Search engine effects on news consumption: Ranking and representativeness outweigh familiarity in news selection

Ulloa, Roberto; Kacperski, Celina Sylwia

Search engine effects on news consumption: Ranking and representativeness outweigh familiarity in news selection

Roberto Ulloa iD
GESIS – Leibniz Institute for the Social Sciences, Germany; University of Konstanz, Germany

Celina Sylwia Kacperski iD
University of Konstanz, Germany; Seeburg Castle University, Austria

Abstract
While individuals' trust in search engine results is well-supported, little is known about their preferences when selecting news. We use web-tracked behavioral data across a 2-month period (280 participants) and we analyze three competing factors, two algorithmic (ranking and representativeness) and one psychological (familiarity), that could influence the selection of search results. We use news engagement as a proxy for familiarity and investigate news articles presented on Google search pages ($n = 1221$). We find a significant effect of algorithmic factors but not of familiarity. We find that ranking plays a lesser role for news compared to non-news, suggesting a more careful decision-making process. We confirm that Google Search drives individuals to unfamiliar sources, and find that it increases the diversity of the political audience of news sources. We tackle the challenge of measuring social science theories in contexts shaped by algorithms, demonstrating their leverage over the behaviors of individuals.

Corresponding author:
Roberto Ulloa, Computational Social Science, GESIS—Leibniz Institute for the Social Sciences, Unter Sachsenhausen 6-8, Cologne, 50667 Köln, Germany.
Email: roberto.ur@protonmail.com

Correction (March 2023): Article updated to correct the affiliation of Roberto Ulloa.

Introduction

Online platforms have expanded the ways in which individuals access and interact with news (Möller et al., 2020). Across a variety of countries, websites have become the most important source for news (75%); about 25% of news users name search engines as their main way of coming across news online, with 25% direct, and another 26% via social media (Newman et al., 2021). There is a need to understand how individuals seek information and how these decisions affect their behavior (Loecherbach et al., 2021; Sharot and Sunstein, 2020). Search engines are of particular interest: with their role as gatekeepers (Nechushtai and Lewis, 2019), they are in the privileged position of fulfilling an individual's need for information by preselecting, sorting, and presenting resources from the millions available on the Internet. Crucially, individuals place a high degree of trust in search engines, which is evidenced by how often they select the top results (Pan et al., 2007; Urman and Makhortykh, 2021).

Despite the strong effect of ranking, it is important to consider other factors that might influence the selection of results. Some studies have found a positive association between search engine use and the diversity of news sources visited by individuals (Fletcher et al., 2021; Fletcher and Nielsen, 2018; Scharkow et al., 2020), while others have analyzed representativeness of sources in search results (Haim et al., 2018; Nechushtai and Lewis, 2019; Puschmann, 2019; Urman et al., 2021). Yet, to the best of our knowledge, no one has yet investigated how representativeness of news sources affects individuals' selection of news articles, even though the probability with which a user selects a domain might be affected by the frequency with which it is displayed on the search result page.

Most interestingly, personal preferences should influence the result selection (e.g. Robertson et al., 2021), but they remain understudied due to the difficulties of collecting behavioral data that offers a comprehensive picture of what individuals see and do online. Specifically, we use web tracking data to test the hypothesis that familiarity with news sources predicts online news selection in search engine results.

In the context of the argument made for theory development in environments shaped by algorithms (Wagner et al., 2021), neither algorithmically decided factors nor psychological phenomena such as familiarity should be explored in isolation. Thus, supported by the above theories and empirical data, we analyze three factors using linear mixed-effect models (controlled for surveyed individual characteristics) to investigate how individuals select news articles from a list of search results: two algorithmic (ranking and source representativeness) and one psychological (familiarity), and study their potential interactions. In a novel manner, we explore web tracking data in the form of a behavioral field study to analyze the relation between these three factors. Access to the full HTML of individuals' visited pages and tab activity of the browser allow us to examine not only their selections, but also the alternative options to which participants were exposed.

Literature review

With the emergence of new forms of accessing information online, concerns have emerged about the impact that online platforms have on society, because they guide individuals to algorithmically chosen content (Rahwan et al., 2019; Wagner et al., 2021).

Researchers have demonstrated that search results influence judgments, decisions, and behaviors ranging from purchases (Ghose et al., 2014) to health information (Kammerer and Gerjets, 2012; Lau and Coiera, 2009) and to voting preferences (Epstein et al., 2017; Epstein and Robertson, 2015; Zweig, 2017).

Crucially, individuals place a high degree of trust in search engines, which is evidenced by how often they select the top results (Pan et al., 2007; Urman and Makhortykh, 2021). The ranking effect is so widely accepted that heavy weights are used in studies that investigate biases in search engines (Kulshrestha et al., 2017; Robertson et al., 2018). The role of ranking finds strong support in psychological theories that have shown that the first (and last) positions of items in lists are recalled more often (Murdock, 1962; Murphy et al., 2006), and that primacy effects impact attitudes and beliefs (Asch, 1946; Sullivan, 2019). Similarly, the strong effect of ranking on choice is in line with findings on default effects, well-supported across a variety of behavioral experimental studies (Hummel and Maedche, 2019; Jachimowicz et al., 2019). However, most research on search engine rankings deals with general search behavior, and to the best of our knowledge, no research has yet identified how rankings impact individuals' exposure to news results specifically, or whether there are differences in ranking effects for news and non-news.

Online actors compete to improve the visibility of their websites on search engines, a practice known as Search Engine Optimization (SEO). On one hand, the details of filtering and ranking algorithms are often kept confidential by search companies; researchers here opt for indirect methods such as search engine audits (Ulloa et al., 2022a) to demonstrate the presence of biases in search results (e.g. Robertson et al., 2018; Ulloa et al., 2022b; Urman et al., 2021). On the other hand, search companies provide guidelines to make it easier for themselves to process online content (Google Developers, 2022). At the same time, news organizations have become increasingly reliant on these digital intermediaries, as they offer short-term opportunities to engage audiences, even if this may result in a loss of control over their professional identity (Nielsen and Ganter, 2018). And although news reports are guided by journalistic norms (Hackett, 1984; Muñoz-Torres, 2012), research indicates that market forces may be influencing the gatekeeping aspect of the media (Hamilton, 2011; Patterson, 2013).

Researchers have raised concerns about the potential effects that online platforms have on society. Three theoretical concepts are commonly named to describe such concerns: filter bubbles, echo chambers, and selective exposure (for a review, see Ross Arguedas et al., 2022). Filter bubbles happen when content is selected according to individuals' previous consumption, creating a feedback loop which hinders the exposure to different views (Pariser, 2011), which is closely related to the effects of search personalization (Hannak et al., 2013). Echo chambers refer to environments in which individuals are exposed to information mostly from like-minded individuals (Bakshy et al., 2015; Dubois and Blank, 2018). Selective exposure describes the tendency of individuals to select information matching their beliefs (Garrett, 2009; Stroud, 2017).

Filter bubbles focus on information filtered by the algorithm, echo chambers by the individual's social network and selective exposure by the individual themselves. The outcome is similar: individuals engage with information that is congruent to their

previous beliefs. Not surprisingly, it has been difficult to tease apart and study the mechanism behind these phenomena.

For the case of search engines, echo chambers are irrelevant, as web search is arguably not affected by the opinion of like-minded individuals. However, search engines have the potential of generating filter bubbles. Most of the recent empirical evidence indicates that such concerns are overstated (e.g. Fletcher and Nielsen, 2018; Scharkow et al., 2020; Stier et al., 2022), although theoretical models show that complex social-influence dynamics could amplify small effects (Keijzer and Mäs, 2022). Still, instead of feedback loops of information which limit the exposure to different perspectives, researchers find that search results are more likely to drive users to diverse sources (Fletcher and Nielsen, 2018; Scharkow et al., 2020; Steiner et al., 2022; Wojcieszak et al., 2022), despite there being evidence of a higher representation of "mainstream" news sources in search results (Nechushtai and Lewis, 2019; Steiner et al., 2022; Trielli and Diakopoulos, 2019).

The cited studies that challenge the existence of filter bubbles can also (explicitly or implicitly) challenge the effects of selective exposure. At the same time, selective exposure is supported by extensive research (for a review, see Stroud, 2017). More recently, Robertson et al. (2021) found ideological preferences in browsing data, that is, strong partisans chose to engage with substantially more partisan news than they were exposed to in their Google Search results. One type of preference that has not yet been studied using browsing data is familiarity to news sources, despite psychological theories that support its examination. The mere-exposure effect explains the development of preferences toward objects as individuals become familiar with them (Montoya et al., 2017). This is similar to the function of the familiarity heuristic, which in consumer behavior explains why individuals are more likely to consume products of the same brand (Park and Lessig, 1981). In addition, familiarity could also be taken as an indicator of trust toward specific news sources, and researchers have established a link between trust and "mainstream" news sources (Fletcher and Park, 2017).

