com	0.2	19january2017snapshot.epa.gov	0.2
wikipedia.org	0.6	thenation.com	0.5	19january2017snapshot.epa.gov	0.2	clintonfoundation.org	0.2
independent.org	0.6	cbsnews.com	0.5	yaleclimateconnections.org	0.2	rainforest-alliance.org	0.2
insideclimatenews.org	0.6	amazon.com	0.3	wwf.panda.org	0.2	eurekalert.org	0.2
amazon.com	0.4	hcn.org	0.3			lavoisier.com.au	0.2
fauquiernow.com	0.4	aycc.org.au	0.3				
pnas.org	0.2	skepticalscience.com	0.3				
telegraph.co.uk	0.2						
rainforest-alliance.org	0.2						
cbsnews.com	0.2						
berkeleyearth.org	0.2						
aycc.org.au	0.2						

Notes: OR (i.e. occurrence rate) was calculated as the rate of domain occurrence by the total number of unique URLs.