Disentangling Web Search on Debated Topics
A User-Centered Exploration

Rieger, Alisa; Kulane, Suleiman; Gadiraju, Ujwal; Pera, Maria Soledad

DOI
10.1145/3627043.3659559

Publication date
2024

Document Version
Final published version

Published in
UMAP 2024 - Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization

Citation (APA)
Rieger, A., Kulane, S., Gadiraju, U., & Pera, M. S. (2024). Disentangling Web Search on Debated Topics: A User-Centered Exploration. In *UMAP 2024 - Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization* (pp. 24-35). (UMAP 2024 - Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization). Association for Computing Machinery (ACM). https://doi.org/10.1145/3627043.3659559

Disentangling Web Search on Debated Topics: A User-Centered Exploration

Alisa Rieger
a.rieger@tudelft.nl
Delft University of Technology
Netherlands

Suleiman Kulane
S.A.N.Kulane@student.tudelft.nl
Delft University of Technology
Netherlands

Ujwal Gadiraju
u.k.gadiraju@tudelft.nl
Delft University of Technology
Netherlands

Maria Soledad Pera
m.s.pera@tudelft.nl
Delft University of Technology
Netherlands

ABSTRACT

When using web search engines to conduct inquiries on debated topics, searchers' interactions with search results are commonly affected by a combination of searcher and system biases. While prior work has mainly investigated these biases in isolation, there is a lack of a comprehensive understanding of web search on debated topics. Addressing this gap, we conducted an exploratory user study ($N = 255$), aimed at advancing the understanding of the intricate *searcher-system interplay.* Particularly, we investigated the relations between (i) search system exposure, searchers' attitude strength, prior knowledge, and receptiveness to opposing views, (ii) search interactions, and (iii) post-search epistemic states. We observed that search interaction was shaped by search system exposure, attitude strength, and prior knowledge, and that attitude change was influenced by the level of confirmation bias and initial attitude strength, but not search system exposure. Insights from this work suggest the need to adapt interventions that mitigate the risks of searcher and system bias to searchers' nuanced pre-search epistemic states. They further emphasize the threat of customizing the search ranking to enhance user satisfaction in the context of debated topics to responsible opinion formation.

CCS CONCEPTS

• **Information systems** → **Web searching and information discovery**; **Users and interactive retrieval**; • **Human-centered computing** → **User studies**.

ACM Reference Format:
Alisa Rieger, Suleiman Kulane, Ujwal Gadiraju, and Maria Soledad Pera. 2024. Disentangling Web Search on Debated Topics: A User-Centered Exploration. In *Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization (UMAP '24), July 01–04, 2024, Cagliari, Italy*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3627043.3659559

UMAP '24, July 01–04, 2024, Cagliari, Italy

ACM ISBN 979-8-4007-0433-8/24/07
https://doi.org/10.1145/3627043.3659559

1 BACKGROUND AND INTRODUCTION

Debated topics are subjects of ongoing socio-scientific discussions among individuals with differing perspectives, often lacking a single, straightforward solution [62]. Some can be extremely one-sided and be supported by scientific evidence, e.g., '*human activities are the primary drivers of climate change*'. Others are known to prompt more discussion, given the existence of reasonable arguments on both sides, e.g., ' *recreational drugs should be legalized*'. Engaging with information on such topics can ultimately impact opinion formation and decision-making processes, and thus be consequential for individuals and society. Due to the complexities of such topics, interacting with related information can be demanding on individuals and trigger emotionally charged and biased behavior, impeding accurate and well-rounded knowledge gain [30, 35, 48, 52, 53].

When seeking information on debated topics, individuals often resort to web search engines [15, 24]. Yet, these information retrieval systems are not explicitly designed to handle queries that call for resources encompassing diverse viewpoints, as is the case with debated topics [57, 64, 74]. Search engines are not neutral but act as algorithmic curators that have been found to absorb, amplify, and reflect biases [9, 46, 69, 75]. In the context of debated topics, this can lead to viewpoint-biased exposure, i.e., Search Engine Result Pages (**SERPs**) on which one viewpoint is over-represented [19]. Searchers also add to the challenges, as cognitive biases can manifest in individuals' search interactions, for instance when they favor attitude-confirming over attitude-opposing information (*confirmation bias*) or resort to high-ranked search results (*position bias*) [8]. Researchers have observed an interplay of exposure and interaction biases, which can lead searchers to adopt the viewpoint expressed by the majority of the highly-ranked SERP results [6, 22, 49].Yet, searchers have learned to rely on search engines as neutral and unbiased providers of information, remaining unaware of the impact that exposure and interaction biases exert on their search experience, opinions, and decisions [27, 54, 66].

The body of work dedicated to understanding the impact of exposure and interaction biases during web search on debated topics, along with the strategies to overcome the associated challenges is growing. Research, however, has predominantly investigated specific aspects of this larger issue in isolation. For instance, recent works studied relations between exposure and interaction [16, 58, 80], user traits and interaction [55, 59], and exposure, interaction and attitude change [20, 38, 65], frequently focusing specifically on individuals with either strong or weak pre-search attitudes. Across these works, a recurring point of discussion is the difficulty of understanding the intricate searcher-system interplay that is characterized by entangled effects of system exposure, search interactions, and user characteristics. We argue that comprehension of the complex dynamics of web search on debated topics requires undertaking a holistic approach with the user placed at the center of this exploration, as depicted in Figure 1.

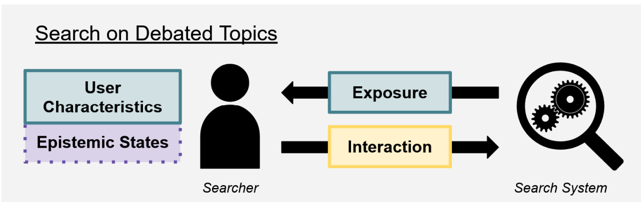

Figure 1: A *user-centered perspective* of search on debated topics. Searchers with varied characteristics are *exposed* to resources on debated topics by a search system. They choose the results to *interact* with, modifying their epistemic states.

To advance knowledge and uncover relations between the different facets of web search on debated topics, we adopt an exploratory, open-ended approach. We study the relations between factors inherent to the *searcher and search system* (user characteristics, exposure bias), *search interaction* (confirmation bias, position bias, search effort), and *post-search epistemic states* (attitude change, knowledge gain), as illustrated in Figure 2. This is inspired by outcomes inferred from research on belief dynamics during web search for information on medical yes-no questions with a known true answer (expert consensus) [76, 78]. In this context, White [76] observed that pre-search beliefs affected interaction and that exposure bias could shift post-search beliefs, but only if participants did not have strong pre-search beliefs. However, the investigated yes-no questions are considerably different from debated topics that involve multiple viewpoints and have no defined correct answer. In addition to exposure bias and beliefs, users' prior knowledge and personality traits might play a role in shaping search interaction and post-search epistemic states [59]. These are recognized factors influencing general web search behavior [5, 34, 41], yet, their combined impact in the context of debated topics remains to be explored.

In this study, we investigate the following research questions (RQs):

- **RQ1** How do attributes of the searcher and search system shape search interaction?
- **RQ2** How do attributes of the searcher and search system shape the post-search epistemic states of the searchers?
- **RQ3** How do search interactions shape the post-search epistemic states of the searchers?

To address these RQs, we conducted a user study with 255 participants and asked them to imagine a scenario where they would use a search system to find arguments on a debated topic. Before the search task, we captured user characteristics, i.e., participants' epistemic states and the relatively stable trait of receptiveness to opposing views [45]. During the search task, we exposed participants to a mock SERP with results portraying different viewpoints (i.e., *opposing*, *supporting*, *neutral*) towards the topic statement. We logged their interactions to then compute metrics that approximate their confirmation bias, position bias, and search effort. Using a between-subjects experimental design, we manipulated exposure by assigning participants to three SERP ranking bias conditions: balanced, biased supporting, and biased opposing. We captured their post-search epistemic states(attitude change and knowledge gain).

The findings of this empirical study exploring relations among attributes of the searcher and search system, search interaction, and user post-search epistemic states in the context of web search on debated topics indicate that search interaction was shaped by search system exposure, attitude strength, and prior knowledge. Attitude change was not directly affected by search system exposure but by participants' confirmation bias and initial attitude strength. We observed various moderating effects, which suggest that effects of exposure, interaction, and their interplay are moderated by prior knowledge, attitude strength, and potentially other pre-search epistemic states. These observations can serve as a foundation for research aiming to identify strategies to overcome the risks related to web search on debated topics, particularly for adapting such strategies to different searchers' needs. The dataset, containing behavioral data from search logs and measures of knowledge, attitude, receptiveness to opposing views from questionnaire responses, as well as the material used for data collection are publicly available.[1]

2 RELATED WORK

Engaging with Debated Topics. Web search engines are used to fulfill wide-ranging information needs, including complex inquiries that require exploration, such as researching information about debated topics [7, 15, 24, 79]. Engaging with information on debated topics can cause changes in individuals' opinions and knowledge, thus their *epistemic states* [22, 38]. *Epistemics* is an interdisciplinary field focusing on information processing, knowledge behavior, and belief formation which incorporates insights from epistemology, cognitive science, and information science, among other disciplines [36]. Throughout this paper, we refer to user states that are related to knowledge and opinions as user or searcher *epistemic states*. Opinions on debated topics can lead to practical decisions, e.g., whether to adopt vegetarianism, equally split parental leave, or vote for a certain party. Given the substantial implications of such decisions for individuals and society, searchers would ideally engage with diverse viewpoints to become informed on the topic, enabling them to form opinions *responsibly* [39, 47, 57, 59].