In our work, we will look at individuals' familiarity with news sources, while considering the ranking of results and representativeness of news sources. Crucially, our measurement of familiarity is not self-reported but estimated from news visits unrelated to those driven by search engines. Methodologically, we improve on the previous works that study these factors by capturing the full content (HTML) of the search pages visits to estimate the exposure, and the browser's Tab Activity to identify the referrals; this is similar to Robertson et al. (2021), though we also identify the visited URLs (of the news article) in the list of search results to include the ranking of the selected results in our statistical model, instead of calculating a rank-weighted partisanship score for the entire search result page.

Research questions

In this study, our aim is to understand how Google Search shapes participants' news search and consumption across a 2-month period of data collection. Based on the theoretical frameworks regarding search engine information proliferation and filter bubbles introduced above, we pose several research questions. The main research question is

related to three relevant factors that we identified in our literature review, that is, rank, representativeness, and familiarity:

> *Research Question 1a (RQ1a).* Do rank, representativeness, and familiarity all influence news article selection?
>
> *Research Question 1b (RQ1b)*. Which has the strongest effect?

The second research question concerns itself with the existence of filter bubbles:

> *Research Question 2a (RQ2a).* Will Google Search create a filter bubble, or will it increase source diversity in news consumption?
>
> *Research Question 2b (RQ2b).* Will Google funnel participants to partisan news sources, or will it increase audience diversity?

Although ranking has been established as major factor on search results selection, it is not clear if this applies equally to news, which leads to a third research question:

> *Research Question 3 (RQ3).* Do ranking effects differ for news and non-news search results?

We also conducted post hoc exploratory analyses into our here developed familiarity metric, and present findings on reliability and consistency across our data set.

Methods

Data sources and participant demographics

We used data from a web tracking study that collected browsing activity of German citizens eligible to vote in the 2021 elections between 19 August 2021 and 27 October 2021. A total of 18,244 members of a commercial market research panel (*dynata*) received an invitation to participate in a longitudinal online survey (three waves) and to install a desktop Chrome/Firefox browser extension (Aigenseer et al., 2019); 7710 entered the survey, 739 completed the survey and installed the desktop extension. Exclusions occurred for the following criteria: (a) quota demographic cells were full, (b) individuals did not meet the necessary eligibility criteria to vote in the 2021 German election, (c) they did not use Firefox or Chrome as their primary web browser, (d) they did not install the extension, or (e) they withdrew early from the survey for other, self-selected reasons. The aim was to approximate a representative sample of the German population. Due to the above exclusions, lower educated individuals, and young and very old age groups are underrepresented.

Depending on participation days, participants were compensated with 25 to 75 EUR. In total, ~8.36 M page visits of the 739 participants (M = 11,311.07, Mdn = 7436, SD = 13,730.45) were recorded. After data cleaning, 1221 Google Search result pages,

comprising 5380 news articles, were included in the analysis; 280 participants were represented with at least one search page in which they selected at least one news article ($N = 1221$, M = 4.36, Mdn = 3, SD = 6.63). Of these, 122 identified as women, 155 as men, and one as non-binary. 60 participants were from East Germany and 218 from West Germany. 34 participants held at least elementary level education, 93 had a mid-level education, and 151 reported a high education level (high school or above). The sample's mean age was 46.73 (Mdn = 48, SD = 13.96, min = 18, max = 74). The sample's mean in terms of political alignment was 5.45 (Mdn = 6, SD = 1.88, left = 1, right = 10). The 280 participants conducted ~4.27M visits (M = 15,261.05, Mdn = 11,286.50, SD = 13,239.92), of which 75,911 (M = 271.11, Mdn = 90.5, SD = 448.59) were to German and English news websites, which we identified using three lists (i.e. AllSides, 2022; Robertson et al., 2018; Stier et al., 2020). These data represent a case study of German news consumers in web search, providing an opportunity to examine the empirical validity of existing theoretical concepts (Yin, 2009).

Data protection

Data collection was approved by an institutional ethics committee and conducted in line with institutional data protection and ethics regulations, including the collection of explicit informed consent with regards to participation in the surveys and in the web tracking procedure. The WebTrack plugin avoided collection of sensitive websites and information (pre-generated deny lists), and participants could at any time temporarily de-activate the tracking by switching on a private mode button within the plugin. Because web tracking data inherently carries the potential for de-anonymization, raw data access is restricted to researchers directly authorized by the project lead. Fully anonymized data sets can be made available upon request to reproduce the here presented analysis and plots.

News articles results and news domain visits data sets

Using the web tracking data, we generated five data sets for our analysis, namely: (A) *search news* (articles) *results*, (B) *all news visits*, (C) *Google-associated* (news visits), (D) *Google-independent* (news visits), and (E) *non-news* (visits). The *search news results* data set contains news articles results that appeared in Google Search pages. The *all news* data set includes all visits to news domains regardless of the way in which participants arrived at them. The *Google-associated* data set includes only those news domain visits that were referred by Google services. The *Google-independent* data set includes only those news visits that were not referred by Google services. The *non-news* data set includes all website visits that did not correspond to any news domains. The left side (I) of Figure 1 presents a schematic representation of the data set that we will further explain.

To generate the *search news results* data set (identified with the letter A, orange, Figure 1), we followed these steps.

First, we extracted (from the *all news* subset, explained below) the news pages at which participants arrived directly from Google results. Second, we parsed the HTML of the pages corresponding to the Top-10 Google Search results (i.e. "organic" results or "blue

links") and extracted related data (including their ranking and URL). Third, we flagged the results that were selected ("clicked on"); a result was considered selected if the URL of a subsequent news visit corresponded to the URL of the result (redirects and cookie consent pages were omitted to the corresponding sequence between the result selection and the news article). Fourth, we discarded search pages resulting from search query terms that included references to news sources; we automatically discarded queries that included news root and main domains. For example, "elections 2021 bild" or "elections 2021 bild.de" would be discarded, because "bild"/"bild.de," referring to the German newspaper Bild, formed part of the query term. We manually inspected the search queries to include typos (e.g. "bild:de" instead of "bild.de") as well as other news source names (e.g. "Frankfurter," which refers to the German newspaper "Frankfurter Allgemeine" and whose domain is faz.net). Fifth, we manually annotated the URLs of the Google results, and (a) discarded results that did not correspond to news articles, and (b) full Google Search pages that did not contain at least two news articles of different domains (otherwise making it impossible to compare news sources). For the annotation, a URL was considered an article if (a) it was hosted by a news domain and (b) the content (including text and video) informed potential readers about a non-fictional topic (opinion pieces such as commentaries or documentaries are included). One of the authors inspected all URLs (reviewing the content when it was unclear); the remaining ambiguous cases were discussed and resolved with the other author. Sixth, we removed repeated search behaviors, that is, a participant searching for the same query term and selecting the same result.

To generate the *all news* data set (identified with the letter B, blue, in Figure 1), we used three lists of news domains (AllSides, 2022; Robertson et al., 2018; Stier et al., 2020) and so identified all news-related visits. The *Google-associated* data set (identified with the letter C, purple, in Figure 1) is a subset of the *all news* data set in which we only included visits that were driven, directly or indirectly, by Google (including subdomains such as maps.google.com, and country-level domains such as google.de or google.co.uk). We defined that a participant arrived directly from page x to page y if the browser tab that displayed x was opened from a tab displaying y; we then pose that this is a direct path from x to y (denoted $x \rightarrow y$). We defined that a participant arrived indirectly from page x to page y if (a) there existed a sequence of direct paths that led from x to y, that is, $\exists \{z_1, ..., z_n\} : x \rightarrow z_1, z_1 \rightarrow z_2, ..., z_{n-1} \rightarrow z_n, z_n \rightarrow y$, and (b) $z_j \in \{z_1, ..., z_n\}$ belonged to the same web domain d. The path $x \rightarrow y$ exists if y was opened from a tab displaying x. The reason for this is that the browser extension can identify the tab from which a visit was initiated (whether if the visit is open in the current tab or in a new one—including tabs on new browser windows), but it cannot identify if the visit was caused by clicking on a link or because the participant manually typed a URL in the navigation bar. For the same reason, we excluded paths originating from a Google home page URL (e.g. https://www.google.com/), as we can safely assume that home pages do not contain news links—it is more likely that a participant's browser is configured to have the Google home page as its default page.

