

Perception versus Reality: Evaluating User Awareness of Political Selective Exposure in News Recommender Systems

ALAIN D. STARKE, Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Netherlands and MediaFutures, University of Bergen, Norway

ANDERS SANDVIK BREMNES, MediaFutures, University of Bergen, Norway

ERIK KNUDSEN, MediaFutures, University of Bergen, Norway

DAMIAN TRILLING, Vrije Universiteit Amsterdam, Netherlands

CHRISTOPH TRATTNER, MediaFutures, University of Bergen, Norway

News Recommender Systems (NRSs) have become increasingly pivotal in shaping the news landscape, particularly in how news is disseminated. This has also led to concerns about information diversity, especially regarding selective exposure in the realm of political news. Users may not recognize that news content presented to them is subject to selective exposure, through users that incorporate political beliefs. Within the U.S. two-party system, our research explores the interactions between NRSs and users' ability to discern news articles that align with their political biases. We performed an online experiment ($N = 160$) to address the issue of user awareness and self-recognition of selective exposure within NRSs. Users were asked to select any number of news articles that matched their political orientation (i.e., Democrat or Republican) from a list of 50 news articles (5 Democrat, 5 Republican, 40 filler articles), which were either ranked saliently towards their political orientation or randomly. Contrary to expectations, our findings reveal no significant difference in article selection between participants exposed to a baseline random order and those who were presented with the more salient and easy to select version. We did observe that Republicans performed worse than Democrats in identifying aligning articles, based on precision and recall metrics.

CCS Concepts: • **Information systems** → **Recommender systems**; • **Human-centered computing** → *Human computer interaction (HCI)*.

Additional Key Words and Phrases: News Recommender Systems (NRSs), User Awareness, Selective Exposure, Political Preferences

ACM Reference Format:

Alain D. Starke, Anders Sandvik Bremnes, Erik Knudsen, Damian Trilling, and Christoph Trattner. 2024. Perception versus Reality: Evaluating User Awareness of Political Selective Exposure in News Recommender Systems. In *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization (UMAP Adjunct '24)*, July 1–4, 2024, Cagliari, Italy. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3631700.3665189>

1 INTRODUCTION

News Recommender Systems (NRSs) have become increasingly pivotal in shaping the way online news are being consumed [8]. These systems aim to personalize news content presented to the user, enhancing both user experience and engagement. Due to the relevancy of such systems, multiple proposals on how to improve these systems have been made over the last few decades [14, 18], leading to significant advances in recommender systems.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2024 Copyright held by the owner/author(s).

Manuscript submitted to ACM

Despite their advancements, NRSs have introduced several challenges and debates. Key among these are concerns about information diversity and user preference modeling [14, 21]. Scenarios in which only content is presented that aligns with users' existing beliefs (i.e., echo chambers) [3], potentially leading to reinforcement of biases [4]. These challenges raise concerns about the role of NRS in shaping the way online news are disseminated, because there may be differences across political affiliation in terms of news consumption, how they are targeted, and susceptibility for false news-related beliefs [1]

A particularly critical aspect centers on selective exposure, especially in the context of news consumption through recommender systems. Selective exposure refers to the tendency of individuals to favor information that aligns with their existing beliefs and attitudes, potentially exacerbating polarization and echo chambers in online spaces [2, 10]. Critics argue that NRSs can lead to an increase in selective exposure, which may pose a threat to democratic values [22]. Several studies have explored the impact of selective exposure in NRSs, examining how these systems can influence users' political perceptions and engagement [10]. Additionally, research suggests that some NRSs can contribute to both an increase and decrease in selective exposure to political articles [15]. However, there is little attention in news recommender research for whether people actually understand they are being shown selective or biased content.

In this paper, we investigate users' proficiency in identifying political articles that align with their preferred ideologies within a recommender news website. We present an experimental study using a modified version of the TREC Washington Post corpus, examining users' proficiency in recognizing news articles aligning with their political preferences. We build upon earlier work from Seddik et al. [15], that also examined selective exposure in the context of NRSs using a similar infrastructure. We extend their work by specifically focusing on the user's capacity to identify news articles that are presented for users with a specific political orientation. Moreover, we use a more varied set of metrics, including precision and recall.