Interaction and Exposure Biases. For the individual, interacting with information on debated topics can be challenging and affected by cognitive biases, shaping their interactions with the search system and impeding responsible opinion formation [8, 38, 57]. E.g., when searching about debated topics, users may prioritize protecting and defending their own beliefs and values over the pursuit of becoming informed by gaining knowledge about diverse viewpoints [30, 32, 35, 52], thus interacting preferably with information that aligns with preexisting beliefs (*confirmation bias*) [8, 73, 75]. Further, searchers have learned to rely on search engines and mostly

[1] https://osf.io/u3s5n/?view_only=86cb2495551943bd87576d9790aef3dd

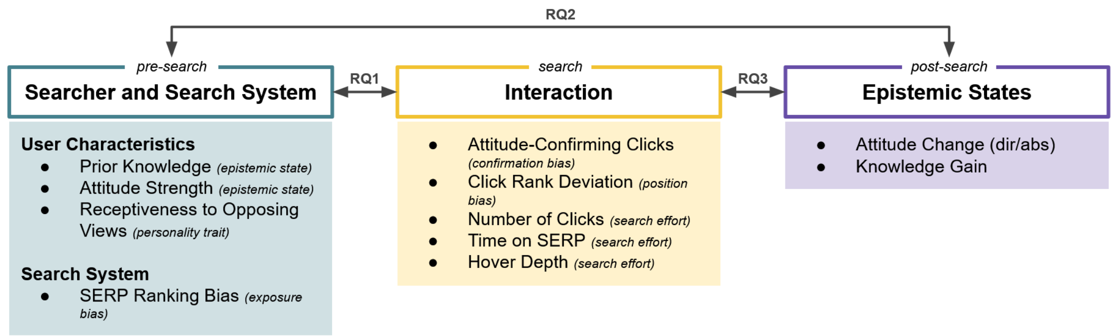

Figure 2: Exploration Overview. To gain a better understanding of web search on debated topics, we investigated RQ1, RQ2, and RQ3, exploring relations between selected aspects of the searcher and search system, the search interaction, and the post-search epistemic states.

interact with highly ranked search results (*position bias*) [34] trusting the search engine to provide relevant, unbiased, and credible information [27, 29, 54, 66]. Yet, recent research by Draws et al. [19] has observed viewpoint-biased rankings in response to queries on debated topics across different search engines, characterized by an over-representation of specific viewpoints in highly ranked search results, even in response to viewpoint-neutral queries. Search engines might further amplify searchers' interaction biases by tailoring the search result ranking to maximize individual relevance, for instance, based on click rates [9, 14, 27, 67]. Such biases commonly remain unnoticed by searchers who, due to the opacity of automatic filtering and ranking processes, face difficulties in determining whether the provided search results are unbiased and complete [44, 66]. Conversely, searchers were even observed to adopt the prevailing viewpoint when exposed to SERPs with a viewpoint-biased ranking—a phenomenon known as the *Search Engine Manipulation Effect* **(SEME)** [22]. For search on debate topics, however, Draws et al. [20] did not find evidence for an effect of exposure bias on attitude change, but instead found that it was linked to search interactions. For search on political topics and news, studies indicate that although exposure plays a considerable role in shaping interactions, individuals still tailor their interactions to align with prior beliefs when faced with belief inconsistent exposure bias [60, 65].

User Characteristics. The interaction with information on debated topics can vary based on user characteristics, such as their epistemic states and more stable user traits. For instance, the manifestation of confirmation bias in an individual's information behavior was found to be influenced by their attitude strength, where those with weaker attitudes seem to be more likely to engage with and open to processing attitude-opposing information [30, 37, 72]. Further, prior research has observed variance in web search behavior depending on users' topic knowledge [23, 41, 77, 81]. Searchers with high topic knowledge tend to employ more efficient search strategies [77], demonstrating a reduced susceptibility to position bias by being more likely to click on lower-ranked items on the SERP [41] and select items based on topic relevance and source credibility [62]. In the context of search on debated topics, there are only preliminary explorations on the role of prior knowledge [59]; more conclusive insights are pending.

In addition to epistemic states, relatively stable user traits are known to shape interactions with information on debated topics [13, 28, 40, 51]. Particularly relevant in the context of debated topics and responsible opinion formation is the trait of *receptiveness to opposing views*–defined as the *willingness to access, consider, and evaluate opposing views in an impartial manner* [45].

3 METHOD

We conducted and exploratory between-subjects study to investigate our research questions. All related material can be found at the URL in Footnote 1.

3.1 Experimental Setup

To probe the dynamics of web search on debated topics, we selected two topics from *ProCon* [1], a resource that presents controversial topics and related arguments, with varying levels of controversy: '*Should abortion be legal?*' (highly controversial, people tend to have strong attitudes) and '*Is obesity a disease?*' (moderately controversial, people tend to have moderate attitudes). We created a custom SERP on which we displayed ten ranked, pre-selected, viewpoint-annotated search results on a given topic. As in conventional SERPs, we displayed the title and snippet for each SERP result. By clicking on a result searchers could access the corresponding linked webpage.

To manipulate the exposure bias, we assigned participants to one of three SERP ranking bias conditions, in which the results were ranked adhering to a viewpoint ranking template: *balanced*, *biased supporting*, and *biased opposing*. In the *balanced* condition, participants were exposed to alternating attitude-confirming and -opposing results. We randomly varied whether an opposing or confirming result would be displayed on the first rank. In the *biased* ranking conditions, the first six were either attitude-confirming (biased supporting) or attitude-opposing (biased opposing) search results, followed by two neutral and two attitude-opposing or attitude-confirming ones, respectively. By including two neutral search results we could derive additional insights into the interplay of exposure and interaction bias by assessing whether participants would invest the extra effort to interact with the two lowest-ranked search results to learn about the underrepresented viewpoint or engage with search results that confirm their opinion.

We selected the ten SERP results per condition from a set of viewpoint-annotated search results that we prepared for each topic. For this, we obtained 30 search results per topic from the *Bing API* [43] which had to fulfill our inclusion criteria (no paywall, content focuses clearly on the topic). We collected annotations (on a 7-point Likert scale, ranging from *strongly opposes* to *strongly supports* the topic statement) from crowd workers recruited via *Prolific* [2]. Each crowd worker annotated ten search results, and each search result was annotated by three crowd workers. To control for quality, we included two attention checks and discarded annotations from crowd workers who failed at least one. Ultimately, we assigned the median value of the three annotations to the search results.

To measure prior knowledge and knowledge gain, we developed knowledge questionnaires. Per topic, we compiled a list of 60-80 statements (technically referred to as items) based on details extracted from the *Wikipedia* pages on the topic (e.g., *Approximately 45% of abortions conducted globally are considered unsafe*, *According to the American Medical Association, obesity is a disease*). We recruited 20 distinct crowd workers per topic to judge these items and respond with *true*, *false*, or *I don't know* labels. We discarded items that proved to be too easy (resulting in all correct responses) or too difficult/ambiguous (resulting in all incorrect responses). We then created the knowledge questionnaires by selecting a subset of items per topic that could capture participants' knowledge reliably [23]. To maximize the internal reliability of the knowledge questionnaires we randomly selected 4,000 samples of 15 items from the remaining item pools per topic and computed Cronbach's α for each sample. We identified a set of 15 items per topic with Cronbach's α exceeding 0.8. The knowledge questionnaires used in our study are publicly available at the link in Footnote 1.

To measure participants' prior **attitude** and attitude change, we asked them to report their agreement with a statement on the assigned topic (i.e., '*obesity is a disease*', '*abortion should be legal*') on a 7-point Likert scale ranging from *strongly disagree (-3)* to *strongly agree (+3)* before and after the search task, adopting the approach used in prior research on attitude change in web search (e.g., [20, 22, 59]). We measured participants' **receptiveness to opposing views** with the 18-item self-report measure developed by Minson et al. [45] which captures *negative emotional reactions toward disagreement*, *intellectual curiosity regarding opposing views*, *derogation of those holding opposing views*, and *belief that it is inappropriate to debate certain issues*.