The *Google-independent* data set (identified with the letter D, green, in Figure 1) was derived by excluding the rows in the *Google-associated* data set from the *all news* data set (i.e. B—C).
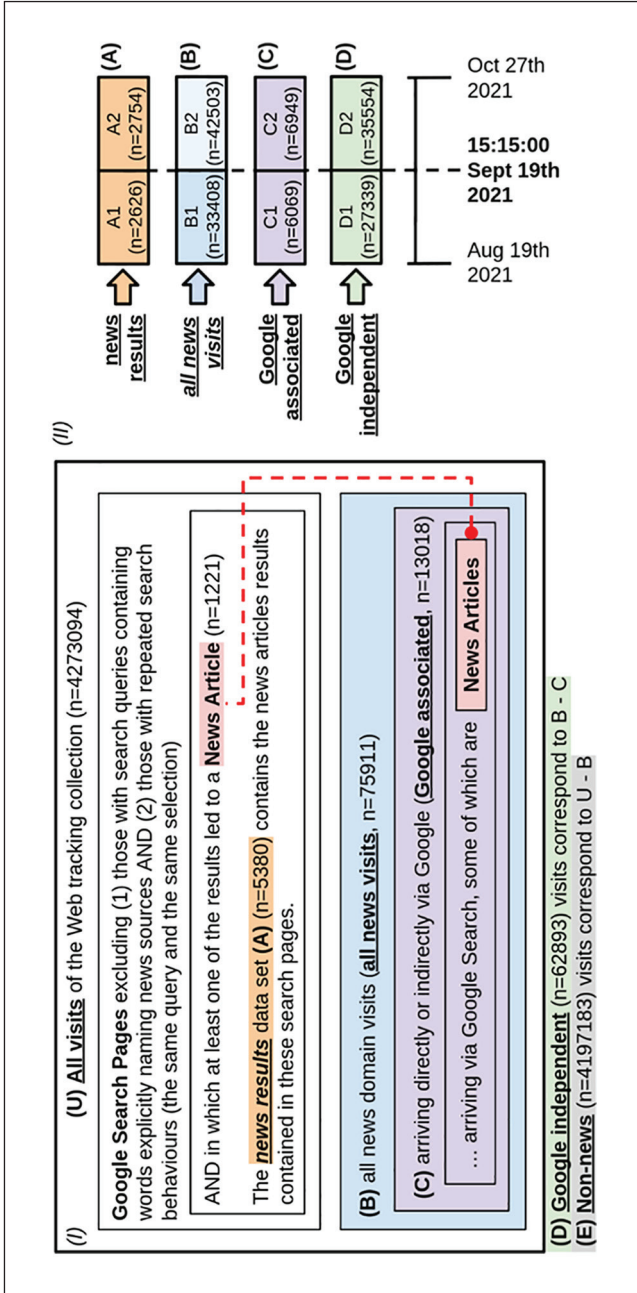
Figure 1. Diagrams of the data sets extracted from the web tracking data. The left figure (I) represents all the pages in the web tracking collection. Rectangles inside represent subsets: Google Search pages–related subsets and news-related subsets. The colored rectangles and upper-case letters in parenthesis highlight relevant data sets: blue (B) for *all news* domain visits; green (D) for the *Google-independent* news domain visits; red (D) for the Google Search domain visits; purple (C) for the *Google-associated* news domain visits; and orange (A) for the set of Google Search results that correspond to news articles. The diagram describes basic conditions used to filter the data sets. The right figure (II) represents the way in which the data sets are split, so that it is possible to create time-independent models using the data before the cut-off date (19 September 2021, 15:15:00) to measure predictors (independent variables) that could explain responses (dependent variables) in the data after the cut-off date. Colors in Figure (I) and Figure (II) are aligned and used consistently throughout figures in this article.

We considered statistical tests time-independent using the portion of the *all news* and *Google-associated* data sets that was collected before the cut-off date (19 September 2021, 15:15:00), and the *news results* data set that was collected after the cut-off date. For example, our main model measures the *news engagement* in the data set B1, and it is used to predict the news article selection in the A2 data set. The cut-off date (19 September 2021, 15:15:00) splits in half Google Search pages corresponding to the *news results*.

The *non-news* data (identified with the letter E, gray, in Figure 1) are derived by excluding the rows in the *all news* data set from the entire Web tracking collection.

Definition of news engagement, representativeness, and ranking

News engagement. We used participants' news visits to construct a behavioral indicator of news source engagement, which measures the strength of the relation between a news source and a participant based on their browsing history. Let $u_1, u_2, ..., u_k$ be the list of URLs corresponding to the news visits of a participant p—then the engagement of p with the domain d is given by

$$\frac{1}{k}\sum_{i=1}^{k} S(u_i, d)$$

where $S(u,d) \epsilon \{0,1\}$ indicates if the domain of the URL u corresponds to $d(1)$, or not (0).

Representativeness. Let $T = t_1, t_2, ..., t_k$ be the list of URLs corresponding to the results shown in one Google Search result page. The representativeness of a news source, which domain is d, is given by

$$\frac{1}{k}\sum_{i=1}^{k} S(t_i, d)$$

where S was defined previously. Note that k is usually 10, as it corresponds to the results displayed on the first result page.

Ranking. The ranking is given by the position of the news article on the Google Search page. Note that, idiomatically, the result with lowest ordinal value (i.e. top-1 result) is referred to as the result with the highest ranking. For the purposes of readability, we keep this convention in all our analyses, including the odds ratios (OR) reported in the binomial linear mixed-effect model regression tables, that is, positive values indicate that a higher ranking increases the probability of a result being selected.

Modeling news article selection

To answer the research question that motivated this investigation, we used a binomial linear mixed-effect model fitting the interaction between the three study factors (ranking, representativeness and engagement) while controlling for the following individual characteristics that were centered and scaled before use: is a woman (yes: 1, no: 0), education

(low: 0, middle: 1, high: 2), from region (East Germany: 0, West Germany: 1), age (continuous categorized), political alignment (10 point scale), and election interference (as a major election occurred during data collection): whether the search was performed on the day of the election (yes: 1, no: 0), and the day before or after the election (yes: 1, no: 0). We defined random intercepts for repeated measures (participant, search page, and news domain). The dependent variable was binary, with 1 indicating that the participant selected ("clicked on") the result, and 0 that they did not.

To avoid time dependencies (and thus correlational effects), we split the collection into before and after a cut-off date (19 September 2021, 15:15:00). For the main analysis, we used the *all news visits* before the cut-off date (B1 in Figure 1) to calculate the news engagement, and the Google *news results* after the cut-off date (A2 in Figure 1). Data set A2 comprises 2754 search results representing 611 search pages, and 194 participants: 90 women, 1 diverse person, 103 men; 41 from East Germany, 152 from West Germany; 23 of low, 65 of middle, and 105 of high education. The sample's mean age was 46.26 (Mdn=47, SD=13.89, min=18, max=73). The sample's average political alignment was 5.34 (Mdn=6, SD=1.93, left=1, right=10). We fit the model two more times according to data sets used to calculate news engagement: *Google-associated* (C1) and *Google-independent* (D1).

To further explore the main result, we fit a time-dependent model using the data of the entire period, that is, we used all the Google *news results* in data set A and calculated the engagement using *all news visits* data set (B), while excluding visits that were, directly or indirectly, driven by the selected *news results* (A) (i.e. excluding cases that we are inferring in the *news engagement* predictor). We also fit the model for the news engagement calculated on the *Google-associated* (C) and *Google-independent* (D) data sets. The "Methods, *Data sources*" section already reports relevant demographic distributions for this data set.

Google Search and filter bubbles

Source diversity of consumed news. To study the diversity of the news sources that the participants visited, we looked at the number of unique domains. As a baseline, we first looked at the diversity in the *Google-independent* data set (data set D). We then reported the diversity attributed to Google Search, that is, we measured the increase of diversity when the domains of the search results selected (i.e. "clicked on") by the participants (and consequently also visited) in the *news results* data set (A) were added to the domain visits of the *Google-independent* data set (D). We also measured the diversity of the exposure in Google Search by looking at all the domains in the *news results* (data set A).

Partisanship audience. To explore how Google shapes the audience that news sources receive, we measured the variance of the political alignment of the participants that visited the news domains. Accordingly, we analyzed two types of visits: *Google-independent* visits (data set D), and the specific news visits resulting from the selection among the Google Search results (Google Search driven visits). We used variance because Bhadani et al. (2022) found it to be the best measurement of dispersion to predict the quality of a news source.

Ranking relevance for news articles

To analyze the importance of ranking for the results containing news articles, we compared the ranking of the news articles selected by the participants (in the *news results* data set A) with the ranking of non-news results selected by the participants, that is, including all Google Search pages in the web tracking collection. First, we parsed the HTML of the pages corresponding to the Top-10 Google Search and extracted all their results (including their ranking and URL). Second, we identified those results that were selected (clicked on), so checked if there exists $x \rightarrow y$ where x is the Google Search page and y the page visited after. Third, we discarded search pages resulting from search query terms that included any domains present in our collection ("xyz wikipedia" would be discarded because "Wikipedia" is included in the term), as we considered this a strong signal of a predisposition toward a certain domain. Fourth, we removed repeated search behaviors, that is, a participant searching for the same query term and selecting the same result twice. We used a two-sample *t*-test to establish if there was a significant difference between selecting the first search result for news and non-news, and reported a Spearman' rank correlation r_s to test whether the search engine ranking was biased toward news domains.