We employ the Recommender Influenced Selective Exposure framework (RISE) [10], which was used in our own earlier work. This is a framework that proposes that NRS' influence on selective exposure is conditional upon its design objectives. The RISE framework allows analysts to model NRS' design objectives as causal variables and identify the conditions under which NRS amplifies or reduces selective exposure in online news environments. One possible design objective could be to reduce selective exposure through nudging political news articles that are not aligned with a user's political preference so that such articles are more salient and prominent on a front page compared to politically like-minded news—a so-called *salience factor* [15]. Indeed, prior research also finds empirical evidence to support the claim that NRS design decisions can influence selective exposure [10, 15].

Selective exposure theory assumes that individuals are likely to prefer information they agree with over information they disagree with, building on the model of dual-process modes of thinking [7]. We can distinguish between how news selection decisions are processed: spontaneous, fast, and effortless (System 1) or slow, deliberative, and effortful (System 2) [11]. Users' selective exposure decisions are generally assumed to operate through System 1. In this paper, we do not study the degree to which a salience factor influences news exposure on the automatic and unconscious level of decision-making (System 1), but rather how a salience nudge influences users' ability to detect news stories that are aligned with their political preferences, that is, decisions we assume operate through System 2. Our research is guided by the following research question:

RQ: To what extent do readers' ability to detect politically aligned news articles vary between Democrats and Republicans?

This research question is addressed in terms of both user choice and evaluation. Through the latter, we extend the work of Seddik et al. [15], which also examined selective exposure and user choices in the context of a personalized news

platform. In terms of building on the RISE framework, we hypothesize that users will be more likely to detect politically aligned news articles if such articles are nudged, through an NRS, to be more salient and prominently featured on the front page of a news site. Formally, we expect for user choices:

H: Readers are more likely to detect politically aligned news articles rank-ordered prominently at the top of a news site, compared to readers that face a randomized order.

We also extend previous work by examining this hypothesis with two classification metrics: precision and recall. These are commonly used in recommender studies for the validation of user models [5]. As a first, we use them in the context of user perception of politically positioned news articles.

2 METHOD

2.1 Materials and Procedure

We designed an online experiment¹ in which participants were asked to select news articles that matched their political orientation. For instance, a participant identifying as a Democrat was expected to spot and select articles favorable to the Democratic party, and vice versa for Republicans. Our system is an adaptation of the infrastructure used in Seddik et al. [15].

We used an artificial news website featuring 50 news articles, see Figure 1. Each article was presented with an image and a headline. Participants could select an article by clicking within its border. Figure 1 shows how the articles were displayed in a grid system, each row containing three items, except for the first three rows. The first row displayed one news article prominently, while rows two and three displayed two news articles each, with one displayed more prominently. This design aimed to increase the exposure of articles presented at the top of the news website.

2.2 Dataset

The news articles were a subset of the TREC Washington Post corpus. This is a well-documented collection, comprising 728,626 news articles and blog posts from Jan 2012 through Dec 2020 [13]. It has been hailed for its comprehensive coverage and diversity, and was used in prior news recommender system research [6, 15, 18]. As we aimed to examine users' proficiency in identifying political articles that align with their preferred political ideologies, we derived them from a subset in a prior study by Seddik et al. [15]. This subset comprised 50 articles: 5 pro-Democratic, 5 pro-Republican, and 40 non-political filler articles, all of which were published on the Washington Post website between 2013 and 2016.²

2.3 Participants

We recruited a total of 200 participants through the crowdsourcing platform Prolific. Participants were US residents, who were fluent in English and a 98% approval rate on the platform. Each participant received GBP 1.05 for a 7-minute study. Unfortunately, only 160 participants met the predetermined criteria for inclusion in the analysis. Most notably, 22 participants reported to not identify as either Democrat or Republican, not even leaning.³

Out of the 160 participants ($M_{age} = 38.88$ years, $SD_{age} = 14.83$, $Min = 19$, $Max = 80$), 79.4% self-identified with varying degrees of allegiance to the Democratic Party, specifically as 'strong', 'not so strong', or 'leaning'. Among the participants that were omitted, 22 identified as an independent with no political preference.