3.2 Procedure

We recruited participants via *Prolific* [2] and paid them 2.1£ (median = 8.02£/h) for their participation. They had to be at least 18 years old and proficient in English. The questionnaire responses were collected using *Qualtrics* [3]. We integrated five attention checks across the pre and post-search questionnaires, in which we instructed participants on which response to select. All data was collected in January 2023 with the following procedure, approved by our institution's ethics committee:

- **Pre-Search:** Given the potential sensitivity of the debated topics (i.e., abortion, obesity), we named them in the opening statement to allow prospective participants to make an informed decision regarding their participation. After receiving participants' informed consent to join the study, we randomly assigned them to one of the two debated topics and asked them to state their attitude on the topic (**attitude strength**). We then assigned them to one of the three **SERP ranking bias** conditions, balancing the distribution of participants with different attitude strengths across conditions. Subsequently, we asked them to fill out the knowledge questionnaire (**prior knowledge**) on the assigned topic and measured their **receptiveness to opposing views**.
- **Search:** For the search task, we asked participants to envision a situation in which they prepared for a mock debate with colleagues by making use of the search engine to find arguments related to the assigned topic. They could access the custom search page and click on search results to retrieve the linked documents, with no time limit. We logged participants' **search interactions** with the *LogUI* framework [42]. Specifically, we logged click events, the ranks and viewpoints of SERP results interacted with, and the time spent on the SERP.
- **Post-Search:** Following the search task, participants could report the identified arguments in a free-text field to simulate the completion of the search task. Further, we asked them to complete the knowledge questionnaire (**knowledge gain**) and report their attitude (**attitude change**) again.

3.3 Variables and Analysis

In Table 1, we describe the variables used in our study to model relations among searcher and search system, search interaction, and post-search epistemic states. For context on study participants, we captured their age and gender. To explore the relations between the searcher and search system, search interaction, and post-search epistemic states, we computed the Pearson correlation matrix between all independent and dependent variables that we considered. To explore searcher and search system-dependent group differences through descriptive statistics and ANOVAs, we grouped participants according to their attitude strength (i.e., weak, moderate, strong), the SERP ranking bias condition (i.e., balanced, biased supporting, biased opposing) and categorized them into three levels (i.e., low, moderate, high) of prior knowledge and receptiveness to opposing views, based on the quartiles of the distribution of the respective variable across all participants. We investigate differences between these groups in search interaction (five ANOVAs: attitude-confirming clicks, click rank deviation, number of clicks, time on SERP, hover depth) and in post-search epistemic states (three ANOVAs: directional attitude change, absolute attitude change, knowledge gain). Due to the exploratory and open-ended nature of this study, we did not set a significance threshold but present effect sizes and p-values of the ANOVA results as indicators of meaningful relationships that warrant further investigation with confirmatory studies in the future. A table providing an overview of all ANOVA results can be found at the link in Footnote 1.

Table 1: Study Variables. Name, values, and description of the variables that we manipulated or measured. For searcher and search system variables we used two representations, to compute correlations (numerical) and investigate group differences (categorical).

	Name	Values	Description
Pre-Search Searchers and Search System	SERP Ranking Bias	balanced (0), biased supporting (1), biased opposing (-1)	Randomly assigned to each participant
	Prior Knowledge	0 to 1 (correlations)	The Proportion of correctly answered questions of the knowledge questionnaire
		low, moderate, high (group differences)	Derived from the distribution of knowledge scores across all participants, those in the lowest quartile are categorized as having low knowledge; those in the highest quartile, high knowledge
	Attitude Strength	undecided (0), weak (1), moderate (2), strong (3)	Reporting to neither agree nor disagree with a topic statement was considered as an *undecided*, to somewhat agree or disagree as a *weak*, to agree or disagree as a *moderate*, and to strongly agree or disagree as *strong* attitude
	Receptiveness to Opposing Views	-1 to 1 (correlations)	Higher values indicate a higher receptiveness
		low, moderate, high (group differences)	Derived from the distribution of receptiveness to opposing views scores across all participants, those in the lowest quartile are categorized as having a low receptiveness, and the highest quartile as having a high receptiveness to opposing views
Search Interaction	Attitude-Confirming Clicks	0 to 1	The proportion of attitude-confirming results among the search results participants clicked on *(only for participants who clicked on one or more search results)*
	Click Rank Deviation	0 to 1	Deviation of the mean rank clicked from the mean rank if the participant would have clicked the top-ranked search results, normalized by the number of clicks *(only for participants who clicked on one or more search results)*; A value of 1 indicates maximal deviation (the participant clicked on the lowest-ranked search results), a value of 0 indicates no deviation (the participant clicked the top-ranked search results)
	Number of Clicks	0 to 10	The number of distinct search results a participant clicked on
	Time on SERP	in seconds	Amount of time that a participant spent on the search task in seconds
	Hover Depth	0 to 10	Lowest ranked search result a participant hovered on
Post-Search Epistemic States	Attitude Change	-6 to 2 (Directional)	The difference between the pre- and the post-search attitude; Negative values indicate an attitude weakening or change to the opposing attitude, while positive values indicate an attitude strengthening. Since directional attitude change could only be calculated for participants who did not report to be *undecided*, the values could range between -6 (change from *strongly agree* to *strongly disagree* or vice versa) and +2 (change from *somewhat agree* to *strongly agree* or *somewhat disagree* to *strongly disagree*)
		0 to 6 (Absolute)	The absolute difference between the pre- and the post-search attitude
	Knowledge Gain	-1 to 1	The difference between pre and the post-search knowledge; Negative values indicate a loss and positive values a gain of knowledge

4 RESULTS

We collected data from 280 participants, of which 25 were excluded from the analysis since they failed at least one attention check. Of the 255 remaining participants, 44.3% reported to be female, 54.1% male, and 1.5% preferred not to share their gender. 49% reported to be aged between 18 and 25, 32.5% between 26 and 35, 11% between 36 and 45, 5.9% between 46 and 55, 0.8% between 56 and 65, and 0.8% more than 65 years old.

We aimed for equal distribution of participants across topics and SERP ranking bias conditions: 124 participants were assigned to '*should abortion be legal*' and 131 to '*is obesity a disease*', and 83 to 86 participants were assigned to each of the SERP ranking bias conditions. Amongst the 255 participants, we observed a mean (M) *attitude strength* of 1.8 with a standard error (SE) of 0.06, a mean *prior knowledge* of 0.54 ($SE = 0.01$), and a mean *receptiveness to opposing views* of 0.06 ($SE = 0.02$) with no differences in all three variables between the three SERP ranking bias conditions. Participants, on average, clicked on 2.76 ($SE = 0.16$) search results and spent 7 min 1 sec ($SE = 29.8sec$) on the search page.

Due to varying search interactions, dependent variables were not always applicable across all participants (e.g., confirmation bias was not applicable for participants who did not click on any search result). Thus, we considered different subsets of participants depending on the dependent variable of interest, detailed in Table 2. Similarly, the direction of SERP ranking bias (supporting, opposing) cannot be defined for participants who reported to be *undecided* when asked for their attitude ($n = 19$). However, to advance our understanding of SEME, we explored the attitude change of undecided participants who were exposed to a biased SERP ($n = 13$, see Table 5).

4.1 Search Interaction

Attitude-Confirming Clicks. We observed weak to moderate correlations between the proportion of attitude-confirming clicks and SERP ranking bias ($r = 0.49, p < .001$), prior knowledge ($r = -0.16, p = .033$), and attitude strength ($r = 0.27, p < .001$) (Table 3). Exploring group differences with an ANOVA, we found variations depending on the SERP ranking bias; $F(2, 107) = 33.74, p < .001, f = 0.79$, with participants exposed to a SERP with an attitude-supporting ranking bias clicking on a higher

Table 2: Subsets of participants considered in the data analysis per dependent variable, since they can only be calculated for participants who fulfilled the requirements. Participants who reported *undecided* attitudes ($n = 19$) were excluded from this analysis due to the inability to determine the direction of *SERP ranking bias* for this group.