Post hoc analyses

Consistency of news engagement. We used two measures of similarity to compare two lists of visits: Jaccard similarity (aka Jaccard Index [JI]) and rank-biased overlap (RBO). Let F and G be, respectively, the sets of news domains corresponding to the lists of visits, then the Jaccard similarity between F and G corresponds to

$$J(F,G) = \frac{|F \cap G|}{|F \cup G|}$$

By definition, $J(F,G) = 1$ if F and G are both empty. Jaccard only considers the overlap in the domains and ignores the importance (rank) of each domain according to, for example, the frequency of visits. Thus, we complemented the JI using RBO. Let F and G be sequences of domains ordered by their rank (i.e. by the number of times each domain was visited), then

$$RBO(F,G,p) = (1-p) \sum_{d=1}^{\infty} p^{d-1} \cdot A(F_d, G_d)$$

where F_d and G_d denote the sequence until the element in the position d, the parameter p determines the weight given to the top results (the lower, the more importance is given to the top results), and A is the proportion of agreement until the element in the position d, formally

$$A(F_d, G_d) = \frac{|F \cap G|}{d}$$

In all cases, the similarity values were calculated within participants and 95% boot-strapped confidence intervals are presented in the plots.

We used these similarity metrics to assess the consistency across time of the *news engagement* measurement as a predictor by comparing the *Google-associated* visits before the cutting date (data set C1) and after the cutting date (C2), and the *Google-independent* visits before the cutting date (data set D1) and after the cutting date (D2).

Correlations of visits. To validate the findings, we used Spearman' rank correlations (denoted by r_s) and explored the connection between the within visits in two data sets. Spearman correlation is well-suited to study the relation between the frequency distributions of the domain visits, as the domains are sorted (i.e. ranked) according to their number of visits (i.e. the most visited domain has a higher rank). We report correlations between domain visits within participant (participant-domain pair) in the following data sets: (a) before and after the cut-off date in all news visits (B1 vs B2), (b) in *Google-independent* visits (D1 vs D2), and (c) in *Google-associated* visits (C1 vs C2). We also report the overall participant visits to news in the following data sets: (a) before and after the cut-off date in *Google-independent* visits (D1 vs D2), and (b) in *Google-associated* visits (C1 vs C2).

Comparison of distributions. We used a Kolmogorov–Smirnov test of equality of distribution (the statistic "difference" is denoted by D_{KS} in the text) (a) to check whether Google Search–driven news visits effectively shifted the diversity of *Google-independent* visits and (b) to check that the splits by cut-off dates were balanced between C1 and C2, and D1 and D2.

Results

Familiarity with news sources does not predict news article selection, while rank and representativeness do

We investigated which of three relevant factors best predict news article selection in Google results: *news* (source) *engagement*, *ranking*, or *representativeness*. We use *news engagement* as a proxy for a participant's familiarity with each news source. The dependent variable (*news selection*) is binary (0,1), representing news articles that appear in the first Google Search results page: we coded each available news result by indicating whether the participant selected it (1) or not (0).

For the main analysis, the web browsing data is split in half to resolve time-dependence between *news engagement* (using data set B1 in Figure 1) and *news selection* (A2); for details, see "Methods." We find a significant effect of *ranking* (OR = 2.95, CI = 2.61, 3.34], $p < .001$), and *representativeness* on article selection (OR = .77, CI = [.66, .90], $p < .001$) (see "Methods, *Modeling news article selection*" for details on the fitted models, and Supplementary Material S1 for a full regression summary). We do not find a significant effect of *news engagement* on article selection (OR = .93, CI = [0.81, 1.08], $p = .351$).

We then conducted a variety of tests to check the robustness of our effects. First, we calculated a *Google-independent news engagement* measure, that is, we excluded all traffic referred by Google when calculating news engagement (using C1), and a *Google-associated news engagement*, that is, we included only traffic referred by Google (using D1). Then, in the linear model, we replaced *news engagement* with either *Google-independent news engagement*, or *Google-associated news engagement*. In both replacements, we corroborated the above results (see Supplementary Material S2 and S3).

Second, we tested whether a lack of power might be the reason for the non-significance of the effect of *news engagement* by using all the data (ignoring the splits by cut-off date). We replicated the same pattern as in the previous models, that is, for the larger model (i.e. using data set B for *news engagement* and A for *news selection*), we found significant effects for *ranking* (OR=3.1, CI=[2.83, 3.39], $p < .001$) and *representativeness* (OR=.75, CI=[0.65, 0.85], $p < .001$), and, additionally, a significant interaction of them (OR=1.15, CI=[1.01, 1.32], $p = .038$). We again did not find a significant effect of *news engagement* on article selection (OR=1.03, CI=[0.95, 1.12], $p = .468$). See Supplementary Material S4–S6 for the full regression summary of models using data sets B, C, and D to calculate news engagement.

Google Search increases source diversity in news consumption

Ranking and *representativeness* are both determined by Google's algorithms—so the results above demonstrate the leverage that Google Search has in driving users toward news sources. We investigated the implications of this leverage by looking more closely at the source diversity of news visits across the entire tracking period.

We logged 75,911 total visits to news domains. Of these, 13,018 were referred directly or indirectly by Google services (see "Methods"), including 1269[1] news articles specifically by Google Search. A total of 62,893 news pages were visited by participants independently of Google.

We first investigated individuals' news consumption assuming an Internet without Google. Plot A in Figure 2 presents the distribution. The visible skew indicates that participants repeatedly visited their own preferred domains: on average, each participant's top-3 news domains concentrate 87.90% (Mdn=95.65, SD=15.95) of their news consumption.

We then investigated how Google Search exposes users to news. Plot B presents the distribution of how often news domains are presented in the search results. The distribution is flatter (compared to Plot A); on average, the top-3 most displayed news domains on each participants' Google Search page concentrate 56.09% (Mdn=47.14, SD=27.11) of their news exposure.

Finally, we tested how Google Search contributes to source diversity of news consumption. Plot C presents the increase in news visits when adding Google Search to participants' otherwise *Google-independent* browsing experience. We find a statistically significant difference between the two, $D_{KS}(278) = .125$, $p < .0251$, with an average increase of ~32.35%, corresponding to 2.88 (Mdn=2, SD=2.94) added unique domain visits (increased from 8.89, Mdn=6, SD=9.25). This increase in source diversity is
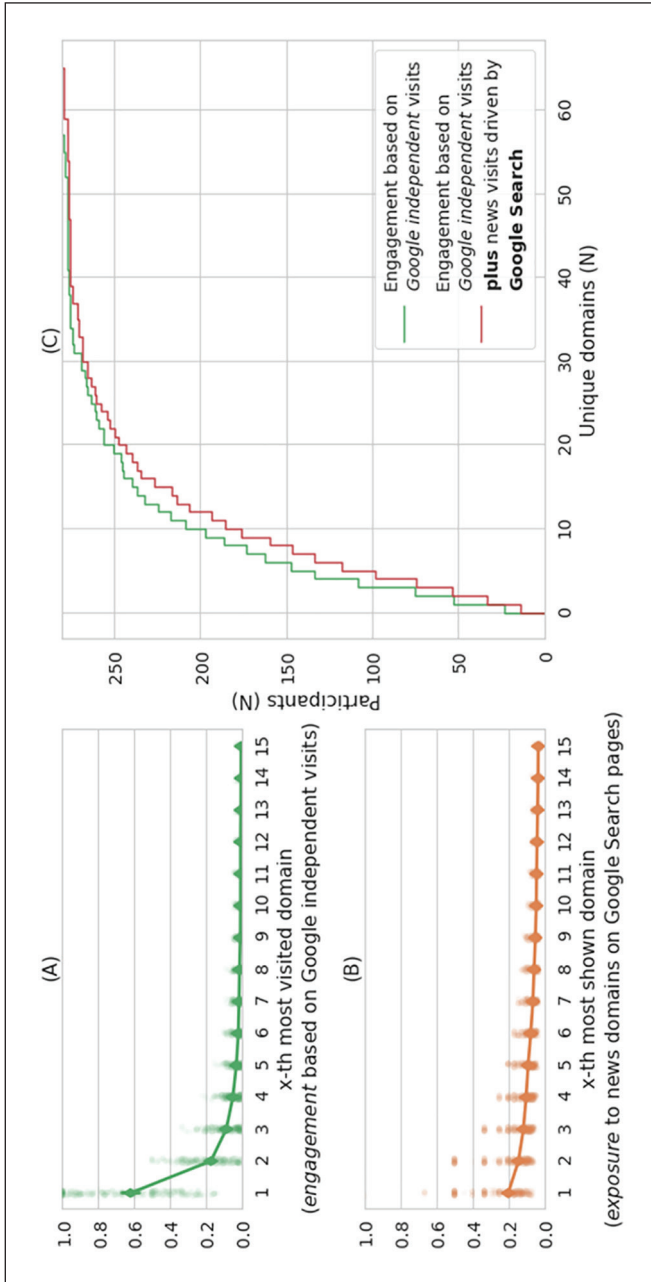
Figure 2. News diversity with and without Google Search. The plots display distributions of (A) rank of most shown domain on the search results, and (C) number of unique domains visited. Plot A shows the proportion of domain visits to total visits (Y-axis) by most visited domains (X-axis, truncated to top-15 out of 57); green data points represent the participants, and the green line is the median. Plot B shows the likelihood of a domain to be shown on the first page of Google Search results (Y-axis) by most shown domains (X-axis, truncated to top-15 out of 63); red data points represent the participants, and the red line is the median. Vertical lines in plot A and B display confidence intervals at 95%. Plot C shows the cumulative distribution of participants (Y-axis) to the number of unique domains (X-axis): the green line represents Google-independent visits and red includes visits driven by Google Search.

