What Drives the Selection of Political Information in Google? Tension Between Ideal Democracy and the Influence of Ranking

Zumofen Guillaumea\* and M. Bühlmannb

aInstitute of Political Science, Année Politique Suisse (APS), Bern University, Bern, Switzerland;

bInstitute of Political Science, Année Politique Suisse (APS), Bern University, Bern, Switzerland;

Zumofen Guillaume

Bern University

Fabrikstrasse 8,

CH-3012 Bern,

+41 79 389 13 05,

guillaume.zumofen@ipw.unibe.ch

ORCID: 0000-0002-9047-9099

@anneepolitique

Zumofen Guillaume is a PhD student at the Institute of Political Science, University of Bern, and research assistant of the 'Année Politique Suisse'. His research interests include political communication, political psychology, voting behavior, online media, and referendum campaign.

Prof. Dr. Marc Bühlmann is director of the Swiss Political Yearbook 'Année Politique Suisse' at the Institute of Political Science at Bern University. He has published on political sociology, electoral and participation behavior, methods, social capital, democratic theory, and comparative politics.

Ranking or Information sources– What Drives the Selection of Political Information in Google?

The emergence of the Internet has altered how individuals obtain information – this also applies to political information. Search engines have taken over the role of political information gatekeepers becoming thus key players in a democracy. Through an online survey experiment, which mimicked a Google web interface, this article examines how citizens select political information in a political campaign from a Google search result page. The results suggest that mostly ranking, but also textual cues from snippets drive an individual’s selection of political information in Google. We discuss these findings from a democratic theory perspective.

Keywords: Search engines, information selection, political campaign, democracy, selective exposure, filter bubble, Google.

# Introduction

Information about politics is crucial for a functioning democracy. In his seminal work on democratic theory, Dahl (1989) has crystallized the five principles of an ideal democracy. One of the principles states that individuals must have access to information to reach informed and enlightened decisions that serve their best personal interest. Although a broad consensus exists on the significance of political information in a democracy, there is no agreement about how it should be provided or how it is processed by individuals.

Recently, there has been growing interest in how the Internet amends democracy, notably regarding political information. Xenos et al. (2018) and Mitchell and Holcomb (2016) note a generational shift from traditional (i.e., offline) to online media, which has driven more people to seek information via the Internet. For the first time, an information structure incorporates almost all existing information available in one place (Schroeder, 2018). A transition of the information environment from low- to high-choice and from push to pull has created an easily accessible, unlimited information supply (Hargittai et al., 2012; Hilbert and Lopez, 2011; Neumann et al., 2012). This information revolution has become a promise for enlightened decision making (Hindman, 2009) and, accordingly, for improving democracy through more informed citizen participation. Facing the information overload, people began to use search engines as a compass to navigate through the overwhelming amount of available information (Lee et al., 2016; Pan et al., 2007; Scharkow and Vogelgesang, 2011). Thus, search engines take over the role of political information gatekeepers and become key players in a democracy (Dutton et al., 2016; Newman et al., 2019; Trevisan et al., 2018).

Scholars have warned about two emerging risks in this new paradigm. First, the filter bubble hypothesis postulates that algorithmic personalization (i.e., pre-selected personalization) filters out information diversity and increase the risk of self-reinforcement (Pariser, 2011). Given this information blindness (Zuiderveen Borgesisus et al., 2016) and the algorithm’s blurredness (i.e., black box) (Schroeder, 2018), Epstein and Robertson (2015) and Epstein et al. (2017) assert that algorithmic personalization is a potential threat for democracy. In line with this, Muddiman (2013) concludes that search engines provide access to mainstream rather than diverse information because they also follow a market model during political campaign. What is more, Hong and Kim (2018) findings confirm the information cascade hypothesis which states that search engines’ users mostly read information read by others as well. At the opposite, a recent growing body of evidence in communication science suggests that the filter bubble fear is fairly exaggerated (Flaxman et al., 2016; Fletcher & Nielsen, 2017; Haim et al., 2017; Haas & Unkel, 2017; Steiner et al., 2020; Unkel & Haas, 2019) concluding that the bubble might have burst (Haim et al., 2018).

Second, the selective exposure hypothesis assumes that citizens select only like-minded source of information (e.g., Stroud, 2011). One might consider that search engines ease a self-selected personalization of information (Zuiderveen Borgesisus et al., 2016). Indeed, citizens have two opportunities to exploit the pull dimension of the Internet when using a search engine. First, they can type what they want to obtain in a search bar. Second, they can freely select a source of information from an ordered list. However, scholars have concluded that trust in search engine creates a digital bandwagon effect, which fosters the selection of information sources ranked in the top of a search engine result page (SERP), rather than a selection of like-minded information (Agichtein et al., 2006; Ghose et al., 2019; Kammerer & Gerjets, 2014; Lorigo et al., 2008; Magno et al., 2006; Pan et al., 2007; Trevisan et al., 2018; Unkel & Haas, 2017).

With that in mind, this study tries to shed light on how citizens select political information in a search engine information environment in a referendum campaign context. It brings information science and social science together. On one hand, from an information and communication science point of view, previous research has highlighted the importance of ranking with individuals selecting more often information ranked on top of the SERP. On the other hand, from a political science point of view, the selection of an information source should be based on one’s own information utility as a democratic citizen rather than at random. This could be a heuristic choice in terms of shortcuts, or an argument-based choice in terms of content.

This brings us to the question of the potentially distorting role of search engines’ algorithms in a democracy. To get politically informed, do citizens simply click on the topmost entries regardless of the expected content utility? Or do they consider textual cues from snippets to select information, regardless of the position on the Google SERP? This study explores these questions and provides new insight into a hot discussion, namely the digitalization of democracy, by extending the search engine-oriented research agenda to a referendum campaign context.

To analyze information selection of search engine users, our paper exploits an online survey experiment, which mimicked a Google webpage. This experiment was conducted during a real-world campaign for a referendum vote on combining tax and pension reform in Switzerland in May 2019. The findings based on a binary logistic regression indicate that citizens tend to mostly select political information sources based on ranking. Heuristic or argument-based selection of political information - disregarding the ranking – remains infrequent and depends on the type of information source. This has important implications for the functioning of a democracy.

# Search Engines and Political Information Selection

Building on Boudreau and MacKenzie (2014) and Lutz (2006), political information can be defined as all the information available to citizens about political actors, institutions, and policies. To be exhaustive, one could also incorporate the perspective that political information can potentially alter citizens’ perceptions. In that case, political information becomes pivotal for citizens’ qualitative, i.e., enlightened, participation in democracy (Strömbäck, 2005). This follows the study by Vowles (2013) asserting that citizens’ ideal participation in democracy relies on citizens having full information. This is especially true in a referendum campaign setting in which citizens are directly involved in policy making. As de Vreese (2007) stated, when facing a referendum, most citizens lack reliable knowledge for making a qualitative choice. Therefore, political information should reduce citizens’ uncertainty and reinforce knowledge during a referendum campaign.

Many recent studies have demonstrated that more than 90% of people use a search engine as a compass for navigating the Internet, including the political information it offers (Lee et al. 2016; Scharkow and Vogelgesang, 2011; Stroud and Muddiman, 2013). Stephens et al. (2014) demonstrated that citizens are motivated to use search engine to obtain political news and information. In Switzerland, Milic et al. (2018) have empirically demonstrated that the political information supply is distributed and accessed via the Internet more frequently. Indeed, the Reuters Digital News Report (Newman et al., 2020) indicated that 77% of Swiss citizens use online media as a source of news.