Dependent Variables	Requirements	n
Attitude-Confirming Clicks, Click Rank Deviation	Number of clicks > 0, Attitude Strength > 0	176
Attitude Change, Number of Clicks, Time on SERP, Hover Depth, Knowledge Gain	Attitude Strength > 0	236

Table 3: Correlation matrix off all aspects we captured of the searcher and search system,search interaction,and post-search epistemic states, based on participants who had complete data rows ($n = 176$). Positive correlations exceeding 0.1 are colored blue; negative ones below -0.1 red. Color shades indicate the correlation strength with light shades for weak (0.1 to 0.24) and dark shades for moderate (0.25 to 0.49) correlations. Coefficients are bolded if $p < .05$.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. SERP ranking bias	1.00											
2. Prior Knowledge	0.02	1.00										
3. Attitude Strength	0.09	0.02	1.00									
4. Receptiveness opp. Views	0.04	-0.09	-0.04	1.00								
5. Attitude-Confirming Clicks	**0.49**	**-0.16**	**0.27**	0.12	1.00							
6. Click Rank Deviation	**-0.37**	0.06	-0.03	0.02	**-0.19**	1.00						
7. Number of Clicks	-0.02	0.14	-0.03	-0.02	-0.14	0.07	1.00					
8. Time on SERP	-0.02	0.03	-0.06	0.02	0.04	-0.04	0.09	1.00				
9. Hover Depth	**-0.30**	0.18	-0.08	-0.02	**-0.22**	**0.39**	**0.26**	**0.10**	1.00			
10. Attitude Change (dir)	-0.06	**0.16**	-0.03	0.01	**0.30**	0.07	0.08	0.00	0.08	1.00		
11. Attitude Change (abs)	-0.02	0.05	**-0.38**	-0.01	**-0.30**	0.02	0.08	0.05	0.02	**-0.41**	1.00	
12. Knowledge Gain	0.02	**-0.26**	-0.07	0.00	0.12	-0.07	0.04	0.07	-0.07	-0.04	0.08	1.00

proportion of attitude-confirming results ($M = 0.72, SE = 0.03$) than participants exposed to a SERP with a balanced ranking ($M = 0.6, SE = 0.02$) or an attitude-opposing ranking bias ($M = 0.28, SE = 0.03$). Further, we observed differences depending on the attitude strength; $F(2, 107) = 6.97, p = .001, f = 0.36$, with those who reported having a strong attitude prior to the search clicking on a higher proportion of attitude-confirming results ($M = 0.66, SE = 0.03$) than those who reported having a moderate ($M = 0.54, SE = 0.03$) and weak attitude ($M = 0.42, SE = 0.03$) (see Figure 3, A). The ANOVA moreover revealed variations depending on the level of prior knowledge; $F(2, 107) = 4.11, p = .019, f = 0.28$, with participants with low prior knowledge clicking on a higher proportion of attitude-confirming search results ($M = 0.66, SE = 0.03$) than participants with moderate ($M = 0.51, SE = 0.03$) and high ($M = 0.46, SE = 0.03$) prior knowledge (see Figure 3, B). Lastly, the results indicate an interaction effect between attitude strength and ranking bias; $F(4, 107) = 3.27, p = .014, f = 0.35$.

Click Rank Deviation. As captured in Table 3, we saw a moderate negative correlation between the SERP ranking bias and click rank deviation ($r = -0.37, p < .001$). This difference between SERP ranking bias conditions was supported by a between groups ANOVA; $F(2, 107) = 15.22, p < .001, f = 0.53$), indicating a higher click rank deviation (i.e., participants clicked on lower-ranked search results) for those exposed to a SERP with an attitude-opposing ranking bias ($M = 0.54, SE = 0.03$) than for those exposed to a balanced ranking ($M = 0.32, SE = 0.02$) or an attitude-supporting ranking bias ($M = 0.27, SE = 0.02$). The ANOVA also revealed a SERP ranking bias and attitude strength interaction; $F(4, 107) = 3.26, p = .014, f = 0.35$, where the click rank deviation was lower for participants with strong attitudes compared to those with weak attitudes ($\Delta = -0.24$) if they were exposed to a SERP with an attitude-supporting ranking bias, but higher ($\Delta = +0.14$) if exposed to a SERP with an attitude-opposing ranking bias (see Figure 3, C).

Number of Clicks, Time on SERP, Lowest Rank Hovered. We saw a moderate negative correlation between SERP ranking bias and hover depth ($r = -0.3, p = .008$) (see Table 3). Our data did not reveal any correlations between attributes of the searcher and search system and the number of clicks and time spent on the SERP. When exploring group differences designated by attributes of the searcher and search system with ANOVAs, we found that participants' number of clicks varied depending on interactions of the SERP ranking bias with participants' attitude strength; $F(4, 163) = 2.73, p = .031, f = 0.26$, as well as their knowledge; $F(4, 163) = 3.06, p = .018, f = 0.27$. The lowest ranked result that participants hovered on varied with the SERP ranking bias; $F(4, 161) = 4.65, p = .011, f = 0.24$. Further, we saw that this effect of the SERP ranking bias was moderated by participants' attitude strength; $F(4, 161) = 2.78, p = .028, f = 0.26$.

4.2 Post-Search Epistemic States

Attitude Change. We found a weak positive correlation between participants' prior knowledge and their directional attitude change ($r = 0.16, p = .045$) and a moderate negative correlation between attitude strength and absolute attitude change ($r = -0.38, p < .001$) (Table 3). We also noticed moderate correlations between participants' attitude-confirming clicks and their directional ($r = 0.3, p < .001$) and absolute ($r = -0.3, p < .001$) attitude change.

The ANOVA exploring group differences revealed a second-order interaction of SERP ranking bias, participants' knowledge, and their receptiveness to opposing views for directional attitude change; $F(7, 163) = 2.55, p = .016, f = 0.33$. When considering *absolute* attitude change, an ANOVA showed group differences depending on participants' attitude strength; $F(2, 163) = 16.81, p < .001, f = 0.45$: those with weak ($M = 0.79, SE = 0.05$) and moderate attitudes ($M = 0.65, SE = 0.07$) were more likely to change their attitudes, whereas those with strong attitudes were highly unlikely to change their attitudes ($M = 0.08, SE = 0.03$) (Figure 3, E).

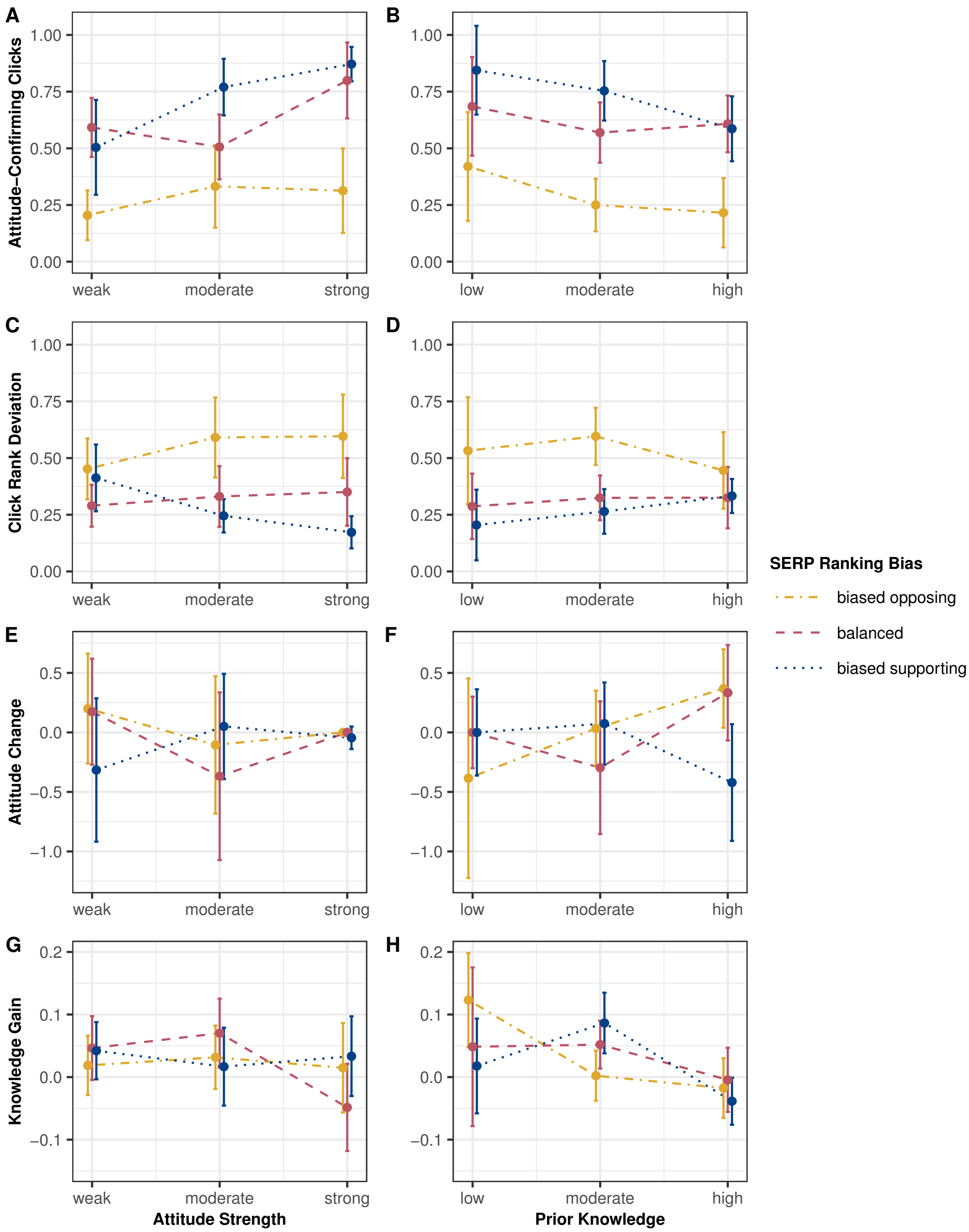

Figure 3: Mean proportion of attitude-confirming clicks (A, B), mean click rank deviation (C, D), mean attitude change (dir) (E, F), and mean knowledge gain (G, H) per SERP ranking bias condition for different levels of attitude strength and prior knowledge with 95% confidence intervals. Note that we use line plots to illustrate interaction effects.