¹This research adhered to the ethical guidelines of the Research council of Norway and the guidelines of the University of Bergen for scientific research. It was judged to pass without further extensive review, for it contained no misleading information, stress tasks, nor would it elicit extreme emotions.

²A link to a file containing the study's collected data can be viewed here: https://osf.io/pwja5/?view_only=67e4dc2d191746ba815acd0beae49dd9

³Other exclusion criteria were: a) selecting no news articles, b) invalid or incomplete post-experiment questionnaires, c) reporting to not having understood the instructions and facing technical issues.

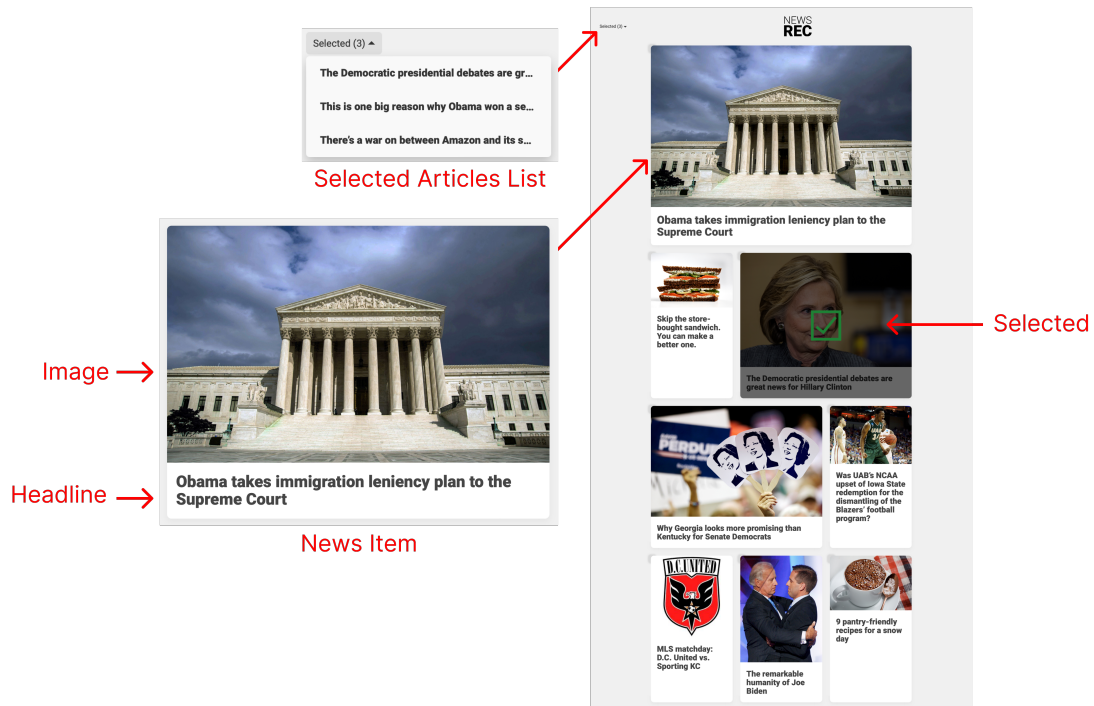


Fig. 1. The main page of the artificial news site used in the experiment.