Bozdag (2013) and Courtois, Sleuchten, and Coenen (2018) have defined a search engine as an information intermediary that facilitates the information-seeking process. In light of the information overload, a search engine filters, prioritizes, and personalizes to set an ordered list of information sources. It is worth noting that a search engine does not generate content itself (Schroeder, 2018); rather, it simplifies users’ access to a wide range of information only after they type in queries to obtain customized, abridged lists of information that could fulfill their search expectations (Flaxman et al., 2016).

The online information is not only high choice (Valentino et al., 2009), but also characterized by the heterogeneity of information sources (Kammerer & Gerjets, 2012). Pirolli (2007) explained that web users anticipate the value of an information source online from heuristic cues (i.e., information scent) and try to match their search expectations with the available “information scent”. Based on the so-called Information Foraging Theory, it is expected that Web users exploit either the ranking or the textual content of the snippet (e.g., URL, summary, headline) as cues to locate their desired information source.

# Hypotheses

Literature in information and communication science provides us with considerable evidence indicating that individuals use ranking as a heuristic cue, rather than the textual content of the snippet, to select information sources. To be precise, they tend to more frequently select search results that appear at the top of the page (Agichtein et al., 2006; Ghose et al., 2019; Kammerer & Gerjets, 2014; Lorigo et al., 2008; Magno et al., 2006; Pan et al., 2007; Trevisan et al., 2018; Unkel & Haas, 2017). First, individuals blindly believe that search engines will rank their most personally relevant result at the top of the results list. Pan et al. (2007) have described this as contemporary trust in search engines. Furthermore, under the satisficing principle, individuals choose satisfactory rather than optimal solutions (Krosnick and Alwin, 1987); thus, they expect search engines to rank the best satisfactory solution at the top of the results list. Second, psychological science’s investigation of the importance of serial position in a rank-ordered list (Haugtvedt and Wegener, 1994) has identified a primacy effect: placing an item at the top of a list reinforces its probability of being selected. Third, due to limited cognitive capacity, humans only consider one choice at a time when dealing with a list; for this reason, items at the top and bottom have an advantage in terms of recall (Mantonakis et al., 2009). Fourth, Höchstötter and Lewandowski (2009) concluded that individuals seldom scroll down the search engine’s results page. This suggests that search results below the fold are rarely selected.

Thus, in line with literature in information and communication science, the first hypothesis states that *when searching for political information online, citizens tend to select more often a search result ranked on top of a SERP* (H1).

Literature in political science provides a different perspective. It indicates that information processing modes are driven by two different paths: A systematic information processing relies on a comprehensive analysis of the content. In comparison, a heuristic information processing rests on peripheral cues to reach a shortcut decision. That is, a SERP provides a short preview of the information, i.e., snippets, with headline, summary and URL. This visual suggestion – with only approximately 200 characters and a link to the full content – supplies various textual cues, rather than argument-based content.

Some recent analyses stated that textual cues are of prime importance in an online information environment (Kessler and Ingelmann, 2019; Sundar et al., 2015). First, Messing and Westwood (2014), Sülflow et al. (2019), and Winter and Krämer (2014) demonstrated that source is a prevalent driver of information selection online. Indeed, these scholars demonstrated that source credibility. i.e., expected quality of the information content, can influence information selection when considering a political information environment. Unkel and Haas (2017) concluded that the credibility of a source, i.e., its reputation, positively influenced the information selection in a SERP. Second, it is also necessary to integrate literature on motivated reasoning and selective exposure (see for a review Yeo et al., 2015) to analyze information selection. According to this theory, individuals have goal-oriented information-seeking strategies. In a search engine information environment, such strategy relies upon the prevalence of textual cues.

Altogether, citizens’ ideal participation in democracy hangs on full information (Vowles, 2013) – if possible – or at least on a selection based on one’s optimal information utility. In a SERP, such selection strategy must be driven by textual cues rather than by an unknown ranking algorithm. For example, these textual cues are could be provided by trustworthy and knowledgeable political actors. In the Swiss direct democracy, the most such relevant actor remains the government (Kriesi, 2005). As a pivotal source of information, the role of the government is to provide factual and impartial political information (Hessami, 2016). Consequently, citizens exploit textual cues from the snippet to locate Web pages from the government. In other words, *when searching for political information online, citizens tend to select governmental Web pages, no matter ranking* (H2a).

What is more, following a recommendation of one’s preferred political party relates to a partisan heuristic. Colombo and Kriesi (2017) and Dermont and Stadelmann-Steffen (2019) indicated that party attachment influences the selection of political information in a referendum. Consequently, we hypothesize that *when searching for political information online, citizens tend to select a political party Web page*, *no matter ranking* (H2b).

Recent literature in political science demonstrated that citizens also rely on policy arguments to form their opinion (Boudreau and MacKenzie 2014; Bullock 2011; Colombo and Kriesi 2017). Overall, individuals using argument-based strategies to form their opinion are more likely to use textual cues to reach information sources such as quality media, which provide topic- or event-related information. To be precise, content and quality of the arguments in terms of coverage vary between different types of media. For example, in contrast to quality newspaper that produce long articles, interviews, and editorial work, free newspaper does not offer a detailed coverage of the referendum campaign (Gerth et al., 2012). At the end, the third hypothesis assumes that *when searching for political information online, citizens tend to select Web pages from quality media, no matter ranking* (H2c).

# Methods and Data

## Overview and Context

The hypotheses were tested using an online survey experiment that mimicked a Google search engine environment by creating a similar layout (see Figure 1). The respondents (*N* = 821) were divided into four treatment groups and one control group. This design allowed for the distinction between the importance of ranking versus textual cues from snippets in a SERP. A binary logistic regression was subsequently used.

The referendum concerned a law to change corporate tax and to enhance the financing of the public retirement provisions (“Steuerreform und AHV-Finanzierung,” STAF). Both topics—corporate tax and retirement provisions—are highly disputed and have been voted on only 2 years earlier. Therefore, a lively campaign and strong predispositions were at work for most people (Heidelberger, 2019; Milic et al., 2018).

## Participants

The sample with 821 respondents is demographically representative, with party affiliation generally matching the political forces in Switzerland in May 2019 (see Table 1 for further details)[[1]](#footnote-1). The respondents were divided into five groups as follows: 154 respondents in Control Group 1, 174 in Treatment Group 2, 162 in Treatment Group 3, 163 in Treatment Group 4, and 168 in Treatment Group 5. Table 1 displays the descriptive statistics, as well as structural consistency tests, used to confirm that the experimental groups were homogenous (*p*-value > .05).[[2]](#footnote-2)