Table 4: Attitude change of opinionated participants ($n = 236$)

Attitude Change	Proportion per SERP ranking bias		
	supporting	opposing	balanced
changed to opposing	0.12	0.04	0.06
renounced	0.03	0.07	0.04
strengthened	0.17	0.22	0.2
weakened	0.04	0.06	0.05
unchanged	0.64	0.62	0.65

Table 5: Attitude change of undecided participants in biased SERP ranking conditions ($n = 13$)

Attitude Change	Proportion
still undecided	0.38
adopted majority viewpoint	0.31
adopted minority viewpoint	0.31

For absolute attitude change, an interaction of the SERP ranking bias and participants' knowledge was found; $F(4, 163) = 2.75, p = .03, f = 0.26$ (Figure 3, F).

We investigated the proportions of participants who adopted the majority viewpoint amongst the high-ranked SERP results. We did not find evidence for SEME since the proportion of participants who changed their attitude towards the opposing attitude was lower for those exposed to a search page with attitude-opposing (4%) than for attitude-supporting ranking bias (12%) and balanced ranking (6%) (see Table 4). Additionally, we explored attitude change for participants who reported being undecided before the search task and were exposed to a viewpoint-biased SERP ($n = 13$). In this group, the same proportion of participants (31%) adopted the majority viewpoint as the minority viewpoint (see Table 5).

Knowledge Gain. We found a moderate negative correlation between participants' level of prior knowledge and their knowledge gain ($r = -0.26, p < .001$). An ANOVA alludes to differences between participants with varied levels of both prior knowledge; $F(2, 163) = 9.69, p < .001, f = 0.34$, and attitude strength; $F(2, 163) = 3.28, p = .04, f = 0.20$. Participants with low prior knowledge gained knowledge ($M = 0.07, SE = 0.009$), those with high prior knowledge did not ($M = -0.02, SE = 0.005$) (see Figure 3, H). Further, participants with strong prior attitudes were less likely to gain knowledge ($M = 0.0, SE = 0.008$) than those with moderate ($M = 0.04, SE = 0.008$) and weak attitudes ($M = 0.03, SE = 0.008$) (see Figure 3, G).

5 DISCUSSION

Reflecting on this exploratory user study, we discuss and contextualize key findings and their implications, as well as acknowledge limitations and outline avenues for future work.

Confirmation and Position Bias. Our data revealed that confirmation bias was strongest in participants with low knowledge or strong attitudes. Both confirmation and position bias were affected by the SERP ranking, where participants exposed to an attitude-supporting ranking bias displayed higher confirmation and position bias than those exposed to attitude-opposing ranking bias. Participants in the attitude-opposing ranking bias condition tended to click on lower-ranked results, and the proportion of clicks on attitude-confirming results, while diminished, was still substantial. This suggests that participants deliberately sought not only results that were not attitude-opposing (i.e., neutral results) but more so, attitude-confirming results within the lowest ranks of the SERP. In contrast, those exposed to attitude-supporting ranking bias engaged less with low-ranked results. From this, we infer that the desire to find information confirming prior attitudes (confirmation bias), rather than to explore diverse perspectives, caused this behavior in participants in the attitude-opposing ranking bias condition. This aligns with previous findings on search interactions being not only shaped by algorithms, determining exposure through selection and ranking but also by users, tailoring interaction to maintain prior beliefs [60, 65]. Yet, Robertson et al. [60] state that this does not imply that exposure biases are less concerning than previously suggested, but that they might cause more indirect effects, persistently over time. Furthermore, our results point towards a potential risk arising from attitude-supporting as opposed to attitude-opposing exposure bias; while attitude-supporting ranking bias does not manipulate individuals into changing their attitudes (i.e., SEME), it amplifies their interaction biases, thus hindering responsible opinion formation.

Answering **RQ1**, we found that searcher and system attributes significantly shape search interaction, with confirmation bias being strongest in participants with low knowledge or strong attitudes, as well as those exposed to an attitude-confirming ranking bias, and position bias being weakest for participants who were exposed to an attitude-opposing ranking bias. Although exposure bias considerably influences search behavior, confirmation bias drives user interaction to somewhat diminish the impact of attitude-opposing, but not attitude-supporting, exposure bias.

Attitude Change & Knowledge Gain. A salient trend regarding attitude change showed that participants with strong prior attitudes were highly unlikely to change their attitude, regardless of viewpoint exposure biases or prior knowledge levels. SERP ranking bias, on the other hand, does not seem to affect attitude change. When exploring attitude change in a more nuanced manner, by comparing the prevalence of different attitude change categories across the SERP ranking bias conditions, we did not find evidence of attitude change that would indicate an effect of exposure bias like SEME. Across SERP ranking bias groups, most participants maintained their initial attitude, not aligned with the observations reported by Epstein and Robertson [22]. However, in this study, the authors considered attitude change for voting decisions, by investigating the impact of search results that favored one candidate in an election over a different candidate. Strong attitudes on debated topics are often rooted in stable moral values [32] and might be less prone to change than election decisions that were found to be impacted by less stable candidate qualities [12]. Drawing from our findings, we infer that the impact of exposure biases on searchers' attitudes towards debated topics, as well as search interactions, is likely mediated by nuanced pre-search epistemic states related to users' attitudes, such as their *importance*, *moral conviction*, or *certainty* [32, 37, 71]. Interestingly, our data indicated that participants with strong attitudes were less likely to gain knowledge than those with moderate and weak attitudes. Corroborating similar findings from prior work, participants with relatively high prior knowledge were less likely to gain more knowledge [23, 61].

Addressing **RQ2**, we discovered that searchers' attitude strength and prior knowledge, but not exposure effects, impact attitude change and knowledge gain. Searchers with relatively strong attitudes were less inclined to change their attitudes and were less likely to gain new knowledge, as did searchers with relatively high prior knowledge. Regarding **RQ3**, our results showed that attitude change is linked to the level of confirmation bias, with position bias or search effort having no significant impact. Specifically, individuals who exhibited higher confirmation bias in their search interactions were less likely to change their attitudes compared to those with lower confirmation bias. Similar to observations by Draws et al. [20], these findings suggest that the influence of search results on attitude change is primarily driven by selective user interaction rather than mere exposure. These observations expand on previous findings on the searcher-system interplay of exposure and interaction, revealing not only its impact on search behavior but also on attitude change. Our study did not uncover any evidence indicating that search interaction influenced knowledge gain.

Additional Observations. We noted first- and second-order interaction effects on aspects of the searcher and search system on search interaction and post-search epistemic states. We found interaction effects of SERP ranking bias and attitude strength on attitude-confirming clicks, click rank deviation, number of clicks, and hover depth, and interaction effects of SERP ranking bias and prior knowledge on the number of clicks and attitude change. Although our sample size was not sufficient to capture further higher-order interactions, we interpret these effects as indicators that exposure effects on search interaction as well as post-search epistemic states are likely shaped by prior knowledge and attitude. This underlines the need to investigate the role of more nuanced pre-search epistemic states related to searchers' attitudes in web search on debated topics, such as their *importance*, *moral conviction*, or *certainty* [32, 37, 71].

Apart from a second-order interaction effect on directional attitude change, we did not see any relations of searchers' *receptiveness to opposing views* to their search interaction and post-search epistemic states. This might be due to the focus of this user trait on engagement with and interpretation of information in the context of *passive exposure* [45]. However, individuals tend to turn to web search when *actively seeking* for information on a topic (locate, select, and access sources) [63]. Given the important role of active user interaction choices as opposed to passive exposure, a user trait that captures not only their receptiveness, but also whether they tend to actively seek out information with opposing views likely plays a more prominent role in search on debated topics (e.g., *intellectual humility* [25, 50, 55]).

5.1 Implications

In our study, we observed that viewpoint biases in exposure and interaction shape web search on debated topics and impede searchers from closing their knowledge gaps on the topic. Our data did not indicate a direct relation between exposure bias and attitude change. In particular, we did not observe a shift towards the opposing attitude in conditions of attitude-opposing ranking bias, i.e. SEME, as was previously reported [22]. However, when aiming for search interactions that enable responsible opinion formation, our findings show that we should place equal or even greater emphasis on the effects of exposure bias aligned with user attitudes. Participants who were exposed to attitude-supporting ranking bias exhibited particularly high confirmation and position bias and, in turn, a low likelihood of engaging with attitude-opposing results, behavior that impedes well-informedness and responsible opinion formation. Yet, from the perspective of search engines that optimize for user satisfaction, their behavior of engaging primarily with highly-ranked result could be interpreted as a signal for well-tailored relevance criteria [4, 33]. This underscores that, when dealing with debated topics, personalization by optimizing relevance criteria with the objective of increasing satisfaction of the individual user can reinforce their biases [9, 11], and thus hinder engagement with diverse viewpoints [26, 57].