2.4 Procedure

All participants were first presented with a short pre-questionnaire on demographics, political preferences and engagement, and news consumption habits. Subsequently, participants engaged in the task of identifying and selecting articles that aligned with their political preferences. Participants were explicitly asked to select at least one news article but were not informed about the total of correct news articles presented, which was 5. Upon completion, they proceeded to a post-questionnaire designed to assess choice difficulty, choice confidence, and gather feedback on technical issues or general reflections. These items were adapted from studies on food recommender systems [16], where choice confidence was an adaptation of choice satisfaction items

2.5 Research Design

We used a 2x2 between-subjects research design. On the one hand, users were deemed to be either Democrats or Republicans based on their self-reported party preference. On the other hand, news articles were rank-ordered towards a user's political orientation or presented randomly. This meant that, effectively, we employed one baseline (irrespective of political orientation) and two treatment conditions that were either tailored towards a Democrat's or Republican's political orientation.

We effectively used three different types of news recommender systems (NRSs). The baseline NRS, serving as the control condition, displayed articles in random order, providing a non-manipulated reference point for comparison with the knowledge-based systems. Meanwhile, the knowledge-based NRSs aimed to increase selective exposure to political news.

The study also featured two treatment NRSs. The first knowledge-based NRS (REC-1) was designed to amplify exposure to pro-Democratic news by featuring such articles predominantly at the top of the website, interspersed with filler articles. Three pro-Democratic articles were also prominently displayed in larger sizes within the first three rows. Conversely, pro-Republican articles were placed towards the bottom, mixed with filler articles. Similarly, the second knowledge-based NRS (REC-2) focused on pro-Republican news. Like REC-1, it featured pro-Republican articles at the top, with three articles enlarged for prominence. Pro-Democratic articles were positioned at the bottom, mixed with filler articles.

All three NRSs displayed the same 50 articles but in different orders. Participants were randomly assigned to one of the three NRSs based on their political preference. Those identifying as ‘not so strong’, ‘strong’, or ‘somewhat aligned’ with the Democratic party were randomly placed in either the baseline or REC-1. Similarly, participants leaning towards the Republican party were placed in either the baseline or REC-2.

- D-BLINE: Pro-Democratic participants who are assigned to the baseline.
- R-BLINE: Pro-Republican participants who are assigned to the baseline.
- D-Increase: Pro-Democratic participants who are assigned to REC-1, aiming to increase their selective exposure to pro-Democratic news.
- R-Increase: Pro-Republican participants who are assigned to REC-2, aiming to increase their selective exposure to pro-Republican news.

2.6 Measures

We assessed the predictive performance of users through standard F-score metrics [20]. Since users were not given a specific number of news articles to select, while there were only 5 correct (out of 50 possible) articles for each user, precision and recall were used to consider the study a user-side classification problem. Users selecting a news article that corresponded to their political ideology was deemed to be a true positive, while *not* selecting one a false negative. Conversely, selecting a news article from the opposite affiliation or a filler article was a false positive, while those not selected were regarded as true negatives. To this end, precision and recall were operationalized as follows, where mostly precision would punish users who would make many guesses:

$$\text{Precision} = \frac{\text{Selected Correct Articles}}{\text{Selected Articles}} \quad (1)$$

$$\text{Recall} = \frac{\text{Selected Correct Articles}}{\text{All Correct Articles}} \quad (2)$$

We further inquired about the user’s evaluation of the task (5-point Likert Scale). The questionnaire items are outlined in Table 1, which were used in a principal component factor analysis. On the one hand, we inquired about each user’s experience of difficulties in identifying the correct articles. This was labeled as choice difficulty, taking items from previous studies in related domains as examples [17, 19]. Furthermore, we inquired about users’ confidence levels of making these choices, labeled as choice confidence, which were related to items used in [9].

Table 1. Results of the Principal Component Analysis. Omitted items are denoted in gray. Two evaluation components or aspects were retained, after performing promax rotation on the items.