**Table 1**. Descriptive Statistics and Structural Consistency Tests

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | | **Operationalization** | | **Stats** | | **1** | | **2** | | **3** | **4** | **5** | **p-value** |
| Sex  (in %) | (0) male;  (1) female | | 50.86;  49.14 | | 51.30;  48.70 | | 48.28;  51.72 | | 48.15;  51.85 | | 55.90;  44.10 | 50.90;  49.10 | 0.629 |
| Age  (in %) | (1) 18-34 years old;  (2) 35-54 years old  (3) more than 55 years old | | 26.80;  37.39;  35.81 | | 29.22;  35.71;  35.06 | | 22.41;  37.93;  39.66 | | 33.33;  34.57;  32.10 | | 25.77;  38.65;  35.58 | 23.81;  39.88;  36.31 | 0.262 |
| Revenue  (mean) | 8-point scale from (0) below CHF 3’000 to (8) above CHF 15’000 gross monthly household income | | 3.22  (1.71) | | 3.09  (1.65) | | 3.11  (1.59) | | 3.33  (1.69) | | 3.51  (1.86) | 3.10  (1.74) | 0.165 |
| Education  (mean) | 12-point scale from (1) no education to (12) university level | | 7.37  (2.93) | | 7.40  (2.97) | | 7.24  (2.92) | | 7.17  (2.89) | | 7.61  (2.85) | 7.43  (3.06) | 0.668 |
| Political interest  (mean) | 4-point scale from (1) not at all interested to (4) very interested | | 2.91  (0.79) | | 2.92  (0.79) | | 2.87  (0.81) | | 2.83  (0.79) | | 3.01  (0.74) | 2.92  (0.80) | 0.282 |
| Political knowledge  (mean) | Additive index from (0) low political knowledge to (4) high political knowledge | | 2.44  (1.20) | | 2.35  (1.21) | | 2.50  (1.13) | | 2.30  (1.21) | | 2.64  (1.21) | 2.38  (1.21) | 0.064 |
| Trust in government  (mean) | 10-point scale from (1) not trust at all to (10) fully trust the government | | 6.44  (1.81) | | 6.41  (1.93) | | 6.21  (1.94) | | 6.52  (1.64) | | 6.49  (1.80) | 6.57  (1.74) | 0.387 |
| Party attachment (in %) | (1) not close to a party  (2) pretty close to a party  (3) very close to a party | | 55.35;  37.64;  7.01 | | 55.33;  40.00;  4.67 | | 58.38;  34.10;  7.51 | | 58.12;  35.62;  6.25 | | 49.69;  40.49;  9.82 | 55.09;  38.32;  6.59 | 0.408 |
| Vote choice  (mean) | 4-point scale from  (1) absolutely no to  (4) absolutely yes | | 2.70  (0.80) | | 2.69  (0.75) | | 2.74  (0.75) | | 2.5  (0.89) | | 2.66  (0.84) | 2.87  (0.74) | 0.026 |
| Internet as a source  (mean) | (1) I never use the internet (5) I daily use the internet | | 2.94  (1.24) | | 2.93  (1.25) | | 2.84  (1.18) | | 2.93  (1.25) | | 3.01  (1.20) | 2.98  (1.31) | 0.732 |
| Google as a source | (0) Google is not a source;  (1) Google is a source | | 80.88;  19.12 | | 77.92;  22.08 | | 85.06;  14.94 | | 82.10;  17.90 | | 82.21;  17.79 | 76.79;  23.21 | 0.288 |
| Operating system | (0) Computer;  (1) Smartphone (in %) | | 60.17;  39.83 | | 57.14;  42.86 | | 62.64;  37.36 | | 65.43;  34.57 | | 54.60;  45.40 | 60.71;  39.29 | 0.284 |

**Note:** To verify structural consistency, we ran a Pearson chi-square test for independence for nominally scaled variables, and a one-way ANOVA test for independence for metrically scaled variable. Standard deviations are in parentheses.

## Procedure

Data were gathered using a bilingual (French and German) survey, distributed by the Qualtrics commercial survey software (Qualtrics, Provo, UT). Respondents were recruited through an online survey invitation link administered by Qualtrics, and they were invited to complete the survey using either a computer or a smartphone. The experiment lasted approximately 11 minutes. It spanned from April 8 to 15, 2019, that is, 6 weeks before the ballot day.

The experiment replicated a Google information-seeking task. First, respondents were instructed to type search queries describing the referendum-related information they were seeking into a mock Google search bar.[[3]](#footnote-3) Second, they were exposed to a predetermined list of 10 Google search results and instructed to pick as many as they felt were needed to adequately inform themselves and formulate an opinion regarding the referendum vote.[[4]](#footnote-4)

**Figure 1.** Mock Google SERP

Une image contenant texte

Description générée automatiquement

To reinforce external validity, the experiment replicated a layout that mimicked a real-world Google page (e.g., similar colors, a mock Google search bar, and a reproduction of Google news story headlines from repeated real-world observations) (see Figure 1). The 10 predetermined results were comprised as follows: two governmental information sources (admin.ch); four media information sources, including the online platform of a national quality newspaper (*Le Temps* for the French-speaking respondents; *NZZ* for the German-speaking respondents), of a free newspaper (*20 Minuten* in both languages), of a regional newspaper (*La Liberté* for the French-speaking respondents; *Der Bund* for the German-speaking respondents), and of the national television (rts.ch for the French-speaking respondents; srf.ch for the German-speaking respondents); information from the most important Swiss economic organization (economiesuisse.ch); a personal blog; information of the respondents preferred party;[[5]](#footnote-5) and easyvote, a neutral platform on which official government is presented in a simplified form for a wide audience (easyvote.ch).

The same news results were used for all groups with only variations in the order in which they were presented. First, Control Group 1 was the reference (i.e. control) group. The 10 search results were randomly assigned. Second, Treatment Group 2 was exposed to 5 selected search results ranked from the 1st to 5th, while the remaining 5 search results were ranked randomly from the 6th to 10th. Third, Treatment Group 3 was organized in reverse mirror of Treatment Group 2. The 5 randomly ranked results in Treatment Group 2 were ranked 1st to 5th, and the search results that were ranked 1st to 5th in Treatment Group 2 were randomly ranked 6th to 10th. Fourth, two governmental sources of information were ranked at the top of the results page, and the 8 remaining search results were randomly ranked 3rd to 10th. Finally, 2 search results were ranked 1st and 2nd with the addition of a sponsored results box (i.e., advertisements). The remaining 8 search results were randomly ranked 3rd to 10th.

## Measures

Following Jang (2014) and Winter and Krämer (2012), a nominally scaled absolute selection rate was used for the binary logistic regression. The absolute selection rate was a binary variable where search results obtained scores of 1 if they were selected and 0 if they were not. On average, respondents picked 2.7 information sources. Then, two variables of interest were incorporated. On one hand, the search result’s position on the Google experiment’s rank-ordered list is a nominally scaled variable, which goes randomly from 0 to 13. Each position in the rank-ordered list gets a nominal value, and there are three other possibilities as follows: being ranked in the top 5 or the last 8, or being a sponsored result (i.e., Google ads). On the other hand, the information source is also nominally scaled. It gathers the different types of political information sources in the mock SERP. At last, control variables such as sociodemographic characteristics, politically related attributes, and information-seeking habits of the individuals were also incorporated. Though individual-related attributes might alter information selection behavior, we refrained from analyzing the role of political-related and demographics attributes as moderator variables. Not only this goes beyond the scope of this article, but also in his seminal work on Google log searches in Australia, Waller (2011) offers persuasive evidence that there are no differences in information-seeking behavior across contrasting population demographics. Therefore, these variables are included as control rather than moderator variables. [[6]](#footnote-6)