Recently, there have been calls for improved search systems that provide better support for complex information needs [64, 66]. During search on debated topics, searchers could be better supported in closing their knowledge gaps and engaging with diverse viewpoints [26, 57]. Given the role of user interaction noted in this and other studies [20, 60, 65], approaches that aim to provide better support for complex information needs should not only facilitate access to diverse viewpoints but prompt and empower searchers to productively interact with results that advance their level of informedness. Our findings of knowledge- and attitude-related differences suggest that for such interventions to effectively support productive searches on debated topics, they would likely need to be personalized. This raises the question on *which searchers* would benefit from *what* kind of support and *when.*

Individuals with strong attitudes and limited knowledge, who, according to our observations, exhibited the most pronounced confirmation bias, might need interventions that motivate engagement with attitude-opposing results (e.g., warning labels [59]) but prove to be challenging to reach, regardless of the intervention. Individuals with high prior knowledge, on the other hand, exhibited low interaction bias. Yet, they might have interacted with attitude-opposing results to counter-argue and discount their content [17, 68], instead of objectively assessing it to close knowledge gaps. These searchers might not require support to interact with attitude-opposing information but instead to objectively assess their content (e.g., boosting strategies for informed search [10] or intellectual humility [55, 56]). While these findings of knowledge- and attitude-related differences suggest the need for personalized interventions to support search for responsible opinion formation, such interventions would require data on an individual's views, raising privacy concerns. Consequently, this would have to be approached in a privacy-aware manner, e.g., by ensuring that users can understand and control what shapes their user model and how it affects the information environment [70].

While this study focuses on web search, understanding the effects of exposure, interaction, and their interplay is highly relevant to similar domains of web interactions with information that can impact individuals' opinions, such as news recommenders, social media platforms, or discussion forums. Recent research on how individuals from Gen Z (born between 1997 and 2012) engage with information online revealed that they tend to encounter rather than actively search for information and that their interactions are strongly driven by social motivations rather than by truth-seeking [31]. However, Hassoun et al. [31] remarked that Gen Z highly values *information sensibility* which the authors define as *a socially-informed awareness of the value of information encountered online.* This could serve as a motivating factor for cultivating web interactions for responsible opinion formation.

5.2 Caveats, Limitations, and Future Work

As with all empirical and exploratory research, our study is not without limitations. We framed the information search task in a single search session with preselected, viewpoint-annotated SERP results and focused on participants' clicking behavior. Factors such as querying, query refinement, or multiple search sessions were beyond the scope of this work. Nonetheless, the qualitative feedback we collected was mostly positive and did not indicate any frustration regarding the restricted interaction options. Although we considered two topics to represent different levels of controversy, we plan to consider diverse tasks and additional topics to enhance the scope of generalizability of our findings. Future research should also gauge the long-term impact of the search session on attitude, knowledge, and decision-making by conducting confirmatory follow-up studies, enabling a more realistic search process, expanding the range of topics and tasks examined, and investigating mediating effects of search interaction measures.

Cognitive biases arising from the task design of crowdsourced user studies can negatively impact data quality [21]. We assessed potential biases in our study with the Cognitive Biases Checklist by Draws et al. [18], identifying that various cognitive biases could have affected the search interaction data by causing participants to diverge from their usual behavior in a real search setting. The pre-search knowledge questionnaire might have caused *anchoring effects* by leading participants to search for answers to the knowledge questions they encountered. Further, the time invested in filling the pre-search questionnaires could have resulted in a *sunk cost fallacy* of sticking to and rushing through the task to receive the final reward, even though participants may not be genuinely interested in the search task. To counter this potential bias, we encouraged genuine behavior within the search task with an added incentive of a bonus payment if they successfully identified three high-quality arguments in the search session.

We used knowledge questionnaires to capture the general depth of knowledge among searchers. However, we came to realize that one-dimensional knowledge questionnaires that capture the depth of topic knowledge do not suffice to measure *informedness* in the context of debated topics. We plan to broaden the scope of our measures to include the breadth of knowledge, i.e., encompassing knowledge about different viewpoints.

6 CONCLUSIONS

We presented the results of our user study to explore the relations between attributes of the searcher and search system, search interaction, and user post-search epistemic states—a first step towards developing a comprehensive and user-focused understanding of web search on debated topics. Our insights can inform the design of interventions that support responsible opinion formation. We observed that search interaction was shaped by search system exposure, attitude strength, and prior knowledge. Attitude change was not directly influenced by search system exposure but instead by participants' confirmation bias and their initial attitude strength. Our findings suggest that the effects of exposure and interaction biases, as well as their interplay on post-search epistemic states, likely depend on nuanced epistemic states related to searchers' attitudes. The knowledge and attitude-dependent differences suggest that interventions to support fruitful search interactions likely require privacy-conscious personalization, adapting the intervention to users' pre-search epistemic states. These insights further underscore that customizing search rankings based on implicit feedback to enhance user satisfaction can have harmful repercussions in the context of debated topics, as it is prone to cultivate attitude-supporting exposure bias, thereby reinforcing confirmation bias and hindering responsible opinion formation. Our findings could extend to other web interactions involving exposure to algorithmically curated information that has the potential to influence opinions, such as interactions on social media platforms or with news recommender systems. This is particularly of interest in the era of transition from the traditional linear information journey—from queries to answers—to more fluid and socially-oriented journeys that alternate between information-encountering and information-seeking [31].

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860621. This work was partially supported by the TU Delft AI Initiative. We also thank all participants from Prolific and the anonymous reviewers for their feedback.