Aspect	Item	Loading
Choice	The task of choosing politically-aligned articles was overwhelming	0.753
	I changed my mind several times before choosing the news articles	0.717
Difficulty	The surrounding items (other articles, ads, etc.) made it difficult to focus on the articles	0.742
	The size and prominence of articles on the screen influenced my selection of news stories	0.747
Choice	I think I chose news articles that align with my political preferences	0.642
	I could easily discern the political leaning of articles from their titles and images	0.807
Confidence	It was challenging to select articles solely based on titles and images	
	I was confident in my ability to identify articles that align with my political views	0.900

3 RESULTS

We examined to what extent users were able to identify news articles that corresponded to their political affiliation. On average, users selected 7.55 news articles ($SD = 5.86$), while 20% of users selected more than 10 articles, with one user selecting 38 articles. To examine [H] (i.e., random vs top rank-ordering) differences across democratic and republican political affiliations, we performed ANOVA analyses on precision and recall. We used dummy variables (0/1) for our predictors: political affiliation (i.e., ‘democrat’) and ranking (i.e., ‘matched’); the latter to address [H].

Our between-subjects ANOVAs revealed mixed results for our [RQ]. For precision, we found that users with a democratic affiliation ($M = 0.58$, $SD = 0.35$) scored significantly higher than Republicans ($M = 0.40$, $SD = 0.32$): $F(1, 156) = 6.15$, $p < 0.05$. Similarly, Democrats ($M = 0.68$, $SD = 0.32$) also scored higher than Republicans ($M = 0.38$, $SD = 0.29$) on recall: $F(1, 156) = 21.80$, $p < 0.001$. This suggested either that Democrats, compared to Republicans were more capable of identifying their ‘own’ news articles, or that these news articles had a clearer political earmark. In contrast, addressing [H], we found no significant effects of the rank-ordering (i.e., matching the political affiliation or random) on precision: $F(1, 156) = 0.03$, $p = 0.87$, nor for recall: $F(1, 156) = 0.05$, $p = 0.83$. In addition, no interaction effects between political affiliation and the rank-ordering were observed for both precision ($p = 0.89$) and recall ($p = 0.41$). Figures 2 and 3 suggest that, when given a task to identify articles, the rank-ordering along with the prominence of the top-ranked articles, did not influence the success rate of users.

To further examine this argument, we examined users’ reported choice confidence and difficulty levels (RQ). A between-subjects ANOVA revealed results in line with precision and recall, as Democrats reported a higher level of choice confidence than Republicans: $F(1, 156) = 6.81$, $p < 0.01$. Again, we observed no effects based on the used rank-ordering, nor an interaction effect between political orientation and the rank-ordering (both: $p > 0.05$). This suggested that users’ expectations corresponded to their actual performance. In contrast, a fourth ANOVA model revealed that choice difficulty did not vary across conditions: $F(3, 156) = 0.49$, $p = 0.69$. This suggested that even ranking all ‘correct’ news articles at the top did not reduce the difficulties of selecting them.

4 DISCUSSION & FUTURE WORK

We have observed notable differences in precision and recall of selected news articles between Democrats and Republicans, with Democrats outperforming Republicans. While recall captures whether users identify all correct articles, precision also weights the false positives, which were quite numerous among some users. The results raise questions about whether these disparities are rooted in personal characteristics or if they reflect differences in the content itself.