## Binary logistic regression

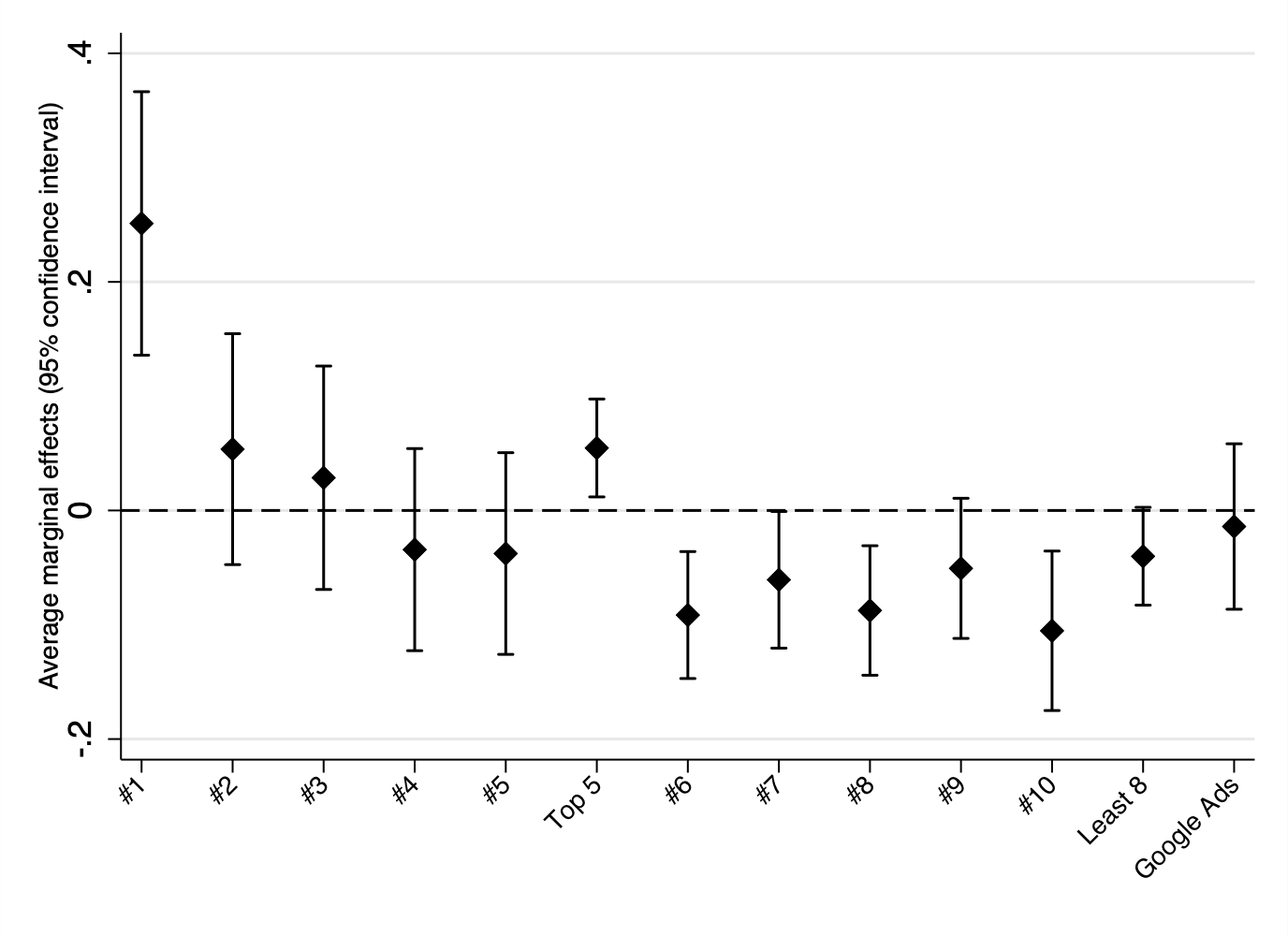
To run a binary logistic regression, the database was reshaped in long format, implying that each respondent is associated with 10 decisions of choosing an information source in the Google rank-ordered list or not. The dependent variable is the absolute selection rate, i.e., a binary variable where search results obtained scores of 1 if they were selected and 0 if they were not.

With respect to the dependency of the data, the method considered that each observation is nested at a different level. Given the treatment group design, the database was formed based on 8,210 observations (Level 1 units) nested in the following two clusters: respondents (Level 2 units) and treatment groups (Level 3 units). The clusters’ homogeneity was measured with the model’s intra-class correlation coefficient (ICC). Results is 0.001 for Level 2 and 0.003 for Level 3. An ICC that is close to 0 for both clusters implies that variation exists mainly within clusters instead of between them. In other words, online information selection behavior is homogenous across respondents and treatment groups. Given that the ICC coefficients are close to 0, it is possible to run a binary logistic regression instead of a multilevel logistic regression.[[7]](#footnote-7)

# Results

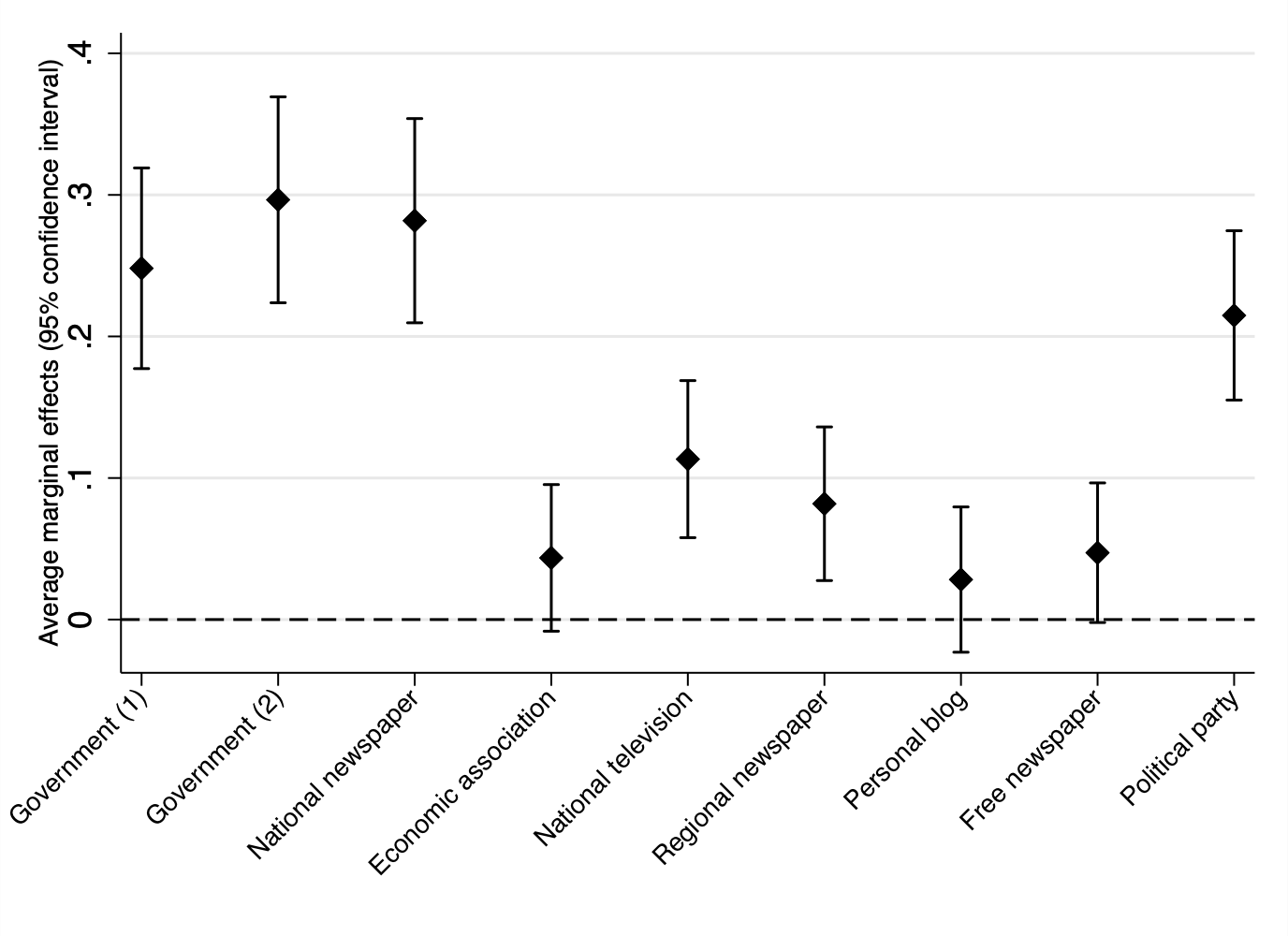
The binary logistic regression emphasizes the online information seeking behavior of citizens facing a SERP. To begin with, figure 2a shows the importance of ranking. Search results ranked at the top of the list tend to be selected more often than those at the bottom. This sharp difference is statistically significant with the extremity of the rank-ordered list: on one side, the probability that an individual will select a search result that is ranked 1st is 4 times higher than with random ranking; on the other side, an individual has a lower probability (3 times lower) of choosing a search result that is ranked 10th in comparison with a random ranking. Then, search results that are ranked higher than 5th (Top 5) are selected more often than those below. We assume that the positive influence of a Top 5 ranking is mostly driven by a 1st position. What is more, the influence of ranking is stronger for Google’s results ranked below the 5th position. Indeed, for a search result, a 6th to 10th ranking significantly reduces the probability of being selected. Subsequently, it is also worth noting that our model identified no differences in selection between a sponsored result (i.e., Google ads) and a random ranking. In summary, H1 is accepted.

