UNDERSTANDING AND MITIGATING COGNITIVE BIAS DURING WEB SEARCH

S. Hitzginger, A. Nussbaumer, C. Gütl, C. Ruß-Baumann
Graz University of Technology, Graz, Austria

Abstract

When conducting research on the internet, Confirmation Bias can cause individuals to selectively retrieve information that confirms their pre-existing beliefs. While current tools are available to counteract this effect, they often require additional data that is not always accessible during regular search sessions. Therefore, we aimed to develop a tool capable of detecting Confirmation Bias by simply tracking general search behavior. Through an online study with 43 participants that simulated real web searches, we identified behavioral patterns, like a difference in number of issued search queries, number of clicked results, and position of clicked results, that can predict bias in retrieved information. With those results, we created a browser extension that can track and analyze relevant behaviors in real-time and alert the user of potentially biased search sessions.

INTRODUCTION

The Internet has revolutionized the way we access and consume information, providing us with unprecedented opportunities for knowledge and connectivity. However, the abundance of ambiguous information also amplifies the likelihood of Confirmation Bias, described as the seeking or interpreting of evidence in ways that are partial to existing beliefs [1]. Phenomenons like *filter bubbles*, which are a product of the intention of social media platforms to present information to users that is interesting to them [2], or echo chambers, which have a similar effect but are mostly caused by the homogeneity of friend groups on social media platforms [3], are common manifestations of Confirmation Bias on the Internet, only partially caused by the user. Likewise, when using a search engine, localization and personalization of search results makes unbiased information retrieval a difficult task. The natural human tendency to prefer belief-consistent information further increases the likelihood of the formation of a one-sided opinion.

An individual's tendency to Confirmation Bias is moderated by various other factors, like the exposure to belief-inconsistent information [4], a challenge-averse personality [5], or the involvement and the perceived threat of a certain topic [4]. There are also a number of tools available that could help individuals reduce their susceptibility to echo chambers [6], filter bubbles [7], or Confirmation Bias in general [8]. However, each of those tools relies on information about the content of visited websites. Therefore, their applicability is limited to pre-classified Websites, or content that can be analyzed and classified automatically.

The aim of this work is to create a tool without such limitations. To do so, instead of analyzing the content or context of information, we focus on tracking behaviour during the use of a search engine, the most common way of navigating the Internet. This way, the detection of bias is no longer dependent on knowledge about visited websites. Previous literature already found indications of correlations between different behavioral patterns and bias during web search[9]. We want to confirm those findings in a more general setting and provide a tool capable of using the results in a real-life application.

STUDY DESIGN

To find behavioral features able to predict Confirmation Bias of users during information retrieval via web search, a study simulating a realistic search setting was conducted. For this purpose, we implemented a custom search engine[1] and hosted it on a web server to make it accessible online. Participants were given a specific search task which they should complete using our custom search engine.

For the topic of the search task, we chose the debate that was occurring at the time of the research on the *Legalization of THC-containing Cannabis for recreational use in Austria*, as it presented contrasting facts and diverse opinions. To get valid results, we attempted to create a real-life situation. Participants were asked to inform themselves about the search topic for as long as they thought necessary to then be able to answer whether Cannabis should be legalized in Austria for recreational use or not (see figure 1).

A total of 61 web articles on the topic of Cannabis were gathered from different news sites. Each of those articles was reformatted to only include text, separated in story title, lead, and content. Four raters judged the opinion expressed by each article, rating it either negative, neutral, or positive concerning Cannabis. Different ratings were then aggregated in the following way:

1. with 4 raters agreeing on the same opinion, this opinion was assigned to the article
2. with 3 raters having the same opinion and the fourth rater being neutral, the opinion was assigned to the article
3. with 3 raters having the same opinion and the fourth having an opposing opinion, the article was removed
4. with 2 raters having the same opinion and the other 2 rating the article as neutral, the opinion of the first two raters was assigned to the article
5. every other article with less agreement was removed

By doing this, the 61 considered articles were reduced to 52 articles which expressed a clear bias. Three more articles were removed to have an equal number of positive

[1] https://github.com/sihi9/cb_explostudy

and negative articles, resulting in 20 articles in each bias category and 9 neutral articles. Those 49 articles could then be found with the search engine.

Figure 1 shows a screenshot of the search task presented to participants. It was implemented using ReactJS for the frontend, ExpressJS[2] for the server, and MeiliSearch[3] for the database and search engine. Participants were able to type arbitrary queries, scroll the search engine result page (SERP), go to different pages of the SERP, and click on results to navigate to a new page showing the entire article. Table 1 shows all different behavior variables that were tracked during the search task.