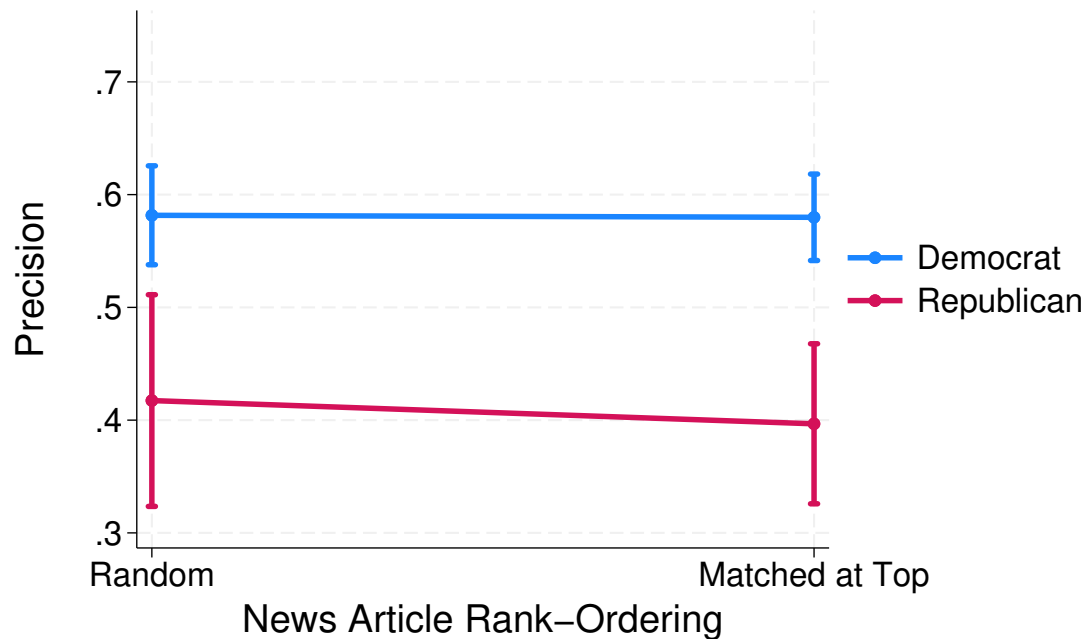


Fig. 2. Precision of articles selected by users. It indicates how many of the selected articles actually matched a user's political orientation. Error bars are 1 SE.

On the one hand, Democrats could generally be better in detecting articles that align with their political preferences. On the other hand, Democratic-aligned content could be more overt in its political messaging, resulting in an easier task for the Democrats. This could have been exacerbated by journalistic choices of the used outlet – but a systematic comparison between different outlets was beyond the scope of the current work. To gain further insight into this possibility, one could repeat the experiment using symmetrically designed content to more accurately isolate the impact of personal versus content characteristics.

Interestingly, we have found no significant differences between participants presented with articles in random versus re-ranked order, across multiple metrics. Thus, we have not found support for our main hypothesis [H]. This challenges the intuitive notion that users would more readily identify articles aligned with their views if these were more prominently displayed. This would particularly apply to choice difficulty, for the news articles could have been located easily. This observation suggests a distinction in user behavior: while recommender systems in a typical browsing scenario might lead users to engage more with like-minded articles at the top, the act of actively spotting articles seems to alter this behavior. In task-oriented settings, users may not exhibit the same preference patterns as observed in casual browsing. The difference in passive consumption and active searching might explain why the experiment yielded no difference between users who were presented with articles in random versus re-ranked order.

Regarding the user's evaluation, we have found the user's choice confidence levels to correspond to the results for precision and recall. This may suggest that many users share a set of cues they use to identify politically motivated