**Figure 2a.** Political information selection on a SERP – Influence of ranking



The importance of political information sources in selection is illustrated in Figure 2b. Users exploited the textual cues from snippets to locate their desired source of information. To begin with, the results pinpoint a sharp increase in the selection of governmental sources of information. Furthermore, concerning the media, respondents tended to rely heavily on national newspaper, regional newspaper, and national television (i.e., quality media). This is especially true for the national newspaper. A third source of information is also highly significant: one’s preferred political party. In other words, respondents based their selection on cues which arose from either the government, a preferred political party, or a quality media. To the contrary, respondents neglected the economic organization, the free newspaper (i.e., tabloid) and the personal blog as sources of information.

**Figure 2b.** Political information selection on a SERP – Influence of textual cues

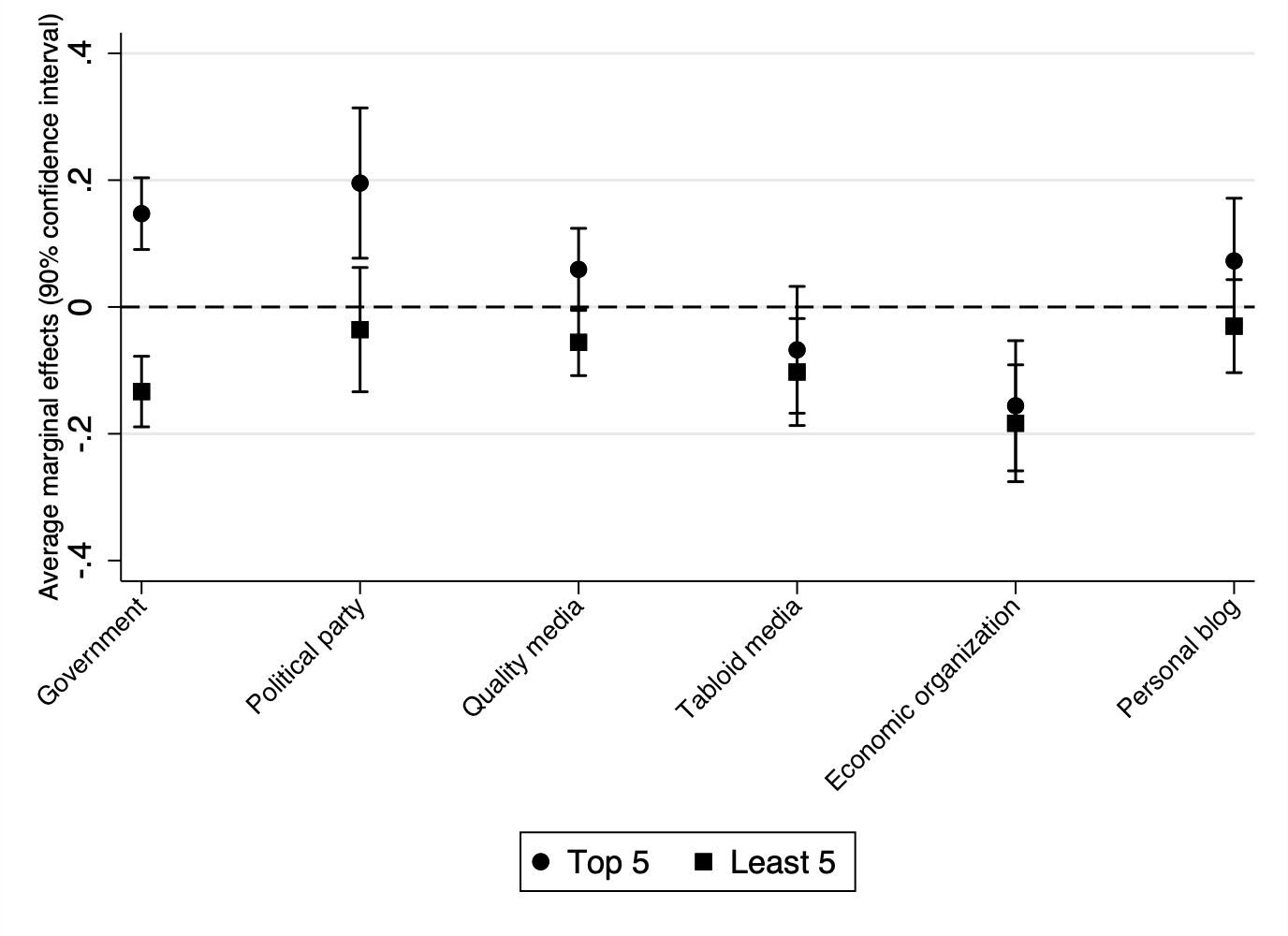


**Note:** The easyvote.ch neutral platform from the government is the base category.

To examine hypotheses 2, it is necessary to measure the interaction effect of ranking with textual cues. On one side, a new categorical variable for ranking with three categories was created: at random, top 5 and least 5. This new subdivision relied upon our preliminary results. On the other side, political sources of information were grouped in six categories. This builds on our preliminary results and on Tate’s (2010) classification of information sources in a SERP. First, the governmental Web pages are grouped into the first category. It includes the two governmental Web pages and the easyvote.ch neutral platform. This matches Tate’s (2010) “informational” category with Web pages supplying factual, i.e., neutral, information. Second, Tate (2010) gather information sources which aim at influencing public opinion into one category termed “advocacy”. Two political information sources from the mock SERP match this classification: political party and economic organization. However, considering the relevance of political party cues in a referendum context (Colombo and Kriesi, 2017; Dermont and Stadelmann-Steffen, 2019) and the hypothesis H2b, it is more precise to separate these two political sources of information, rather than grouping them into an “advocacy” category. Third, Tate (2010) explains that users also encounter “news” Web pages, which present topic- or event-related information. To obtain a finer-grained measure, the analysis divides the “news” category into quality “news” Web pages, i.e., national newspaper, regional newspaper, and national television, and tabloid “news” Web page, i.e., free newspaper. Finally, the personal blog Web page fits into the “personal” Web page category.