The bias expressed during the search task was evaluated based on the articles that were clicked by a participant, summing up the ratings of clicked articles divided by the number of clicked articles, resulting in a bias rating in the range [-1, 1].

For a more detailed insight on Confirmation Bias, a small questionnaire with 6 questions about different aspects and personal attitude towards Cannabis, assessed with a 5-point-likert-scale ranging from -2 to 2, was also presented to participants, once prior to the search task, and once after the search task. A screenshot of the survey is shown in figure 2. The 6 answers were aggregated and normalized, with 2 items being inverted, to gain two variables representing the opinions towards Cannabis before and after the search task in the range [-1, 1], comparable to the bias expressed by clicked articles. Confirmation Bias can then be interpreted as an alignment of bias in viewed articles and the attitude before the search task.

RESULTS

Confirmation Bias

A total of 59 participants started the study. After removing incomplete attempts, participants that took less than 1 minute on the search tasks, and participants that did not click on a single search result, 43 participants remain, most of them being male (n = 27) and between 20 and 29 years old (n = 29).

The bias of viewed articles is almost normally distributed ($M = 0.05, SD = 0.43$). The attitude is skewed to the right (positive attitude towards Cannabis) with means of $M = 0.38$ before-, and $M = 0.36$ after the search task. To compute Confirmation Bias, the minimum distance to either the median of attitude (Mdn = 0.42), or the median of bias (Mdn = 0.08), was used, and multiplied by -1 if bias and attitude were not aligned. Figure 3 shows the relation between the variables, with the medians marked as black lines. The correlations between Confirmation Bias and behaviour variables are shown in table 2.

The attitude after the search task was used to evaluate the impact of Confirmation Bias. For only 5 participants, Confirmation Bias actually strengthened their belief in the previous attitude.

[2] https://expressjs.com/

[3] https://www.meilisearch.com/

Bias of viewed articles

Another approach was to only evaluate the absolute value of bias expressed by viewed articles. Correlations between the absolute bias and behavioral variables revealed significant correlations between bias and the number of queries, the number of clicked results, the average position (index) of clicked results, and the average page of clicked results, as can be seen in table 3. A linear regression with forced entry was calculated to get a prediction of the bias from usage variables. Due to the high correlation between average index and average page, only the average index was chosen for the regression, because the correlation coefficient is only slightly smaller than of the average page, and it would be more versatile later on, because the index is independent of results shown per search engine result page, and the index theoretically contains more information. The regression showed that average index, number of queries and number of clicked results could predict the absolute bias quite well ($R^2 = .446, F(3, 38) = 10.456, p < .001$). Table 4 shows the coefficients for each variable.

DISCUSSION

Results have shown that Confirmation Bias on the internet might not be as big of a problem as previously assumed. In fact, there is no lack of literature with similar results [4][10], showing that exposure to new information alone is often sufficient to cause a moderation of opinion. While this is good news, it does not mean that there is no Confirmation Bias on the internet. However, our results support previous findings, which suggest that susceptibility to Confirmation Bias depends both on the individual and the topic[4], and is generally not as high as expected. Therefore, identifying Confirmation Bias during web search seems to be a difficult problem, which would require larger sample sizes than this study could provide.

However, our results show interesting behavioral patterns when it comes to a bias in the selection of articles. As expected, showing more engagement in the search task, expressed through issuing more queries, clicking more results, and browsing further through the search engine result pages, is correlated with less bias in the selection of articles. With a simple linear regression only including three factors, we could explain a substantial amount of variance in the bias. The correlations shown in table 3 also hint at further relationships, which can potentially be proven with a larger sample size.

An interesting pattern also arises when comparing significant relationships of Confirmation Bias and behaviour with general bias and behaviour. While the number of clicked results and number of queries are the most significant predictors for bias in clicked articles, they seem to have no impact on Confirmation Bias. On the other hand, time spent on result pages, as well as the standard deviation of time spent on results, are the most significant predictors of Confirmation Bias, but do not significantly correlate with bias of clicked articles, although a tendency towards significance can be