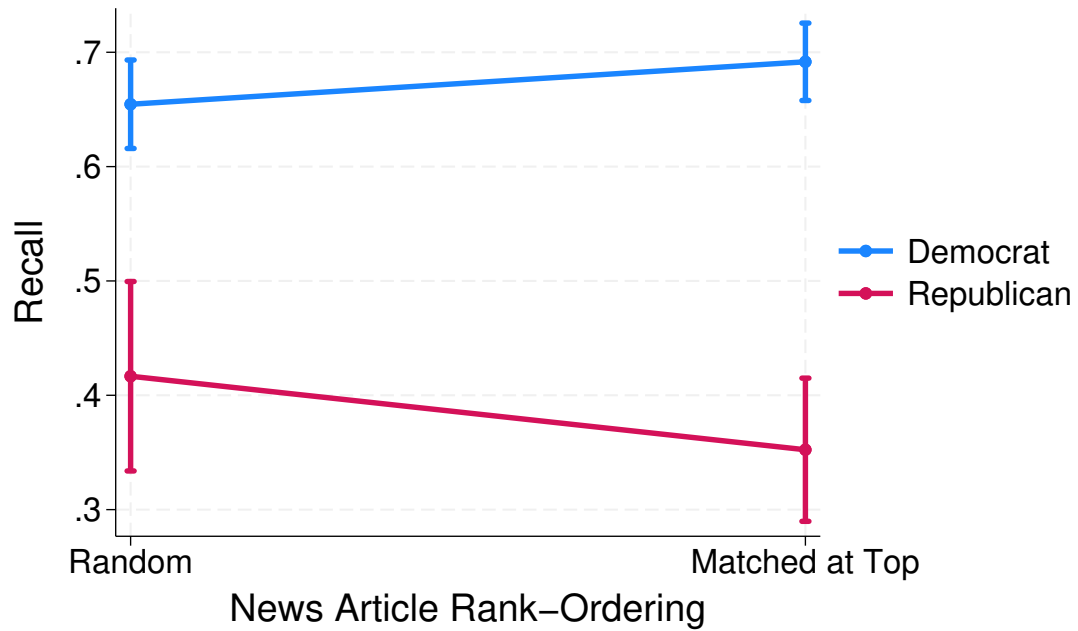


Fig. 3. Recall of correct articles selected by users. It indicates to what extent all of the relevant news articles were selected. Error bars are 1 SE.

selective exposure or framing in news, for the actual performance corresponds to user perceptions. Which cues are actually used should be revealed using more in-depth methods, either through survey questions or qualitative methods.

This work is set up for studies where the political affiliation is included in the user model. As is shown by our study how a match between the content and the user can lead to a variety of choices, further personalization (either in an ‘echo chamber’ way or not) could lead to stronger results regarding selective exposure [3].

An important limitation of this work is the relatively small sample of Republican users. During recruitment, we did not set any restrictions but ended up with a relatively left-leaning sample, which has made comparative analyses more difficult. Nonetheless, we still observed significant differences.

Following this work, we seek to present a more realistic scenario in a follow-up study. A limitation of the results is that the recommender system used in this study has merely re-ranked news articles, tapping into a relatively small corpus of news articles. We propose to examine this further by presenting interventions on selective exposure or politically microtargeting users (cf. [23]) with news articles from current affairs. Recent studies have successfully implemented such NRSs which import the latest news articles from multiple news sources [4, 12]. Moreover, future work should also examine countries with multi-party systems, in which the notion of political alignment is more complex – both from a recommendation perspective as well as from a user perspective.

ACKNOWLEDGMENTS

This work was supported by the NEWSREC Project, with project number 324835. It also received support by the Research Council of Norway with funding to MediaFutures: Research Centre for Responsible Media Technology and Innovation, through the Centre for Research-based Innovation scheme, project number 309339.