Figure 3 displays the interaction effect of textual cues from snippets with the ranking in the ordered SERP. It measures the tension between ideal democracy, i.e., selection based on citizen’s utility, and the influence of the ranking’s algorithm. To begin with, the logistic regression with interaction effects confirms the findings in Figure 2b. That is, citizens are using textual cues to locate only three political information sources: governmental, political party and quality media. To the contrary, the personal blog, the economic organization, and the tabloid media are mostly ignored by citizens. This is in line with literature in political science (e.g. Milic et al. 2014). Then, the results emphasize the strong influence of ranking on the selection of governmental and political party Web pages. In other words, a statistical difference in selection rate between a top 5, least 5 and random ranking exists for governmental and political party Web pages. For example, the probability to select a governmental Web page is 65% higher with a top 5 ranking, and 32% lower with a least 5 ranking in comparison with a random ranking. Similarly, the probability to select a political party Web page is 70% higher with a top 5 ranking. Consequently, H2 a and H2b are rejected. The selection of governmental and political party Web pages is dependent of ranking in a SERP. Finally, the selection of quality media seems to be independent from ranking. The binary logistic regression identifies no statistical difference in selection rate between a top 5, a least 5 and a random ranking. H2c is accepted.

**Figure 3.** Political information selection on a Google results page - Interaction effect between ranking and political information source



**Note:** Basis category is random ranking.

At last, it is also worth mentioning that a slightly higher absolute selection rate was observed for respondents who have higher education, as well as for those who use the Internet as a source of political information. To deepen the analysis, distinction was made between computer and smartphone users, as well as Google versus non-Google users. Nevertheless, the results with these two differentiations remained identical to the results already mentioned.

# Discussion

The ever-growing importance of the Internet has affected the literature on political information selection; though a broad literature in information and communication science rejected fear related to the filter bubble hypothesis and emphasized the seminal importance of ranking in online information selection on a SERP, only few scholars examined the tension between ideal information seeking in a participatory democracy, i.e., information selection based on citizen’s utility, and the influence of ranking on political information selection. This analysis has shown that mostly ranking, and only to a lesser degree the textual cues from snippets are important to the selection of political information in a SERP.

The results indeed demonstrate that ranking influences the information selection strategy. The findings are in line with literature in information and communication science (Agichtein et al., 2006; Ghose et al., 2019; Kammerer & Gerjets, 2014; Lorigo et al., 2008; Magno et al., 2006; Pan et al., 2007; Trevisan et al., 2018; Unkel & Haas, 2017). The results also emphasize the tremendous importance of ranking extremity. That is, a 1st position strongly increases the selection rate, and a 10th position strongly reduces the selection rate. In addition, the results indicate a cut-off at the 5th position. This is in line with Höchstötter and Lewandowski (2009) who identified that search results below the fold are rarely selected. In other words, (political) information selection stops where the screen ends. One might conclude that this prevalence of ranking is driven by a contemporary trust in search engine (Pan et al., 2007) and the satisficing principle (Krosnick and Alwin, 1987). That is, citizens aim at a satisfactory solution rather than an optimal solution. Thus, they blindly trust that search engines provide them with their most personally relevant result at the top of the list.

Examining the combined influence of political information sources with ranking, I accept H2c (quality media). The selection of quality media Web pages is independent from ranking. Thus, citizens use textual cues from snippets to locate such quality “news” information sources. This finding is in line with literature in political science (Boudreau and MacKenzie 2014; Bullock 2011; Colombo and Kriesi 2017; Dermont and Stadelmann-Steffen, 2019). Citizens tend to opt for quality media (e.g., national newspaper, regional newspapers, and television) to nurture their argument-based strategy. Such media provide longer and more informative articles, interviews, and editorial work based on quality journalism. At the end, the selection of quality “news” Web pages from online media no matter their ranking in a SERP is good news for democracy.

However, at the opposite, H2a (government) and H2b (political party) are rejected. Citizens ignore “informational” and political party Web pages if they are not ranked in a top position. As for the parties, these findings disagree with Colombo and Kriesi (2017) who concluded that political party cues play a seminal role in a referendum context. This study concludes that citizens locate political party cues only if they are in a top position in the SERP. That is, trust in search engine and the satisficing principle trump the role of political party cue in a SERP. Furthermore, citizens neglect “informational” Web pages, i.e., governmental Web pages in our study, if they are not in a top position. This conclusion not only conflicts with Kriesi (2005) who argue that the government is the most trustworthy political actor, but also raises the question of the role of search engines’ algorithms in the pre-selection and dissemination of political information.

What are the implications from a democratic theory perspective? On one hand, one might argue that the role of the algorithm is not to choose, but rather to reflect mass leanings towards a few sources (Granka, 2013). That is, a “well-designed” algorithm provides citizens with the political information sources they “want” to obtain based on their previous searches in addition to other people’s searches on the same topic. Then, many scholars in information and communication science recently demonstrated that search engines provide a high diversity of political information concluding that the algorithmic filter bubble burst (Haim et al., 2018; Steiner et al., 2020; Unkel & Haim, 2019). Furthermore, this study concludes that citizens use textual cues to locate “news” Web pages no matter ranking. These quality “news” Web pages are seminal to form an opinion with an argument-based strategy – an important element of Dahl’s (1989) ideal democracy where individuals chose information to reach informed and enlightened decisions that serve their best personal interest.

On the other hand, one might postulate that online political information seeking behavior and algorithmic blurredness is a potential threat for democracy given the lack of transparency of the algorithm pre-selection (Epstein et al. 2017). Though “informational” Web pages provide factual and neutral political information and political party Web pages provide shortcut decision, this empirical study demonstrated that citizens consult them, only if they are ranked in a top position, i.e., top 5 ranking. That is, their selection is dependent of the algorithmic personalization. Alternatively, given that only 3% of searches are potentially related to political information (Waller, 2011), the question is whether the algorithm is sufficiently fed information-wise to deliver diverse political information sources. That is, Muddiman (2013) and Hong and Kim (2018) concluded that search engine provide a highly concentrated distribution of information.

It is worth focusing on the ranking and selection of governmental Web page to further this debate. Though it is assumed that more than 200 parameters compute the ranking of a SERP (Unkel & Haim, 2019), it is highly likely that a government Web page appears in a top position in a SERP during a referendum campaign. Not only has the referendum already been debated in Parliament, but logically, a government Web page features many more hyperlinks than a referendum committee Web page create solely for the campaign. Thus, the “well-designed” algorithm automatically ranked this “informational” Web page in a top position increasing citizens’ selection of factual information. Of course, one could argue that, given the importance and the trustfulness of such information (Kriesi 2005), such an algorithm pre-selection of a source is of no harm to democracy. However, taking Dahl’s (1989) idea of the “ideal democracy” seriously, no information should be ranked over another. Citizens must have the opportunity to find sources that help them to find informed decisions that correspond to their individual personal interest. Thus, because we do not know exactly how the algorithm works, the fact that citizens tend to choose top-ranked sources is a potential threat to Dahl's (1989) ideal democracy. The threat would be even greater if it were possible to place one-sided or less informative sources on the topmost entries.

At the end, this experimental study provides interesting results about the role of search engine for searching political information. It extends research on search engine to direct democratic votes. Based on Dahl’s (1989) ideal democracy, and in line with Steiner et al. (2020), and Unkel and Haim (2019), the findings call for higher algorithmic transparency to ensure. To ensure that citizens benefit from the Internet high-choice and interactive environment, a “democratically well-designed” algorithm must be provided. However, the lack of transparency impedes a thorough analysis by neutral investigator.