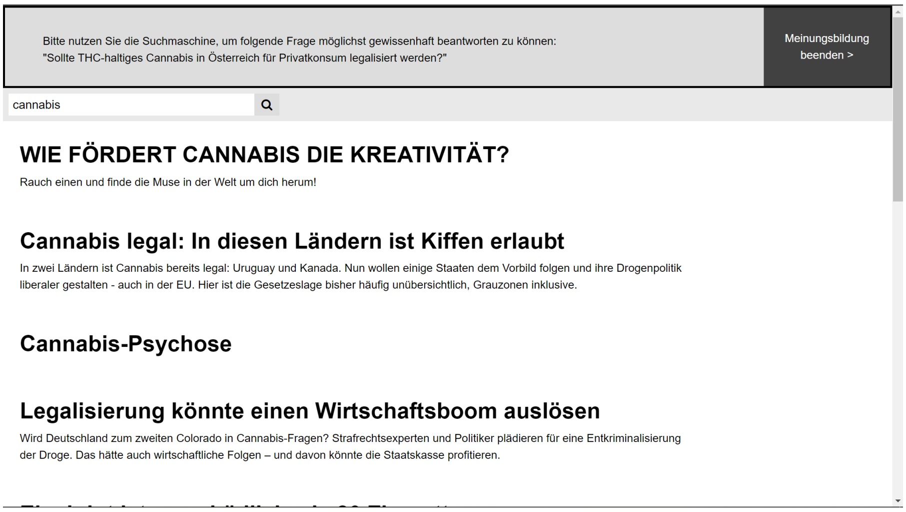

Figure 1: Screenshot of search task

Table 1: Behaviors tracked during search task

Metric	Description
Duration	Total duration spent on the information retrieval task
Time on SERP	Total time spent on a search engine result page (SERP)
Time on results	Total time spent on result pages
Number of queries	Number of different search queries used during information retrieval
Average query duration	Average time spent on each query
Standard deviation of query duration	Standard deviation of query duration
Number of clicked results	Number of clicked results in total
Average time per result	Average time spent on each result page
Standard deviation of time per result	Standard deviation of time spent on results
Average index	Average index (i.e., position of the result on the SERP) of clicked results
Average page	Average page of clicked results

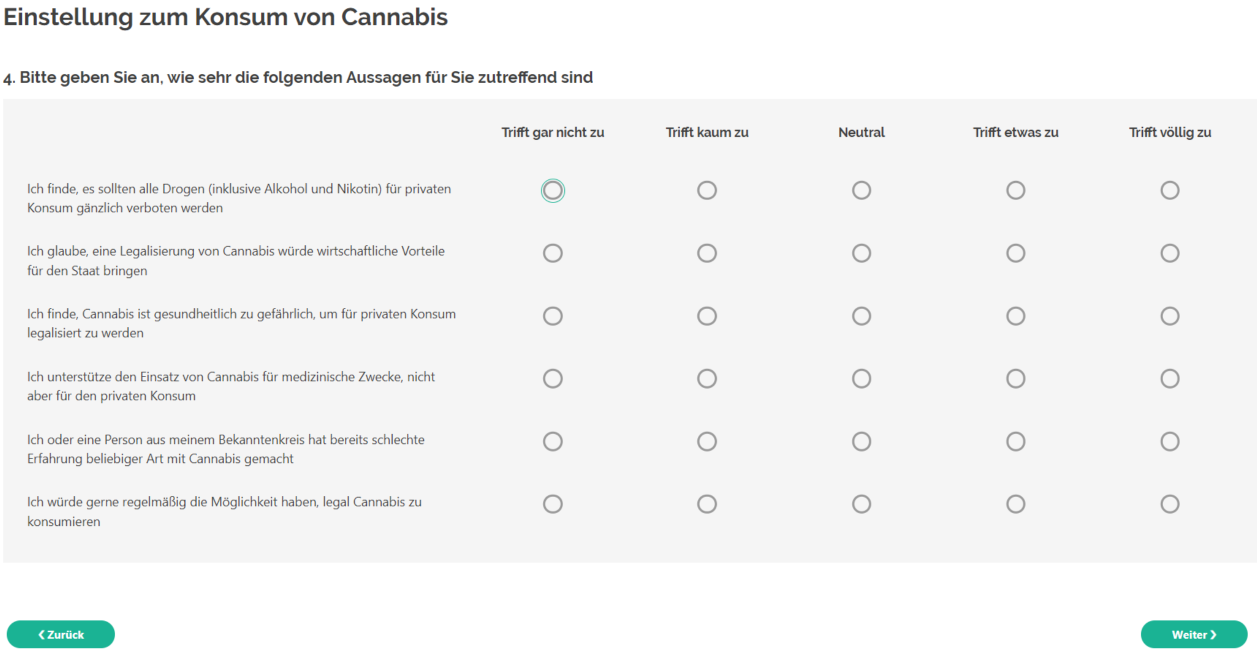

Figure 2: Screenshot of survey

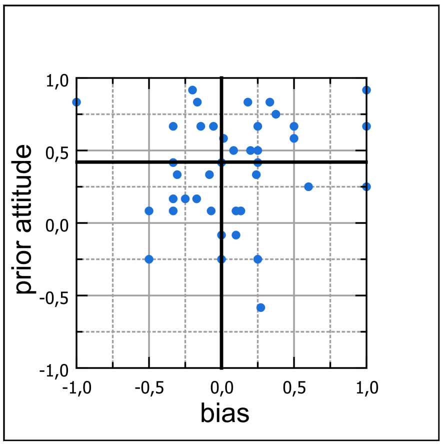

Figure 3: Scatterplot of bias and previous attitude, with black lines indicating the median of each variable.