achieved with only 2.02% additional visits (1269 to the 62,924 *Google-independent* visits).

Google Search increases audience diversity in terms of partisanship

We explored the political alignment of participants in relation to their news consumption to better understand how Google Search shapes the audiences of different news sources. Again, we first looked at browsing behavior supposing an Internet environment without Google. Audience diversity had a mean of 2.92 (SD = 1.81), indicating a relatively homogeneous readership to each news source. We then looked at browsing behavior through Google Search, which had a mean of M = 3.59 (SD = 2.84), significantly higher, $t(236) = 2.22$, *two-sided p* = .027. This indicates that Google Search drove a more diverse partisanship audience to news sources. This difference is even larger if we look at the page views (visits) instead of unique participants; $t(271) = 4.49$, *two-sided p* < .0001. Figure 3 illustrates this result.

Comparing news and non-news: ranking matters less for news articles

Consistent with previous literature (Wojcieszak et al., 2021), we found that news consumption differs widely across participants. Across the tracking period, the median frequency of news visits in our subsample was 90.5 (M = 271.24, SD = 448.92, N = 280); this represents 2.32% (Mdn = 1.02, SD = 3.89, N = 280) of the overall browsing. 10.36% (29 of 280) of participants had less than 10 news visits and 6.76% (19 of 280) of participants visited more than 1000 news sites in the ~2 months period. Over this period, 8.21% (23 of 280) did not visit any news sites by means other than Google (i.e. without being referred by Google).

We investigated how Google Search ranking differs in its effect on participant's choices in news articles versus non-news articles. We first looked at the top position (first ranking). We found that non-news are selected in 46.71% of cases (Mdn = 45.61, SD = 16.32), while news are selected 31.96% of the time (Mdn = 25, SD = .33), a difference that is statistically significant $t(557) = 6.68$, *two-sided p* < .0001. Plot (A) in Figure 4 shows the proportions for each Google Search ranked result (black and red lines). We hypothesized that this effect might be driven by Google ranking news articles lower than non-news on average. Contrary to this hypothesis, we find that the top results are slightly more likely to showcase news, $r_s(1928) = -.11$, *two-sided p* < .0001.[2]

Familiarity: a matter of consistency and reliability

In this section, we will further explore the underlying patterns behind our proxy for familiarity: news engagement. In the first results section, we did not find significant evidence supporting the hypothesis that news engagement impacts news selection when individuals are using Google Search. In this section, we show that, comparing behavior across two time periods, participants were not very consistent regarding the news sources they consumed. For the time periods, we employed the same cut-off date as for our first model testing news selection.
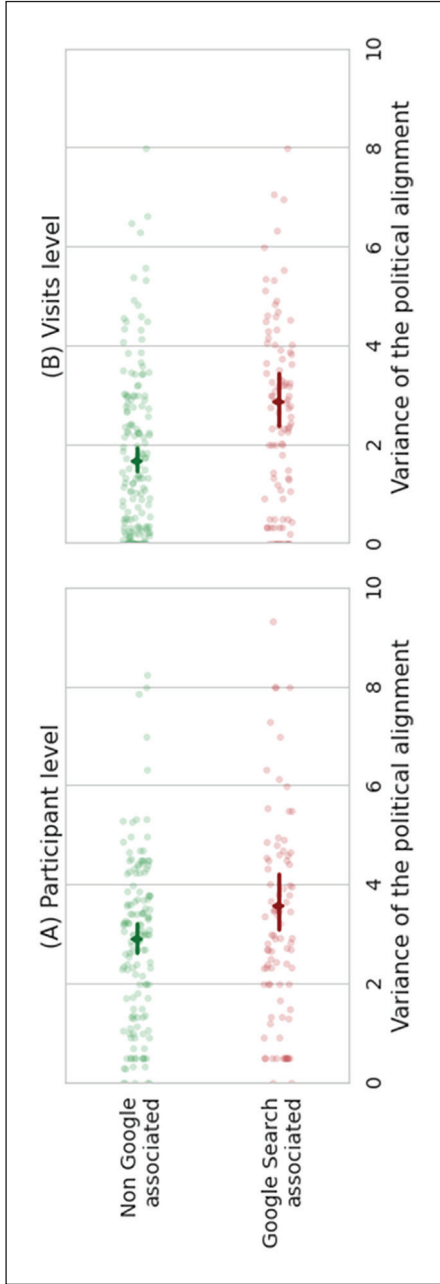
Figure 3. Partisanship audience of *Google-independent* and Google Search visits. The left plot (A) shows the variance of the political alignment (X-axis) of the unique participants that visited each news domain according to the association of the visit to Google (Y-axis): *Google-independent* and Google Search associated. The right plot (B) shows the equivalent plot for each visit (page view). Confidence intervals at 95% are represented.
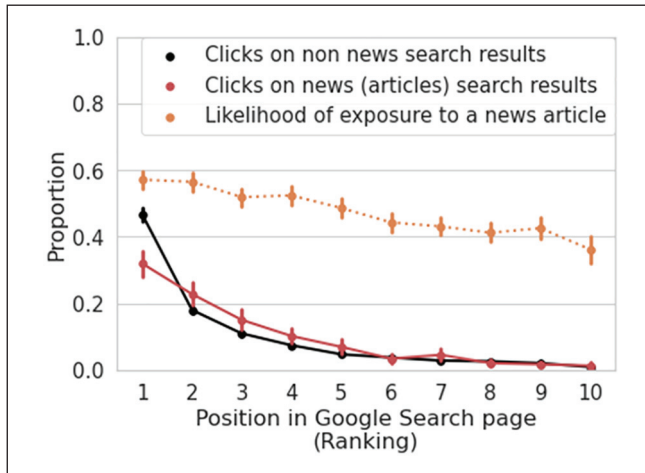
Figure 4. Selection of results according to Google Ranking. The *X*-axis indicates Google ranking of results presented on the search page. The *Y*-axis displays either the proportion of result selections across participants (black and red lines), or the likelihood of being exposed to a result leading to news articles in pages that contain at least one news article (orange dotted line). Vertical lines represent confidence intervals at 95%.

First, we found a moderate correlation of the visits for each participant-domain[3] pair, $r_s(62,438)=.44$, $p=.0001$. Knowing the leverage that Google has in driving users to news domains, we explored whether the consistency of news consumption is shaped differently depending whether Google services (Google Search, Google Maps, and Google News) do or do not drive participants to news. When looked at individuals' news engagement supposing an Internet experience without Google (*Google-independent* news), removing all Google referred browsing, we found a stronger correlation; $r_s(62,438)=.52$, $p<.0001$.

These correlations are difficult to interpret because most domains were never visited by most of the participants (i.e. a large amount of 0s in the vectors might inflate the correlations). Thus, we used the JI, a measure of similarity between data sets to establish how similar participants' news visits were for the two timespans. We defined *thresholds* t and for a given t, we only included participants that had at least t visits before and t visits after the cut-off date. We found that participants' behavior was robust (JI > .42, green bars in Plot A of Figure 5), especially considering that news visits are rare for some participants. We also calculated the corresponding RBO (Rank Biased Overlap) (Webber et al., 2010), a similarity metric which uses a parameter p to add a weight of the rank of each domain based on the participants' number of visits. The results for various values of p and t are shown in the green trace in Plot B of Figure 5. The similarity is higher when the ranking is considered, especially for the top positions (lower values of p).