REFERENCES

[1] 2023. Britannica ProCon.org. https://www.procon.org/. [Online; accessed 2023-02-01].
[2] 2023. Prolific. https://www.prolific.co/. [Online; accessed 2023-08-10].
[3] 2023. Qualtrics. https://www.qualtrics.com/. [Online; accessed 2023-08-10].
[4] Eugene Agichtein, Eric Brill, and Susan Dumais. 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 (SIGIR '06)*. Association for Computing Machinery, New York, NY, USA, 19–26. https://doi.org/10.1145/1148170.1148177
[5] Hosam Al-Samarraie, Atef Eldenfria, and Husameddin Dawoud. 2017. The Impact of Personality Traits on Users' Information-Seeking Behavior. *Information Processing & Management* 53 (2017), 237–247. https://doi.org/10.1016/j.ipm.2016.08.004
[6] Ahmed Allam, Peter Johannes Schulz, and Kent Nakamoto. 2014. The Impact of Search Engine Selection and Sorting Criteria on Vaccination Beliefs and Attitudes: Two Experiments Manipulating Google Output. *Journal of Medical Internet Research* 16 (2014), e2642. https://doi.org/10.2196/jmir.2642
[7] Kumaripaba Athukorala, Dorota Głowacka, Giulio Jacucci, Antti Oulasvirta, and Jilles Vreeken. 2016. Is exploratory search different? A comparison of information search behavior for exploratory and lookup tasks. *Journal of the Association for Information Science and Technology* 67, 11 (2016), 2635–2651.
[8] Leif Azzopardi. 2021. Cognitive biases in search: a review and reflection of cognitive biases in Information Retrieval. In *Proceedings of the 2021 conference on human information interaction and retrieval*. 27–37.
[9] Ricardo Baeza-Yates. 2018. Bias on the Web. *Commun. ACM* 61 (2018), 54–61.
[10] Markus Bink and David Elsweiler. 2024. Balancing Act: Boosting Strategies for Informed Search on Controversial Topics. In *Proceedings of the 2024 Conference on Human Information Interaction and Retrieval (CHIIR '24)*. Association for Computing Machinery, New York, NY, USA, 254–265. https://doi.org/10.1145/3627508.3638329
[11] Robin Burke. 2022. Personalization, Fairness, and Post-Userism. *Perspectives on Digital Humanism* (2022), 145.
[12] Matthew K. Buttice and Walter J. Stone. 2012. Candidates Matter: Policy and Quality Differences in Congressional Elections. *The Journal of Politics* 74 (2012), 870–887. https://doi.org/10.1017/S0022381612000394
[13] John T. Cacioppo, Richard E. Petty, and Katherine J. Morris. 1983. Effects of Need for Cognition on Message Evaluation, Recall, and Persuasion. *Journal of Personality and Social Psychology* 45 (1983), 805–818. https://doi.org/10.1037/0022-3514.45.4.805
[14] Cansu Canca. 2022. Did you find it on the internet? ethical complexities of search engine rankings. *Perspectives on Digital Humanism* (2022), 135.
[15] Noel Carroll. 2014. In Search We Trust: Exploring How Search Engines Are Shaping Society. *International Journal of Knowledge Society Research (IJKSR)* 5 (2014), 12–27. https://doi.org/10.4018/ijksr.2014010102
[16] Federico Maria Cau and Nava Tintarev. 2024. Navigating the Thin Line: Examining User Behavior in Search to Detect Engagement and Backfire Effects. *arXiv preprint arXiv:2401.11201* (2024).
[17] Jason K. Clark and Duane T. Wegener. 2013. Chapter Four - Message Position, Information Processing, and Persuasion: The Discrepancy Motives Model. In *Advances in Experimental Social Psychology*, Patricia Devine and Ashby Plant (Eds.). Vol. 47. Academic Press, 189–232. https://doi.org/10.1016/B978-0-12-407236-7.00004-8
[18] Tim Draws, Alisa Rieger, Oana Inel, Ujwal Gadiraju, and Nava Tintarev. 2021. A Checklist to Combat Cognitive Biases in Crowdsourcing. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* 9 (2021), 48–59. https://doi.org/10.1609/hcomp.v9i1.18939
[19] Tim Draws, Nirmal Roy, Oana Inel, Alisa Rieger, Rishav Hada, Mehmet Orcun Yalcin, Benjamin Timmermans, and Nava Tintarev. 2023. Viewpoint Diversity in Search Results. In *Advances in Information Retrieval (Lecture Notes in Computer Science)*, Jaap Kamps, Lorraine Goeuriot, Fabio Crestani, Maria Maistro, Hideo Joho, Brian Davis, Cathal Gurrin, Udo Kruschwitz, and Annalina Caputo (Eds.). Springer Nature Switzerland, Cham, 279–297. https://doi.org/10.1007/978-3-031-28244-7_18
[20] Tim Draws, Nava Tintarev, Ujwal Gadiraju, Alessandro Bozzon, and Benjamin Timmermans. 2021. This is not what we ordered: Exploring why biased search result rankings affect user attitudes on debated topics. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 295–305.
[21] Carsten Eickhoff. 2018. Cognitive Biases in Crowdsourcing. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. ACM, Marina Del Rey CA USA, 162–170. https://doi.org/10.1145/3159652.3159654
[22] Robert Epstein and Ronald E. Robertson. 2015. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences* 112, 33 (Aug. 2015), E4512–E4521. https://doi.org/10.1073/pnas.1419828112
[23] Ujwal Gadiraju, Ran Yu, Stefan Dietze, and Peter Holtz. 2018. Analyzing knowledge gain of users in informational search sessions on the web. In *Proceedings of the 2018 conference on human information interaction & retrieval*. 2–11.
[24] Lisa Gevelber. 2018. *It's all about 'me'—how people are taking search personally.* Technical Report. https://www.thinkwithgoogle.com/marketing-strategies/search/personal-needs-search-trends/
[25] Tim Gorichanaz. 2022. Relating Information Seeking and Use to Intellectual Humility. *Journal of the Association for Information Science and Technology* 73 (2022), 643–654. https://doi.org/10.1002/asi.24567
[26] Tim Gorichanaz. 2023. Virtuous Search: A Framework for Intellectual Virtue in Online Search. *Journal of the Association for Information Science and Technology* n/a (2023). https://doi.org/10.1002/asi.24832
[27] Jutta Haider and Olof Sundin. 2019. *Invisible Search and Online Search Engines: The Ubiquity of Search in Everyday Life.* Taylor & Francis. https://doi.org/10.4324/9780429448546
[28] Uriel Haran, Ilana Ritov, and Barbara A. Mellers. 2013. The Role of Actively Open-Minded Thinking in Information Acquisition, Accuracy, and Calibration. *Judgment and Decision Making* 8 (2013), 188–201.
[29] Eszter Hargittai, Lindsay Fullerton, Ericka Menchen-Trevino, and Kristin Yates Thomas. 2010. Trust Online: Young Adults' Evaluation of Web Content. *International Journal of Communication* 4 (2010), 27.
[30] William Hart, Dolores Albarracín, Alice H. Eagly, Inge Brechan, Matthew J. Lindberg, and Lisa Merrill. 2009. Feeling Validated versus Being Correct: A Meta-Analysis of Selective Exposure to Information. *Psychological Bulletin* 135 (2009), 555–588. https://doi.org/10.1037/a0015701
[31] Amelia Hassoun, Ian Beacock, Sunny Consolvo, Beth Goldberg, Patrick Gage Kelley, and Daniel M. Russell. 2023. Practicing Information Sensibility: How Gen Z Engages with Online Information. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3544548.3581328

[32] Lauren C. Howe and Jon A. Krosnick. 2017. Attitude Strength. *Annual Review of Psychology* 68 (2017), 327–351. https://doi.org/10.1146/annurev-psych-122414-033600

[33] Thorsten Joachims. 2002. Optimizing Search Engines Using Clickthrough Data. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '02)*. Association for Computing Machinery, New York, NY, USA, 133–142. https://doi.org/10.1145/775047.775067

[34] Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, and Geri Gay. 2017. Accurately Interpreting Clickthrough Data as Implicit Feedback. *ACM SIGIR Forum* 51 (2017), 4–11. https://doi.org/10.1145/3130332.3130334

[35] Dan M. Kahan. 2015. The Politically Motivated Reasoning Paradigm.

[36] Matthew Kelly. 2021. Epistemology, Epistemic Belief, Personal Epistemology, and Epistemics: A Review of Concepts as They Impact Information Behavior Research. *Journal of the Association for Information Science and Technology* 72 (2021), 507–519. https://doi.org/10.1002/asi.24422

[37] Silvia Knobloch-Westerwick and Jingbo Meng. 2009. Looking the Other Way: Selective Exposure to Attitude-Consistent and Counterattitudinal Political Information. *Communication Research* 36 (2009), 426–448. https://doi.org/10.1177/0093650209333030

[38] Silvia Knobloch-Westerwick, Benjamin K. Johnson, and Axel Westerwick. 2015. Confirmation Bias in Online Searches: Impacts of Selective Exposure Before an Election on Political Attitude Strength and Shifts. *Journal of Computer-Mediated Communication* 20 (2015), 171–187. https://doi.org/10.1111/jcc4.12105

[39] Hilary Kornblith. 1983. Justified Belief and Epistemically Responsible Action. *The Philosophical Review* 92 (1983), 33–48. https://doi.org/10.2307/2184520

[40] Mark R. Leary, Kate J. Diebels, Erin K. Davisson, Katrina P. Jongman-Sereno, Jennifer C. Isherwood, Kaitlin T. Raimi, Samantha A. Deffler, and Rick H. Hoyle. 2017. Cognitive and Interpersonal Features of Intellectual Humility. *Personality and Social Psychology Bulletin* 43 (2017), 793–813. https://doi.org/10.1177/0146167217697695

[41] Jingjing Liu and Xiangmin Zhang. 2019. The Role of Domain Knowledge in Document Selection from Search Results. *Journal of the Association for Information Science and Technology* 70 (2019), 1236–1247. https://doi.org/10.1002/asi.24199

[42] David Maxwell and Claudia Hauff. 2021. LogUI: Contemporary Logging Infrastructure for Web-Based Experiments. In *European Conference on Information Retrieval*. Springer, 525–530.

[43] Microsoft Bing. 2023. Web Search API. https://www.microsoft.com/en-us/bing/apis/bing-web-search-api. [Online; accessed 2023-08-10].

[44] Boaz Miller and Isaac Record. 2013. Justified Belief in the Digital Age: On the Epistemic Implications of Secret Internet Technologies. *Episteme* 10 (2013), 117–134. https://doi.org/10.1017/epi.2013.11

[45] Julia A. Minson, Frances S. Chen, and Catherine H. Tinsley. 2019. Why Won't You Listen to Me? Measuring Receptiveness to Opposing Views. *Management Science* (2019). https://doi.org/10.1287/mnsc.2019.3362

[46] Safiya Umoja Noble. 2018. Algorithms of Oppression: How Search Engines Reinforce Racism. In *Algorithms of Oppression*. New York University Press. https://doi.org/10.18574/nyu/9781479833641.001.0001

[47] Rik Peels. 2016. *Responsible Belief: A Theory in Ethics and Epistemology*. Oxford University Press.

[48] Gloria Phillips-Wren and Monica Adya. 2020. Decision Making under Stress: The Role of Information Overload, Time Pressure, Complexity, and Uncertainty. *Journal of Decision Systems* 29 (Aug. 2020), 213–225. https://doi.org/10.1080/12460125.2020.1768680

[49] Frances A. Pogacar, Amira Ghenai, Mark D. Smucker, and Charles L.A. Clarke. 2017. The Positive and Negative Influence of Search Results on People's Decisions about the Efficacy of Medical Treatments. In *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval (ICTIR '17)*. Association for Computing Machinery, New York, NY, USA, 209–216. https://doi.org/10.1145/3121050.3121074

[50] Tenelle Porter, Abdo Elnakouri, Ethan A. Meyers, Takuya Shibayama, Eranda Jayawickreme, and Igor Grossmann. 2022. Predictors and Consequences of Intellectual Humility. *Nature Reviews Psychology* 1 (2022), 524–536. https://doi.org/10.1038/s44159-022-00081-9

[51] Tenelle Porter and Karina Schumann. 2018. Intellectual Humility and Openness to the Opposing View. *Self and Identity* 17 (2018), 139–162. https://doi.org/10.1080/15298868.2017.1361861

[52] Tenelle Porter, Karina Schumann, Diana Selmeczy, and Kali Trzesniewski. 2020. Intellectual Humility Predicts Mastery Behaviors When Learning. *Learning and Individual Differences* 80 (2020), 101888. https://doi.org/10.1016/j.lindif.2020.101888

[53] Emmanuel M. Pothos, Stephan Lewandowsky, Irina Basieva, Albert Barque-Duran, Katy Tapper, and Andrei Khrennikov. 2021. Information Overload for (Bounded) Rational Agents. *Proceedings of the Royal Society B: Biological Sciences* 288 (Feb. 2021), 20202957. https://doi.org/10.1098/rspb.2020.2957

[54] Kristen Purcell, Lee Rainie, and Joanna Brenner. 2012. Search engine use 2012. (2012).