# Limitations and Suggestions for Future Research

Though this study deepens the understanding of political information selection in a SERP, it would be necessary to also study what individuals type in the Google search bar. Not only the algorithmic pre-selection is arguably driven by citizens’ search queries, but also search queries inform about the self-selection intentions of citizens. Are web users opting for generic political search terms? Or are they already indicating their self-selection intentions to the algorithm? This is a research agenda that would benefit from additional scholarly attention.

# References

Agichtein, E., Brill, E., & Dumais, S. (2006). Improving web search ranking by incorporating user behavior information. In: *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp.19-26).Seattle, Washington, USA. August 6-10.

Boudreau, C., & Mackenzie, S. A. (2014). Informing the electorate? How party cues and policy information affect public opinion about initiatives. *American Journal of Political Science* 58(1), 48-62.

Bullock, J. (2011). Elite influence on public opinion in an informed electorate. *American Journal of Political Science* 105(3), 496-515.

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, 15(3), 209-227.

Colombo, C., & Kriesi, H. (2017). Party, policy – or both? Partisan-biased processing of policy arguments in direct democracy. *Journal of Elections, Public Opinion and Parties* 27(3), 235-253.

Courtois, C., Sleuchten, L., & Coenen, L. (2018). Challenging Google Search filter bubbles in social and political information: Disconfirming evidence from a digital methods case study. *Telematics and Informatics* 35, 2006-2015.

Dahl, R. A. (1989). *Democracy and Its Critics*. Michigan: Yale University Press.

De Vreese CH (2007) When referendum campaigns matter – and why? Paper presented at the 57th Conference of the International Communication Association ICA, San Francisco, May.

Dermont, C., & Stadelmann-Steffen, I. (2019). The role of policy and party information in direct-democratic campaigns. *International Journal of Public Opinion Research* 32(3), 442-466.

Dutton, W. H., Reisdorf, B. C., Dubois, E., & Blank, G. (2017). *Search and politics: The uses and impacts of search in Britan, France, Germany, Italy, Poland, Spain, and the United States (Quello Center Working Paper No. 5-1-17).* Michigan Sate University.

Eagly, A., & Chaiken, S. (1993). *The Psychology of Attitudes*. Fort Worth, TX: Harcourt.

Epstein, R., & Robertson, R. E. (2015). The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. In: *Proceedings of the* *National Academy of Sciences of the United States of America* 112(33), 4512-4521.

Epstein, R., Robertson, R. E., Lazer, D., & Wilson, C. (2017). Suppressing the Search Engine Manipulation Effect (SEME). In: *Proceedings of the ACME on Human-Computer Interaction* 1(42). Issue 2.

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubble, echo chambers, and online news consumption. *Public Opinion Quarterly*, 80, 298-320.

Fletcher, R., & Nielsen, R. K. (2017) Are news audience increasingly fragmented? A cross-national comparative analysis of cross-platform news audience fragmentation and duplication. *Journal of Communciation* 67(4), 476-498.

Gerth, M. A., Dahinden, U., & Siegert, G. (2012). Coverage of the Campaigns in the Media. In: Kriesi, H. (ed.). *Political Communication in Direct Democratic Campaigns. Enlightening or Manipulating?* (pp.108-124). Palgrave Macmillan.

Ghose, A., Ipeirotis, P. G., & Li, B. (2019). Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science* 60(7), 1632-1654.

Granka, L. (2013). Using online searches to predict US presidential elections. *PS: Political Science and Politics* 46(2), 271-279.

Haim, M., Arendt, F., & Scherrer, S. (2017). Abyss or shelter? On Google’s role when googling for suicide. *Health communication* 32(2), 253-258.

Haim, M., Graefe, A., & Brosius, H. B. (2018). Burst of the filter bubble?: Effects of personalization on the diversity of Google News. *Digital Journalism.*

Haas, A., & Unkel, J. (2017). Ranking versus reputation: Perception and effects of search result credibility. *Behaviour & Information Technology* 36(12), 1285-1298.

Hargittai, E., Neuman, W. R., & Curry, O. (2012). Taming the Information Tide: Perceptions of Information Overload in the American Home. *The Information Society* 28(3),161–73.

Haugtvedt, C. P., & Wegener, D. T. (1994). Message order effects in persuasion: An attitude strength perspective. *Journal of Consumer Research* 21(1), 205-218.

Heidelberger, A. (2019). APS-Inserateanalyse der eidgenössischen Abstimmung vom 19. Mai 2019. Bern: Année Politique Suisse, Institut für Politikwissenschaft der Universität Bern.

Hessami, Z. (2016). How do voters react to complex choices in a direct democracy? Evidence from Switzerland. *Kyklos* 69(2):, 263-293.

Hilbert, M., & López, P. (2011). The World’s Technological Capacity to Store, Communicate, and Compute Information. *Science* 332(6025), 60–65.

Hindman, M. (2009). *The Myth of Digital Democracy*. Oxford: Princeton University Press.

Höchstötter, N., & Lewandowski, D. (2009). What users see – structures in search engine results page. *Information Sciences* 179(12), 1796-1812.

Hong, S, & Kim, N. (2018). Will the internet promote democracy? search engines, concentration of online news readership, and e-democracy. *Journal of Information Technology & Politics* 15(4), 388-399.

Jang, M. J. (2014). Challenges to Selective Exposure: Selective Seeking and Avoidance in a Multitasking Media Environment. *Mass Communication & Society* 17(5), 665-688.

Kammerer, Y., & Gerjets, P. (2012). How search engine users evaluate and select web search results: The impact of the search engine interface on credibility assessments. In D. Lewandowski (Ed.), *Web search engine research* 4, 251-279. Emerald Group Publishing Limited.

Kammerer, Y., & Gerjets, P. (2014). The Role of Search Result Position and Source Trustworthiness in the Selection of Web Search Results When Using a List or a Grid Interface. *International Journal of Human-Computer Interaction* 30(3), 171-191.

Kessler, S. H., & Engelmann, I. (2019). Why do we click? Investigating reasons for user selection on a news aggregator website. *European Journal of Communication Research* 44(2), 225-247.

Knobloch-Westerwick, S., Mothes, C., Johnson, B. K,. Westerwick, A., & Donsbach, W. (2015). Political online information searching in Germany and the U.S.: Confirmation bias, source credibility, and attitude impacts. *Journal of Communication* 65, 489-511.

Kriesi, H. (2005). *Direct Democratic Choice. The Swiss Experience*. Lanham: Lexington Books.

Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey management. *The Public Opinion Quarterly* 51(2), 201-219.

Lee, B., Kim, J., & Scheufele, D. A. (2016). Agenda setting in the Internet Age: The reciprocity between online searches and issue salience. *International Journal of Public Opinion research* 28(3), 440-455.

Lorigo, L., Haridasan, M., Brynjarsdottir, H., Xia, L., Joachims, T., Gay, G., Granka, L., Pellacini, F, & Pan, B. (2008) Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology* 59(7), 1041-1052.

Lutz, G. (2006). Participation, Information and Democracy. The Consequences of Low Levels of Participation and In- formation for the Functioning of Democracy. Hamburg: Lit Verlag.

Magno, G., Araujo, C. S., Wagner, M., & Almeida, V. (2016). Stereotypes in search engine results: Understanding the role of local and global factors. In *Presented in the Workshop on Data and Algorithmic Transparency (DAT’16)*. November.

Mantonakis, A., Rodero, P., Lesschaeve, I., & Reid, H. (2009). Order in choice: Effects of serial position on preferences. *Association for Psychological Science* 20(11), 1309-1312.