Table 2: Correlations of factors with Confirmation Bias

Variable	Pearson correlation	*p*
duration	-.27	.08
time on results	-.316*	.039
number queries	-.005	.975
average query duration	-.298	.052
std. dev. query duration	-.052	.739
number clicked results	-.056	.72
average time per result	-.21	.176
std. dev. time per result	-.312*	.042
average index	.018	.907
average page	.031	.843

$^{**}p < .001, ^{*}p < 0.05$

Table 3: Correlations of factors with absolute bias

Variable	Pearson correlation	*p*
duration	-.182	.243
time on results	-.19	.222
number queries	-.329*	.031
average query duration	-.108	.49
std. dev. query duration	-.259	.093
number clicked results	-.495**	.001
average time per result	-.071	.651
std. dev. time per result	-.202	.195
average index	-.324*	.034
average page	-.35*	.021

$^{**}p < .001, ^{*}p < 0.05$

Table 4: Regression coefficients

Variable	*B*	*t*	*p*
constant	0.79	8.66	<.001
number queries	-0.038	-3.107	.004
number results	-0.042	-3.667	.001
average index	-0.019	-2.6	.013

observed. Spending more time reading articles might also be a form of showing engagement, and is therefore similar to the expected results. The correlation with the standard deviation of time spent on results however is more of a surprise. Apparently, spending more time on some results and less time on others is correlated with a higher Confirmation Bias. One could theorize that this is because participants susceptible to Confirmation Bias stop reading certain articles which do not confirm their opinion, but more research will be necessary to support such hypothesis.

PRACTICAL APPLICATION

Unfortunately, the results of the study are not as conclusive as hoped, probably due to lack of participants. To still prove the concept of a tool capable of detecting a bias in web search, we decided to focus on the bias of clicked articles, because it is a simpler construct with more predictive power than Confirmation Bias. However, the concept could, without much effort, be adapted to predict Confirmation Bias if future research is able to find factors with sufficient predictive power.

The tool[4] is built as a browser extension, both for Google Chrome and Firefox, as they are widely used Browsers that use two different APIs which are also used in most other browsers. Therefore, the tool can be extended for other browsers as well with minimal adjustments.

The extension tracks relevant browsing behavior, which can be classified in two relevant actions: search actions, and result-clicked-actions.

Search actions are defined as the issuing of a new search query. They can be tracked via the browser's history API, which receives a new entry whenever a query is issued with one of the supported search engines. To prove the concept, the tool currently supports two search engines, Google[5] and Ecosia[6]. To track search actions from different search engines, only the regex which detects the search query from the URL needs to be adapted.

To detect a result-clicked-action, content scripts are used. They contain JavaScript code, which is injected by the extension to specific sites, in our case the result pages of supported search engines. The code is able to add onClick-methods to the click-events of the HTML elements which can be clicked by the user to navigate to a search result. Those methods, in combination with the runtime API, can send messages to the extension, informing it that a result was clicked, as well as the exact position of the result.

The study used to analyze behaviour focuses on one specific search task. To use the results of the study, the extension has to be able to separate the search task into separate sessions. This is done by using two assumptions: queries belonging to the same search sessions are more likely to occur in temporal proximity, and use similar words. Time difference between queries is tracked automatically by the

[4] https://github.com/sihi9/cb_extension
[5] https://www.google.com/
[6] https://www.ecosia.org/

history API and does not require additional logic. To evaluate semantic similarity, word-vector similarities computed by the spaCy[7] library are used. When a new search is issued, the semantic similarity to all other search sessions is calculated. If there is a good fit with any other session, the new query is also assigned to this session. Otherwise, if not much time has passed since the last action in a session, and there is at least some semantic similarity between the new query and queries of the previous session, the new query is assigned to this session. Thresholds and decay functions are chosen partially based on the results of the study, and partially through trial and error. They can be improved by more research on browsing and search behavior, as well as improved similarity detection.