We finally turned our attention to participants' *Google-associated* news visits, that is, news visits directly or indirectly referred by a Google service; in this case we found a much lower correlation, $r_s(62,438)=.29$, $p<.0001$. Consequently, the similarity of
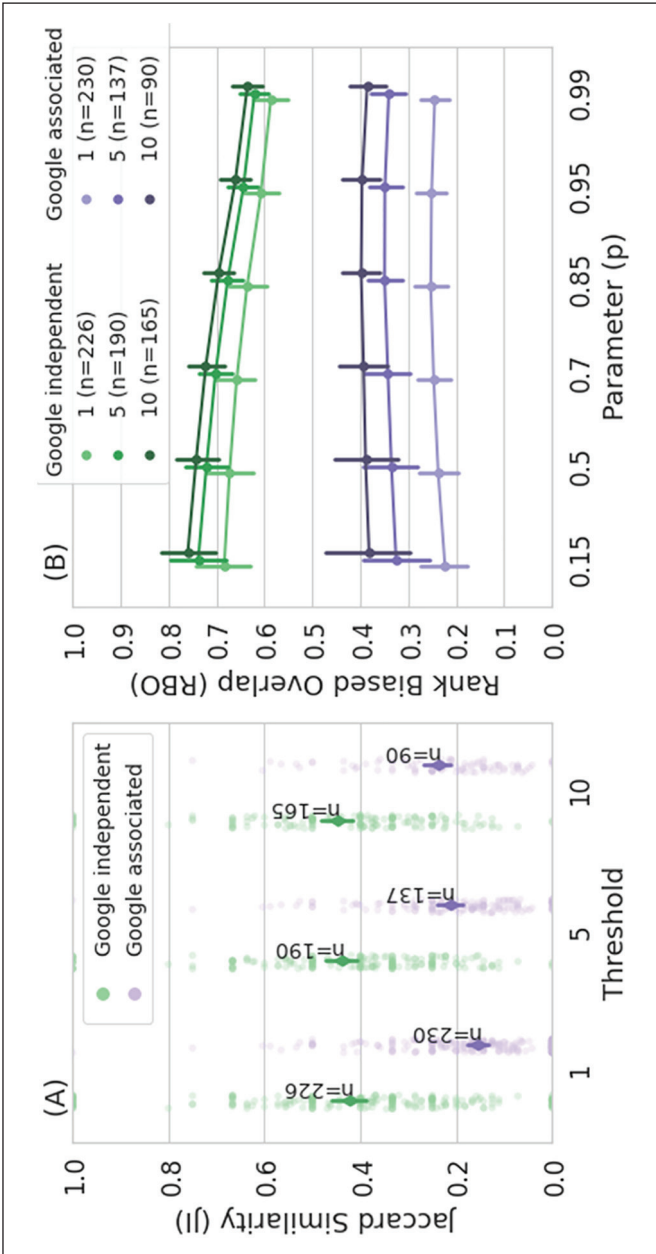
Figure 5. Similarity of the engagement measure across time. Plots A and B measure the average similarity of news engagement between visits before and after the cut-off date (calculated per participant). Plot A shows the Jaccard Similarity (JI, Y-axis) of the news engagement for *Google-independent* news visits (green), and for *Google-associated* visits (purple); it also shows the JI similarity of the news domains that were shown in Google results (exposure, orange). The X-axis indicates different thresholds for filtering out participants based on their total visits (the number of participants that fulfill the threshold, *n*, is displayed over the bar). Plot B shows the corresponding RBO scores (Y-axis) with results for different values of the parameter *p* (X-axis). The legend for B includes the corresponding thresholds and participant numbers (*n*). Vertical lines represent confidence intervals at 95%.

participants' news visits was lower for the two timespans: JI < .24 (purple bars in Plot A of Figure 5) and RBO < 0.4 (purple in Plot B of Figure 5). Here, the RBOs were stable for varying p parameter values, indicating that the consistency did not improve even for the most visited domains.

The volume of news consumption of the participants was consistent. For the Google-independent visits, we ran a Kolmogorov–Smirnov test, and we did not find evidence supporting different distributions between the visits per participants before and after the cut-off date; $D_{KS}(278) = .043$, $p < .960$ (see distribution, Plot A, in Figure 6), and the correlation was very high, $r_s(278) = .85$, $p < .0001$ (see scatterplot, Plot C, in Figure 6). Similar results hold for *Google-associated* visits: we did not find evidence supporting different distributions, $D_{KS}(278) = .043$, $p < .960$ (see Plot B in Figure 6) as well as a high correlation, $r_s(278) = .56$, $p < .0001$ (plot D in Figure 6).

The lack of consistency between news domain visits within the Google ecosystem is in line with results we present in the previous sections, which have suggested that Google Search facilitates more diverse news selection for its users. The overall results suggest that while participants' overall volume of news consumption is consistent, and that participants do, on their own, have preferences for familiar news sources (as indicated by the similarity of sources visited across time), this familiarity loses relevance when they browse for news while integrated in the Google ecosystem, and in particularly when choosing their news in Google Search.

Discussion

We tested the hypothesis of news engagement (as a proxy for familiarity with news sources) as a predictor of news article selection in the Google search engine (RQ1). We did not find evidence supporting this. Instead, we found a significant effect for two factors that are decided by the search engine alone: the position in which the result is presented (ranking) and the number of times the news source appears (representativeness). While ranking has previously been demonstrated to play a strong role (Pan et al., 2007; Urman and Makhortykh, 2021), we show for the first time that its effect is weaker for news article selection compared to non-news selection (RQ3). This may well suggest a more careful decision-making process of individuals when selecting news (e.g. reading the titles and excerpts more attentively).

Research has indicated a higher representation of "mainstream" news sources in search results (Puschmann, 2019), while, at the same time, a positive effect in the diversity of news consumption (Fletcher et al., 2021; Fletcher and Nielsen, 2018). Our results align with this seemingly counter-intuitive evidence: representativeness reduced the likelihood of news article selection (RQ1). This might be an indication that once individuals have decided not to visit a result belonging to a specific news source, they also discard subsequent results from the same source, suggesting that the individual is actively avoiding such sources (Mukerjee and Yang, 2021).

In line with previous research (RQ2a), we found that Google Search increases the diversity of participants' news consumption (Fletcher et al., 2021; Fletcher and Nielsen, 2018; Scharkow et al., 2020). It is possible that participants use Google Search when they are actively looking for novel news sources, though we also show that Google
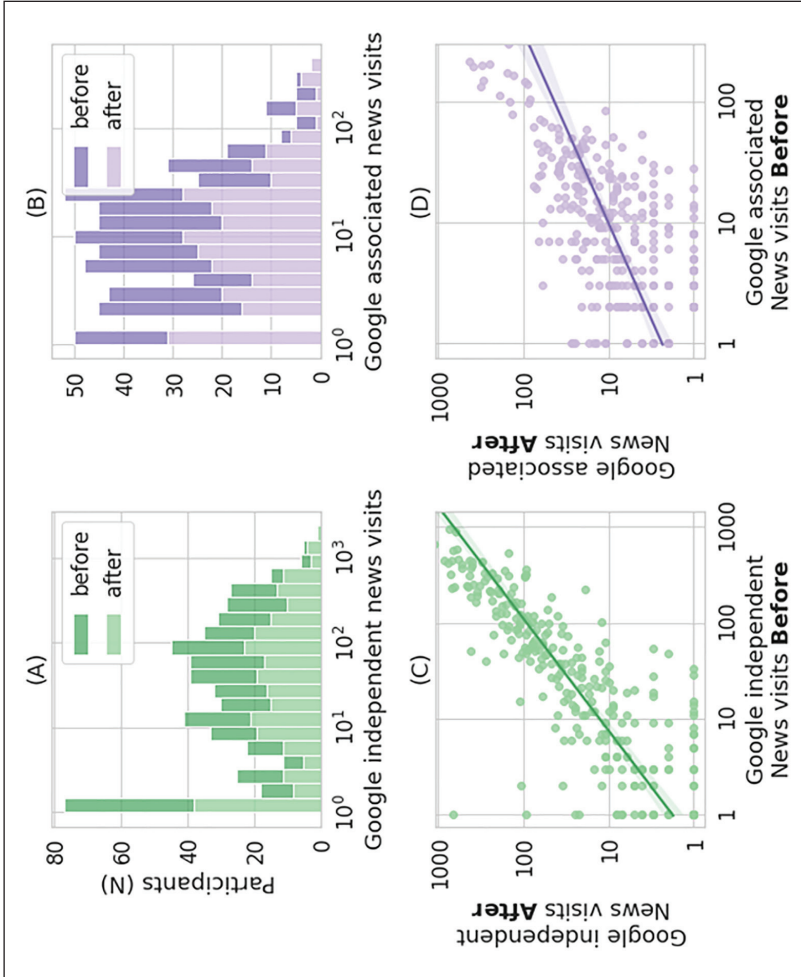
Figure 6. News related web activity. Plots A and B show the distribution of participants (Y-axis) according to their number of *Google-independent* (A) *and Google-associated* (B) domain visits (X-axis, log-scale); the dark and light sections (stacked) distinguish between the visits before and after the cut-off date. Plots C and D show the scatterplots of the visits before and after for *Google-independent* news visits (C) and *Google-associated* news visits (D), both axes are logarithmic. The semi-transparent bands are confidence intervals at 95%.