[55] Alisa Rieger, Frank Bredius, Mariët Theune, and Maria Soledad Pera. 2024. From Potential to Practice: Intellectual Humility During Search on Debated Topics. In *Proceedings of the 2024 Conference on Human Information Interaction and Retrieval (CHIIR '24)*. Association for Computing Machinery, New York, NY, USA, 130–141. https://doi.org/10.1145/3627508.3638306

[56] Alisa Rieger, Frank Bredius, Nava Tintarev, and Maria Soledad Pera. 2023. Searching for the Whole Truth: Harnessing the Power of Intellectual Humility to Boost Better Search on Debated Topics. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3544549.3585693

[57] Alisa Rieger, Tim Draws, Nicolas Mattis, David Maxwell, David Elsweiler, Ujwal Gadiraju, Dana McKay, Alessandro Bozzon, and Maria Soledad Pera. 2024. Responsible Opinion Formation on Debated Topics in Web Search. In *Advances in Information Retrieval*, Nazli Goharian, Nicola Tonellotto, Yulan He, Aldo Lipani, Graham McDonald, Craig Macdonald, and Iadh Ounis (Eds.). Springer Nature Switzerland, Cham, 437–465. https://doi.org/10.1007/978-3-031-56066-8_32

[58] Alisa Rieger, Tim Draws, Mariët Theune, and Nava Tintarev. 2021. This Item Might Reinforce Your Opinion: Obfuscation and Labeling of Search Results to Mitigate Confirmation Bias. In *Proceedings of the 32st ACM Conference on Hypertext and Social Media*. ACM, Virtual Event USA, 189–199. https://doi.org/10.1145/3465336.3475101

[59] Alisa Rieger, Tim Draws, Mariët Theune, and Nava Tintarev. 2023. Nudges to Mitigate Confirmation Bias during Web Search on Debated Topics: Support vs. Manipulation. *ACM Transactions on the Web* (2023). https://doi.org/10.1145/3635034

[60] Ronald E. Robertson, Jon Green, Damian J. Ruck, Katherine Ognyanova, Christo Wilson, and David Lazer. 2023. Users Choose to Engage with More Partisan News than They Are Exposed to on Google Search. *Nature* (2023), 1–7. https://doi.org/10.1038/s41586-023-06078-5

[61] Nirmal Roy, Manuel Valle Torre, Ujwal Gadiraju, David Maxwell, and Claudia Hauff. 2021. Note the highlight: incorporating active reading tools in a search as learning environment. In *Proceedings of the 2021 conference on human information interaction and retrieval*. 229–238.

[62] Ladislao Salmerón, Yvonne Kammerer, and Pilar García-Carrión. 2013. Searching the Web for Conflicting Topics: Page and User Factors. *Computers in Human Behavior* 29 (2013), 2161–2171. https://doi.org/10.1016/j.chb.2013.04.034

[63] Reijo Savolainen. 2016. Elaborating the Conceptual Space of Information-Seeking Phenomena. *Information Research: An International Electronic Journal* 21 (2016).

[64] Chirag Shah and Emily M. Bender. 2022. Situating Search. In *ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '22)*. Association for Computing Machinery, New York, NY, USA, 221–232. https://doi.org/10.1145/3498366.3505816

[65] Laura Slechten, Cédric Courtois, Lennert Coenen, and Bieke Zaman. 2022. Adapting the Selective Exposure Perspective to Algorithmically Governed Platforms: The Case of Google Search. *Communication Research* 49 (2022), 1039–1065. https://doi.org/10.1177/00936502211012154

[66] Catherine L. Smith and Soo Young Rieh. 2019. Knowledge-Context in Search Systems: Toward Information-Literate Actions. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval (CHIIR '19)*. Association for Computing Machinery, New York, NY, USA, 55–62. https://doi.org/10.1145/3295750.3298940

[67] Olof Sundin, Dirk Lewandowski, and Jutta Haider. 2022. Whose Relevance? Web Search Engines as Multisided Relevance Machines. *Journal of the Association for Information Science and Technology* 73 (2022), 637–642. https://doi.org/10.1002/asi.24570

[68] Charles S. Taber, Damon Cann, and Simona Kucsova. 2009. The Motivated Processing of Political Arguments. *Political Behavior* 31 (2009), 137–155.

[69] Kjerstin Thorson and Chris Wells. 2016. Curated Flows: A Framework for Mapping Media Exposure in the Digital Age. *Communication Theory* 26 (2016), 309–328. https://doi.org/10.1111/comt.12087

[70] Eran Toch, Yang Wang, and Lorrie Faith Cranor. 2012. Personalization and Privacy: A Survey of Privacy Risks and Remedies in Personalization-Based Systems. *User Modeling and User-Adapted Interaction* 22 (2012), 203–220. https://doi.org/10.1007/s11257-011-9110-z

[71] Zakary L. Tormala and Derek D. Rucker. 2018. Attitude Certainty: Antecedents, Consequences, and New Directions. *Consumer Psychology Review* 1 (2018), 72–89. https://doi.org/10.1002/arcp.1004

[72] Johan L. H. van Strien, Yvonne Kammerer, Saskia Brand-Gruwel, and Henny P. A. Boshuizen. 2016. How Attitude Strength Biases Information Processing and Evaluation on the Web. *Computers in Human Behavior* 60 (2016), 245–252. https://doi.org/10.1016/j.chb.2016.02.057

[73] Dáša Vedejová and Vladimíra Čavojová. 2022. Confirmation Bias in Information Search, Interpretation, and Memory Recall: Evidence from Reasoning about Four Controversial Topics. *Thinking & Reasoning* 28 (2022), 1–28. https://doi.org/10.1080/13546783.2021.1891967

[74] Henning Wachsmuth, Martin Potthast, Khalid Al-Khatib, Yamen Ajjour, Jana Puschmann, Jiani Qu, Jonas Dorsch, Viorel Morari, Janek Bevendorff, and Benno Stein. 2017. Building an Argument Search Engine for the Web. In *Proceedings of the 4th Workshop on Argument Mining*. Association for Computational Linguistics,

Copenhagen, Denmark, 49–59. https://doi.org/10.18653/v1/W17-5106

[75] Ryen White. 2013. Beliefs and Biases in Web Search. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '13)*. Association for Computing Machinery, New York, NY, USA, 3–12. https://doi.org/10.1145/2484028.2484053

[76] Ryen W. White. 2014. Belief Dynamics in Web Search. *Journal of the Association for Information Science and Technology* 65 (2014), 2165–2178. https://doi.org/10.1002/asi.23128

[77] Ryen W. White, Susan T. Dumais, and Jaime Teevan. 2009. Characterizing the Influence of Domain Expertise on Web Search Behavior. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining (WSDM '09)*. Association for Computing Machinery, New York, NY, USA, 132–141. https://doi.org/10.1145/1498759.1498819

[78] Ryen W. White and Eric Horvitz. 2015. Belief Dynamics and Biases in Web Search. *ACM Transactions on Information Systems* 33 (2015), 18:1–18:46. https://doi.org/10.1145/2746229

[79] Ryen W. White and Resa A. Roth. 2009. Exploratory Search: Beyond the Query-Response Paradigm. *Synthesis Lectures on Information Concepts, Retrieval, and Services* 1 (2009), 1–98. https://doi.org/10.2200/S00174ED1V01Y200901ICR003

[80] Zhangyi Wu, Tim Draws, Federico Cau, Francesco Barile, Alisa Rieger, and Nava Tintarev. 2023. Explaining Search Result Stances to Opinionated People. In *Explainable Artificial Intelligence (Communications in Computer and Information Science)*, Luca Longo (Ed.). Springer Nature Switzerland, Cham, 573–596. https://doi.org/10.1007/978-3-031-44067-0_29

[81] Ran Yu, Ujwal Gadiraju, Peter Holtz, Markus Rokicki, Philipp Kemkes, and Stefan Dietze. 2018. Predicting user knowledge gain in informational search sessions. In *The 41st international ACM SIGIR conference on research & development in information retrieval*. 75–84.