The bias of each separate session can then be analyzed, using the results from the regression of the bias in the study. The threshold between high and low bias is also chosen based on the results of the study, where a bias of 0.8 yielded a good division. All of those values can easily be adjusted once future research offers new insights into the behavior correlated with a bias in web search.

If a bias is detected, the user is informed of the bias with a warning symbol. Clicking on the symbol opens a popup which allows the user to view all of his potentially biased search sessions, remove them from the tracking, or continue searching for more articles on the session. A screenshot of the popup is shown in figure 4.

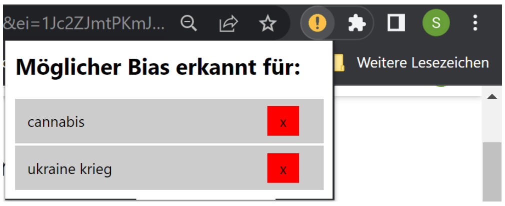

Figure 4: Screenshot of browser extension showing two biased search sessions

CONCLUSION

Confirmation Bias can occur in different forms and shapes. On the internet, during web search, people might tend to prefer content that confirms their preexisting beliefs. There are already tools to help counteract such phenomenons, but most of them require information about content, which is not always easy to acquire. In this paper, we showed the applicability of a different approach, which predicts bias only based on behaviour during web search. Due to the diversity of factors influencing Confirmation Bias and the lack of participants in this study, a lot of open question still remain on the best approach on the extent to which different behaviors correlate with bias. Nonetheless, we did show that there are moderate correlations, both between bias in selected articles and Confirmation Bias. We also provide a proof of concept for a tool capable of tracking browsing behavior and predicting biased search sessions using the results of our study.

[7] https://spacy.io/

ACKNOWLEDGEMENTS

This work has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101070014 (OpenWebSearch.EU, https://doi.org/10.3030/101070014) .

REFERENCES

[1] R. S. Nickerson, "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises," *Review of General Psychology*, vol. 2, pp. 175–220, Jun. 1998, Publisher: SAGE Publications Inc, ISSN: 1089-2680. DOI: 10.1037/1089-2680.2.2.175. https://doi.org/10.1037/1089-2680.2.2.175

[2] M. Cinelli, G. De Francisci Morales, A. Galeazzi, W. Quattrociocchi, and M. Starnini, "The echo chamber effect on social media," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 118, no. 9, e2023301118, Mar. 2021, ISSN: 0027-8424. DOI: 10.1073/pnas.2023301118.

[3] E. Bakshy, S. Messing, and L. A. Adamic, "Exposure to ideologically diverse news and opinion on Facebook," *Science*, Jun. 2015. https://www.science.org/doi/10.1126/science%20.aaa1160

[4] Q. V. Liao and W.-T. Fu, "Beyond the filter bubble: interactive effects of perceived threat and topic involvement on selective exposure to information," in *CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA: Association for Computing Machinery, Apr. 2013, pp. 2359–2368, ISBN: 978-1-45031899-0. DOI: 10.1145/2470654.2481326.

[5] S. A. Munson and P. Resnick, "Presenting diverse political opinions: how and how much," in *CHI '10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA: Association for Computing Machinery, Apr. 2010, pp. 1457–1466, ISBN: 978-1-60558929-9. DOI: 10.1145/1753326.1753543.

[6] S. Nagulendra and J. Vassileva, "Understanding and controlling the filter bubble through interactive visualization: A user study," *HT 2014 - Proceedings of the 25th ACM Conference on Hypertext and Social Media*, Sep. 2014. DOI: 10.1145/2631775.2631811.

[7] S. Faridani, E. Bitton, K. Ryokai, and K. Y. Goldberg, "Opinion space: A scalable tool for browsing online comments," *Conference on Human Factors in Computing Systems - Proceedings*, vol. 2, pp. 1175–1184, Jan. 2010. DOI: 10.1145/1753326.1753502.

[8] S. A. Munson and P. Resnick, "Encouraging Reading of Diverse Political Viewpoints with a Browser Widget," *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*, 2013.

[9] R. W. White and E. Horvitz, "Belief dynamics and biases in web search," *ACM Trans. Inf. Syst.*, vol. 33, no. 4, May 2015, ISSN: 1046-8188. DOI: 10.1145/2746229. https://doi.org/10.1145/2746229

[10] P. M. Dahlgren, "A critical review of filter bubbles and a comparison with selective exposure," *Nordicom Review*, vol. 42, no. 1, pp. 15–33, Jan. 2021. DOI: 10.2478/nor-2021-0002.