Messing, S., & Westwood, S. J. (2014). Selective exposure in the age of social media: endorsements trump partisan source affiliation when selecting news online. *Communication Research* 41(8), 1042-1063.

Milic, T., Rousselot, B., & Vatter, A. (2014). Handbuch der Abstimmungsforschung. Zürich: NZZ.

Milic, T., Reiss, T., & Kübler, D. (2018). *VOTO-Studie zur eidgenössischen Volksabstimmung vom 10. Juni 2018*. ZDA, FORS, LINK: Aarau/Lausanne/Luzern.

Mitchell, A., & Holcomb, J. (2016). State of the News Media 2016. Pew Research Center. June.

Muddiman, A. (2013). Searching for the Next U.S. President: Differences in Search Engine Results for the 2008 U.S. Presidential Candidates. *Journal of Information Technology & Politics* 10(2), 138-157.

Neuman, W. R., Park, A. J., & Panek, E. (2012). Info Capacity: Tracking the Flow of Information into the Home: An Empirical Assessment of the Digital Revolution in the US from 1960–2005. *International Journal of Communication* 6(20), 1022–41.

Newman, N., Fletcher, R., Klageropoulos, A., & Kleis Nielsen, R. (2019). *Reuters Institute Digital News Report 2019*. Oxford, England: Reuters Institute for the Study of Journalism.

Newman, N., Fletcher, R., Schulz, A., Andi, S., & Kleis Nielsen, R. (2020) *Reuters Institute Digital News Report 2020*. Oxford, England: Reuters Institute for the Study of Journalism.

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.

Pariser, E. (2011). *The Filter Bubble: How the new personalized web is changing what we read and how we think.* New York, NY: Penguin.

Petty, R., & Cacioppo, T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology* 19, 123-205.

Pirolli, P. (2007). *Information foraging theory: Adaptative interaction with information*. New York, NY: Oxford University Press.

Qualtrics (2017). Version 2017 of Qualtrics. Copyright © 2017. Qualtrics. Provo, Ut, USA. [Computer software]

Scharkow, M., & Vogelgesang, J. (2011). Measuring the public agenda using search engine queries. *International Journal of Public Opinion Research* 23(1), 104-113.

Schroeder, R. (2018). The internet in everyday life II: Seeking information. In Schroeder, R. (eds) *Social theory after the internet* (pp. 101-125). London: UCL Press.

Sommet, N., & Morselli, D. (2017). Keep calm and learn multilevel logistic modelling: A simplified three-step procedure using Stata, R, MPlus and SPSS. *International Review of Social Psychology* 30(1), 203-218.

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*.

Stephens, M., Yoo, J., Mourao, R., Hong, V., Baresch, B., & Johnson, T. (2014) How App Are People to Use Smartphones, Search Engines, and Social Media for News?: Examining Information Acquisition Tools and Their Influence on Political Knowledge and Voting. *Journal of Information Technology & Politics* 11(4), 383-396.

Strömbäck, J. (2005). In search of a standard: Four models of democracy and their normative implications for journalism. *Journalism Studies* 6, 331-345.

Stroud, N. J. (2011) *Niche news: The politics of news choice.* New York, NY: Oxford University Press.

Stroud, N. J., & Muddiman, A. (2013). Exposure to news and diverse views in the Internet age. *Journal of Law and Policy for the Information Society* 8(3), 605-623.

Sülflow, M., Schäfer, S., & Winter, S. (2019). Selective attention in the news feed: An eye tracking study on the perception and selection of political news on Facebook. *New media & Society* 21(1), 168-190.

Sundar ,S. S., Jia, H., Waddell, T. F., & Huang, Y. (2015). Towards a theory of interactive media effects (TIME). Four models for explaining how interface features affect user psychology. In Sundar SS (ed.), *The Handbook of the psychology of communication technology* (pp.47-86). Malden, MA: Wiley Blackwell.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science* 50(3), 755-769.

Tate, M. A. (2010). *Web wisdom: How to evaluate and create information quality on the Web* (2nd Ed.). Boca Raton, FL: CRC Press.

Trevisan, F., Hoskins, A., Oates, S., & Mahouly, A. (2018). The Google voter: search engines and elections in the new media ecology. *Information, Communication & Society* 21(1), 111-128.

Unkel, J., & Haas, A. (2017). The effects of credibility cues on the selection of search engine results. *Journal of the Association for Information Science and Technology* 68(8), 1850-1862.

Unkel, J., & Haim, M. (2019). Googling politics: Parties, sources, and issue ownership on Google in the 2017 German federal election campaign. *Social Science Computer Review* published online.

Valentino, N. A., Banks, A. J., Hutchings, V. L., & Davis, A. K. (2009). Selective exposure in the Internet age: The interaction between anxiety and information utility. *Political Psychology* 30(4), 591-613.

Vowles, J. (2013). Campaigns claims, partisan cues, and media effects in the 2011 British electoral system referendum. *Electoral Studies* 32, 253-264.

Waller, V. (2011). Not Just Information: Who Searches for What on the Search Engine Google? *Journal of the American Society for Information Science and Technology* 62(4), 761–775.

Winter, S., & Krämer, N. C. (2012). Selecting science information in Web 2.0: How source cues, message sidedness, and need for cognition influence users' exposure to blog posts. *Journal of Computer-Mediated Communication* 18, 80–96.

Winter, S., & Krämer, N. C. (2014). A question of credibility – effects of source cues and recommendations on information selection on news sites and blogs. *The European Journal of Communication Research* 39(4), 435-456.

Xenos, M. A., Scheufele, D., Brossard, D., Choi, D. H., Cacciatore, M., Yeo, S., & Su Y. F. L. (2018). News media use and the informed public in the digital age. *International Journal of Communication* 12, 706-724.

Yeo, S. K., Cacciatore, M. A., & Scheufele, D. (2015). News selectivity and Beyond: Motivated reasoning in a changing Media Environment. In Jandura O, Petersen T, Mothes C and Schielicke AM (Eds.). *Publizistik und gesellschaftliche Verantworttung*. Berlin: Verlag Springer.

Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles?. *Internet Policy Review* 5(1).

1. The sample’s party affiliation is 24.80% SVP, 16.27% SP, 12.99% FDP, 6.04% CVP, 6.04% Greens, 6.96% Green Liberals, 3.67% BDP, 12.11% other remaining parties, and 11.02% with no affiliations. [↑](#footnote-ref-1)
2. Variable voting choice was one exception to this. Groups 3 and 5 differed significantly at the 0.05 level, but not at the 0.01 level. [↑](#footnote-ref-2)
3. The exact wording of the instructions was as follows: “The vote concerning the tax policy and AHV financing reform takes place in a few weeks. The campaign just started. Then, you probably have only limited knowledge on the topic. We give you the opportunity to use a Google search engine to search for information and to form an opinion related to the vote. Type in the search bar what kind of information you want to obtain related to the vote.” [↑](#footnote-ref-3)
4. The exact wording of the instructions was as follows: “Click on the sources you would like to read.” [↑](#footnote-ref-4)
5. The respondents were asked which party they were voting for. The search result “Party” was adjusted specifically for each respondent. [↑](#footnote-ref-5)
6. The Google experiment has proved to be not only robust, but also externally valid. We verified what respondents typed in the mock Google search bar and analyzed Google trends during the real-world referendum. Please contact the author for further information. [↑](#footnote-ref-6)
7. To verify the robustness of our findings, we also ran a multilevel logistic regression and obtained similar results. For further details on multilevel logistic regression, see Sommet and Morselli (2017). [↑](#footnote-ref-7)