Search facilitates a discovery process by presenting a variety of news sources among the results. In addition, we show that Google Search increases the political audience diversity that news sources receive (RQ2b). Given that Google has its own news quality controls in place (Google Developers, 2021), the finding can explain recent research showcasing that political audience diversity can be used as a sign of news source reliability, and that it should be incorporated into ranking algorithms (Bhadani et al., 2022). Instead, our results suggest that it is Google Search (including its ranking) that drives this effect. More broadly, researchers should consider that online news browsing behavior is heavily shaped by online platforms, for example, we demonstrated that there are differences in the consistency of our familiarity metric depending on whether it is measured including traffic referred by Google or not. Considering the dominance of a few large search engines and their role in driving users to news, the situation may lead to a concentration of power and influence within the media landscape in which news organizations prioritize visibility on search engines (and similar platforms) by following SEO guidelines, instead of focusing on journalistic norms (Hackett, 1984; Muñoz-Torres, 2012).

Furthermore, we build on previous research investigating the phenomenon of mere-exposure (Montoya et al., 2017) and trust in news sources (Fletcher and Park, 2017), and investigated familiarity (RQ1), that is, the participants' acquaintance with the news sources they are presented in the search results. We proposed news engagement as a proxy and measured it using a section of the browsing history that is independent from the analyzed news articles that are selected (and visited) in the search results. This allowed us to quantify the existent relationship between the individual preferences toward news sources through the number of visits of the individual to each news domain; thus, capturing three modes of news engagement: routinary visits, social media referrals, and intentional search (Möller et al., 2020).

To summarize, we find no evidence in favor of filter bubbles, nor do we find an effect of selective exposure toward news sources based on familiarity. We find that individuals place their trust in Google, and that in turn, Google steers them toward sources that are different to their routinary visits. In this process, we tackle the challenge of developing reliable measurement models when integrating social science theory with digital behavioral data (Wagner et al., 2021).

Limitations and future directions

We would like to point out several limitations. First, our engagement metric does not fully capture when individuals are familiar and trust a given source, that is, when they do not regularly consume news or mainly do it offline. In addition, the engagement for the time independent analysis was calculated using an arguably short period of time (~1 month). Future data collections should enable analyses across timespans of multiple months up to years, though we highlight that our results did not change when considering a time-dependent analysis (~2 months).

Second, our analysis is limited to Google, which we chose due to its market dominance (StatCounter, 2021). Our findings should not be generalized to other search engines due to differences in how search results are displayed, though it is likely that effects specifically related to ranking and representativeness will remain similar

across different interfaces. Our findings should also not be generalized to contexts beyond the top-10 organic results of web searches, such as carousels and top stories, news aggregators (Google News), or recommendation systems such as Google Discover. The latter does not provide results based on prompts input by the user, but presents content (including news) personalized from individuals' preferences, which may be extracted from previous search behavior. Since search choices are heavily influenced by ranking, the algorithmic representations of preferences might also reflect the ranking effects as individual traits. This raises questions about the extent to which preferences can be captured by these systems, or if they merely expand the influence of the search engine. It highlights a challenge for study of new media: how can we better understand interdependencies between the digital services, rather than studying them in isolation?

Third, characteristics of our sample should be kept in mind when interpreting our results. The sample size used in the above research is relatively small. Out of the 739 individuals that participated in the web tracking study overall, only 280 are represented in our subsample, due to a relatively low number of news visits (including the ones driven by Google Search) in web tracking data—consistent with previous literature (Scharkow et al., 2020; Wojcieszak et al., 2021), and the strict data quality constraints for including a search page for the analysis. Our sample is relatively uniform, only including German individuals with a Chrome or Firefox browser installed on their desktop computers; while this reduces noise, it also affects generalizability of the findings. Finally, many individuals refuse to participate in web tracking studies due to privacy concerns (Makhortykh et al., 2021), which might indicate that the sample is pre-selected based on factors that we don't yet fully understand and cannot control for.

Despite these limitations, the logged web browsing behavior that we capture occurs in a real environment with minimal intervention; we argue that cognitive awareness of the presence of the web tracker is likely to only affect the very initial browsing behavior. Moreover, web tracking studies that include the website content remain rare, and our method is exceptional as it deterministically identifies referrals by tracing the tab activity of the browser and matching the presence of the URLs among the results.

Some promising directions emerge for new media research. Our results suggest that ranking is less important for news selection: research should further confirm if individuals are more selective when choosing news from the top-10 search results, which could for example be achieved by analyzing the duration of the interaction with the individual choices or by applying eye-tracking methodologies. In addition, we find that representativeness is negatively associated with news selection. A further inspection of the reasons is of interest, for example, this behavior might signal an active avoidance of specific news sources (Mukerjee and Yang, 2021). Finally, it is important to examine search choices within the broader context of the layout of web search pages, including features like video carousels and top stories, to gain a deeper understanding of how individuals make decisions for information acquisition.

Acknowledgements

Data availability

Scripts to reproduce the analysis and visualizations are available at https://github.com/gesiscss/news_selection and fully anonymized data sets can be shared upon reasonable request to the corresponding author.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

ORCID iDs

Roberto Ulloa (iD) https://orcid.org/0000-0002-9870-5505

Celina Sylwia Kacperski (iD) https://orcid.org/0000-0002-8844-5164

Supplemental material

Supplemental material for this article is available online.

Notes

1. The number of search pages ($N=1221$) is different from the selected news articles ($N=1269$) because participants could have selected more than one article on a given search page. The referrals to news articles reported included here only include those in data set A (see "Methods" section).
2. This test includes only Google Search pages with exactly 10 results to avoid over-representation of top-results due to incomplete pages.
3. In this data set, we only included news domains with at least one visit.

References

Aigenseer V, Urman A, Christner C, et al. (2019, September 24) *Webtrack–Desktop Extension for Tracking Users' Browsing Behaviour using Screen-Scraping*. GESIS Computational Social Science (CSS) Seminar, Köln.

AllSides (2022) Media bias ratings. Available at: https://www.allsides.com/media-bias

Asch SE (1946) Forming impressions of personality. *Journal of Abnormal and Social Psychology* 41: 258–290.

Bakshy E, Messing S and Adamic LA (2015) Exposure to ideologically diverse news and opinion on Facebook. *Science* 348(6239): 1130–1132.

Bhadani S, Yamaya S, Flammini A, et al. (2022) Political audience diversity and news reliability in algorithmic ranking. *Nature Human Behaviour* 6: 495–505.

Dubois E and Blank G (2018) The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, Communication & Society* 21(5): 729–745.

Epstein R and Robertson RE (2015) The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences of the United States of America* 112(33): E4512–E4521.

Epstein R, Robertson RE, Lazer D, et al. (2017) Suppressing the Search Engine Manipulation Effect (SEME). *Proceedings of the ACM on Human-Computer Interaction* 1: 42.

Fletcher R and Nielsen RK (2018) Automated serendipity: the effect of using search engines on news repertoire balance and diversity. *Digital Journalism* 6(8): 976–989.

Fletcher R and Park S (2017) The impact of trust in the news media on online news consumption and participation. *Digital Journalism* 5(10): 1281–1299.

Fletcher R, Kalogeropoulos A and Nielsen RK (2021) More diverse, more politically varied: how social media, search engines and aggregators shape news repertoires in the United Kingdom. *New Media & Society*. Epub ahead of print 7 July. DOI: 10.1177/14614448211027393.

Garrett RK (2009) Echo chambers online? Politically motivated selective exposure among Internet news users1. *Journal of Computer-Mediated Communication* 14(2): 265–285.

Ghose A, Ipeirotis PG and Li B (2014) Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science* 60(7): 1632–1654.

Google Developers (2021) Understanding the sources behind Google News. *Google Search Central Blog*, 1 June. Available at: https://developers.google.com/search/blog/2021/06/google-news-sources

Google Developers (2022) SEO starter guide: the basics. Google Search Central. *Documentation*. Available at: https://developers.google.com/search/docs/fundamentals/seo-starter-guide

Hackett RA (1984) Decline of a paradigm? Bias and objectivity in news media studies. *Critical Studies in Mass Communication* 1(3): 229–259.

Haim M, Graefe A and Brosius H-B (2018) Burst of the filter bubble? *Digital Journalism* 6(3): 330–343.

Hamilton JT (2011) *All the News That's Fit to Sell*. Princeton, NJ: Princeton University Press.