REFERENCES

- [1] Mitch Dobbs, Joseph DeGutis, Jorge Morales, Kenneth Joseph, and Briony Swire-Thompson. 2023. Democrats are better than Republicans at discerning true and false news but do not have better metacognitive awareness. *Communications Psychology* 1, 1 (2023), 46.
- [2] Ivan Dylko, Igor Dolgov, William Hoffman, Nicholas Eckhart, Maria Molina, and Omar Aaziz. 2017. The dark side of technology: An experimental investigation of the influence of customizability technology on online political selective exposure. *Computers in Human Behavior* 73 (2017), 181–190.
- [3] R Kelly Garrett. 2009. Echo chambers online?: Politically motivated selective exposure among Internet news users. *Journal of computer-mediated communication* 14, 2 (2009), 265–285.
- [4] Lucien Heitz, Juliane A Lischka, Rana Abdullah, Laura Laugwitz, Hendrik Meyer, and Abraham Bernstein. 2023. Deliberative Diversity for News Recommendations: Operationalization and Experimental User Study. In *Proceedings of the 17th ACM Conference on Recommender Systems*. 813–819.
- [5] Dietmar Jannach, Markus Zanker, Alexander Felfernig, and Gerhard Friedrich. 2010. *Recommender systems: an introduction*. Cambridge University Press.
- [6] Jia Hua Jeng, Alain Starke, Christopher Trattner, et al. 2023. Towards Attitudinal Change in News Recommender Systems: A Pilot Study on Climate Change. (2023).
- [7] Daniel Kahneman. 2011. *Thinking, fast and slow*. macmillan.
- [8] Mozghan Karimi, Dietmar Jannach, and Michael Jugovac. 2018. News recommender systems—Survey and roads ahead. *Information Processing & Management* 54, 6 (2018), 1203–1227.
- [9] Bart P Knijnenburg, Martijn C Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. 2012. Explaining the user experience of recommender systems. *User modeling and user-adapted interaction* 22 (2012), 441–504.
- [10] Erik Knudsen. 2023. Modeling news recommender systems’ conditional effects on selective exposure: evidence from two online experiments. *Journal of Communication* 73, 2 (2023), 138–149.
- [11] Milton Lodge and Charles S Taber. 2013. *The rationalizing voter*. Cambridge University Press.
- [12] Felicia Loecherbach and Damian Trilling. 2020. 3bij3—Developing a framework for researching recommender systems and their effects. *Computational Communication Research* 2, 1 (2020), 53–79.
- [13] National Institute of Standards and Technology (NIST). 2020. TREC Washington Post Corpus. <https://trec.nist.gov/data/wapost/>
- [14] Shaina Raza and Chen Ding. 2022. News recommender system: a review of recent progress, challenges, and opportunities. *Artificial Intelligence Review* (2022), 1–52.
- [15] Khadiga MA Seddik, Erik Knudsen, Damian Trilling, and Christoph Trattner. 2023. Understanding How News Recommender Systems Influence Selective Exposure. In *Proceedings of the International Workshop on Behavior Change and Persuasive Recommender Systems co-located with the 17th ACM Conference on Recommender Systems (RecSys 2023)*.
- [16] Alain Starke, Edis Asotic, and Christoph Trattner. 2021. “serving each user”: Supporting different eating goals through a multi-list recommender interface. In *Proceedings of the 15th ACM Conference on Recommender Systems*. 124–132.
- [17] Alain Starke, Martijn Willemsen, and Chris Snijders. 2021. Promoting energy-efficient behavior by depicting social norms in a recommender interface. *ACM Transactions on Interactive Intelligent Systems (TiIS)* 11, 3-4 (2021), 1–32.
- [18] Alain D Starke, Sebastian Øverhaug Larsen, and Christoph Trattner. 2021. Predicting feature-based similarity in the news domain using human judgments. In *Proceedings of the 9th International Workshop on News Recommendation and Analytics (INRA 2021) co-located with 15th ACM Conference on Recommender Systems (RecSys 2021)*.
- [19] Alain Dominique Starke, Justyna Sedkowska, Mihir Chouhan, and Bruce Ferwerda. 2022. Examining Choice Overload across Single-list and Multi-list User Interfaces. In *9th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems (InRS’22)*.
- [20] Cornelis Joost Van Rijsbergen. 1977. A theoretical basis for the use of co-occurrence data in information retrieval. *Journal of documentation* 33, 2 (1977), 106–119.
- [21] Sanne Vrijenhoek, Gabriel Bénédict, Mateo Gutierrez Granada, and Daan Odijk. 2023. RADio*—An Introduction to Measuring Normative Diversity in News Recommendations. *ACM Transactions on Recommender Systems* (2023).
- [22] Sanne Vrijenhoek, Lien Michiels, Johannes Kruse, Alain Starke, Nava Tintarev, and Jordi Viader Guerrero. 2023. Normalize: The first workshop on normative design and evaluation of recommender systems. In *Proceedings of the 17th ACM Conference on Recommender Systems*. 1252–1254.
- [23] Frederik Zuiderveen Borgesius, Judith Möller, Sanne Kruikemeier, Ronan Ó Fathaigh, Kristina Irion, Tom Dobber, Balazs Bodo, and Claes H de Vreese. 2018. Online political microtargeting: Promises and threats for democracy. *Utrecht Law Review* 14, 1 (2018), 82–96.