Hannak A, Sapiezynski P, Molavi Kakhki A, et al. (2013) Measuring personalization of web search. In: *Proceedings of the 22nd international conference on world wide web—WWW '13*, pp. 527–538. Available at: https://dl.acm.org/doi/10.1145/2488388.2488435

Hummel D and Maedche A (2019) How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics* 80: 47–58.

Jachimowicz JM, Duncan S, Weber EU, et al. (2019) When and why defaults influence decisions: a meta-analysis of default effects. *Behavioural Public Policy* 3(2): 159–186.

Kammerer Y and Gerjets P (2012) Chapter 10: how search engine users evaluate and select web search results: the impact of the search engine interface on credibility assessments. In: Lewandowski D (ed.) *Web Search Engine Research*, vol.4. Bingley: Emerald Group, pp. 251–279.

Keijzer MA and Mäs M (2022) The complex link between filter bubbles and opinion polarization. *Data Science* 5(2): 139–166.

Kulshrestha J, Eslami M, Messias J, et al. (2017) Quantifying search bias: investigating sources of bias for political searches in social media. In: *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, pp. 417–432. Available at: https://arxiv.org/abs/1704.01347

Lau AYS and Coiera EW (2009) Can cognitive biases during consumer health information searches be reduced to improve decision making? *Journal of the American Medical Informatics Association* 16(1): 54–65.

Loecherbach F, Welbers K, Moeller J, et al. (2021) Is this a click towards diversity? Explaining when and why news users make diverse choices. In: *13th ACM web science conference 2021*, pp. 282–290. Available at: https://dl.acm.org/doi/10.1145/3447535.3462506

Makhortykh M, Urman A, Gil-Lopez T, et al. (2021) To track or not to track: examining perceptions of online tracking for information behavior research. *Internet Research* 32(7): 260–279.

Möller J, Van De Velde RN, Merten L, et al. (2020) Explaining online news engagement based on browsing behavior: creatures of habit? *Social Science Computer Review* 38(5): 616–632.

Montoya RM, Horton RS, Vevea JL, et al. (2017) A re-examination of the mere exposure effect: the influence of repeated exposure on recognition, familiarity, and liking. *Psychological Bulletin* 143(5): 459–498.

Mukerjee S and Yang T (2021) Choosing to avoid? A conjoint experimental study to understand selective exposure and avoidance on social media. *Political Communication* 38(3): 222–240.

Muñoz-Torres JR (2012) Truth and objectivity in journalism. *Journalism Studies* 13(4): 566–582.

Murdock BB Jr (1962) The serial position effect of free recall. *Journal of Experimental Psychology* 64: 482–488.

Murphy J, Hofacker C and Mizerski R (2006) Primacy and recency effects on clicking behavior. *Journal of Computer-Mediated Communication* 11(2): 522–535.

Nechushtai E and Lewis SC (2019) 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.

Newman N, Fletcher R, Schulz A, et al. (2021) Digital news report. *Reuters Institute for the Study of Journalism*. Available at: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital_News_Report_2021_FINAL.pdf

Nielsen RK and Ganter SA (2018) Dealing with digital intermediaries: a case study of the relations between publishers and platforms. *New Media & Society* 20(4): 1600–1617. DOI: 10.1177/1461444817701318.

Pan B, Hembrooke H, Joachims T, et al. (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: What the Internet Is Hiding from You*. New York: Penguin Books.

Park CW and Lessig VP (1981) Familiarity and its impact on consumer decision biases and heuristics. *Journal of Consumer Research* 8(2): 223–230.

Patterson TE (2013) *Informing the News: The Need for Knowledge-Based Journalism*. New York: Vintage Books.

Puschmann C (2019) Beyond the bubble: assessing the diversity of political search results. *Digital Journalism* 7(6): 824–843.

Rahwan I, Cebrian M, Obradovich N, et al. (2019) Machine behaviour. *Nature* 568(7753): 477–486.

Robertson RE, Green J, Ruck D, et al. (2021) *Engagement outweighs exposure to partisan and unreliable news within Google Search*. arXiv:2201.00074v3. Available at: https://arxiv.org/pdf/2201.00074.pdf

Robertson RE, Jiang S, Joseph K, et al. (2018) Auditing partisan audience bias within Google Search. *Proceedings of the ACM on Human-Computer Interaction* 2: 148.

Ross Arguedas A, Robertson C, Fletcher R, et al. (2022) Echo chambers, filter bubbles, and polarisation: a literature review. Reuters Institute for the Study of Journalism. Available at: https://ora.ox.ac.uk/objects/uuid:6e357e97-7b16-450a-a827-a92c93729a08

Scharkow M, Mangold F, Stier S, et al. (2020) How social network sites and other online interme-
diaries increase exposure to news. *Proceedings of the National Academy of Sciences of the
United States of America* 117(6): 2761–2763.

Sharot T and Sunstein CR (2020) How people decide what they want to know. *Nature Human
Behaviour* 4(1): 14–19.

StatCounter (2021) Search engine market share worldwide. StatCounter Global Stats. Available at:
https://gs.statcounter.com/search-engine-market-share

Steiner M, Magin M, Stark B, et al. (2022) Seek and you shall find? A content analysis on the
diversity of five search engines' results on political queries. *Information, Communication &
Society* 25: 217–241.

Stier S, Kirkizh N, Froio C, et al. (2020) Populist attitudes and selective exposure to online news:
a cross-country analysis combining web tracking and surveys. *The International Journal of
Press/Politics* 25(3): 426–446.

Stier S, Mangold F, Scharkow M, et al. (2022) Post post-broadcast democracy? News exposure in
the age of online intermediaries. *American Political Science Review* 116(2): 768–774.

Stroud NJ (2017) Selective exposure theories. In: Kenski K and Jamieson KH (eds) *The Oxford
Handbook of Political Communication*. Oxford: Oxford University Press, pp. 531–547.

Sullivan J (2019) The primacy effect in impression formation: some replications and extensions.
Social Psychological and Personality Science 10(4): 432–439.

Trielli D and Diakopoulos N (2019) Search as news curator: the role of Google in shaping attention
to news information. In: *Proceedings of the 2019 CHI conference on human factors in com-
puting systems*, pp. 1–15. Available at: https://dl.acm.org/doi/abs/10.1145/3290605.3300683

Ulloa R, Makhortykh M and Urman A (2022a) Scaling up search engine audits: practical insights
for algorithm auditing. *Journal of Information Science*. Epub ahead of print 2 May. DOI:
10.1177/01655515221093029.

Ulloa R, Richter AC, Makhortykh M, et al. (2022b) Representativeness and face-ism: gen-
der bias in image search. *New Media & Society*. Epub ahead of print 19 June. DOI:
10.1177/14614448221100699.

Urman A and Makhortykh M (2021) *You are how (and where) you search? Comparative analysis
of web search behaviour using web tracking data*. arXiv:2105.04961. Available at: http://
arxiv.org/abs/2105.04961

Urman A, Makhortykh M and Ulloa R (2021) Auditing source diversity bias in video search results
using virtual agents. In: *Companion proceedings of the web conference 2021*, pp. 232–236.
Available at: https://dl.acm.org/doi/10.1145/3442442.3452306

Wagner C, Strohmaier M, Olteanu A, et al. (2021) Measuring algorithmically infused societies.
Nature 595(7866): 197–204.

Webber W, Moffat A and Zobel J (2010) A similarity measure for indefinite rankings. *ACM
Transactions on Information Systems* 28(4): 1–38.

Wojcieszak M, De Leeuw S, Casas A, et al. (2021) Null effects of news exposure: a causal test
of the (un)desirable effects of a "news vacation" and "news binging." Research Square.
Available at: https://www.researcher-app.com/paper/9140076

Wojcieszak M, Menchen-Trevino E, Goncalves JFF, et al. (2022) Avenues to news and diverse
news exposure online: comparing direct navigation, social media, news aggregators, search
queries, and article hyperlinks. *The International Journal of Press/Politics* 27(4): 860–886.

Yin RK (2009) *Case Study Research: Design and Methods*. Thousand Oaks, CA: SAGE.

Zweig K (2017) Watching the watchers: Epstein and Robertson's „Search Engine Manipulation
Effect". *AlgorithmWatch*, 7 April. Available at: https://algorithmwatch.org/en/watching-the-
watchers-epstein-and-robertsons-search-engine-manipulation-effect/

Author biographies

Roberto Ulloa is a postdoctoral researcher at GESIS – Leibniz Institute for the Social Sciences and at the University of Konstanz. His interests include the role of institutions in shaping societies, and online platforms as forms of digital institutions.

Celina Sylwia Kacperski is a postdoctoral researcher and behavioral scientist at the University of Konstanz and Seeburg Castle University. Her research interests lie at the intersection of psychology and technological innovation, with a strong focus on incentivizing sustainable behaviors.