Deliberative Diversity for News Recommendations: Operationalization and Experimental User Study

Lucien Heitz
heitz@ifi.uzh.ch
Department of Informatics and Digital Society Initiative, University of Zurich
Zurich, Switzerland

Juliane A. Lischka
juliane.lischka@uni-hamburg.de
Universität Hamburg
Hamburg, Germany

Rana Abdullah
rana.abdullah@uni-hamburg.de
Universität Hamburg
Hamburg, Germany

Laura Laugwitz
laura.laugwitz@uni-hamburg.de
Universität Hamburg
Hamburg, Germany

Hendrik Meyer
hendrik.meyer-1@uni-hamburg.de
Universität Hamburg
Hamburg, Germany

Abraham Bernstein
bernstein@ifi.uzh.ch
Department of Informatics, University of Zurich
Zurich, Switzerland

ABSTRACT

News recommender systems are an increasingly popular field of study that attracts a growing interdisciplinary research community. As these systems play an essential role in our daily lives, the mechanisms behind their curation processes are under scrutiny. In the area of personalized news, many platforms make design choices driven by economic incentives. In contrast to such systems that optimize for financial gain, there can be norm-driven diversity systems that prioritize normative and democratic goals. However, their impact on users in terms of inducing behavioral change or influencing knowledge is still understudied. In this paper, we contribute to the field of news recommender system design by conducting a user study that examines the impact of these normative approaches. We a.) operationalize the notion of a deliberative public sphere for news recommendations, show b.) the impact on news usage, and c.) the influence on political knowledge, attitudes and voting behavior. We find that exposure to small parties is associated with an increase in knowledge about their candidates and that intensive news consumption about a party can change the direction of attitudes of readers towards the issues of the party.

CCS CONCEPTS

• **Information systems** → **Recommender systems**; **Personalization**.

KEYWORDS

deliberative diversity, journalism, recommender system

ACM Reference Format:
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 *Seventeenth ACM Conference on Recommender Systems (RecSys '23), September 18–22, 2023, Singapore, Singapore.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3604915.3608834

RecSys '23, September 18–22, 2023, Singapore, Singapore

ACM ISBN 979-8-4007-0241-9/23/09.
https://doi.org/10.1145/3604915.3608834

1 INTRODUCTION

News recommender systems (NRS) curate content by making automated content suggestions to users [15]. This algorithmic content curation serves as a gatekeeper in online public spheres that determines what is seen and by whom. In the political domain, political news can shape public opinion by influencing individuals' knowledge, beliefs, and understanding of political issues and actors [19]. These attitudinal effects can translate into behavioral effects, such as political participation and voting decisions [22].

While real-world political news environments often provide a visibility bonus to the incumbent party [3], voters need to be extensively informed about each party's candidates and campaign issues to make an informed voting decision in multi-party political systems [4]. According to prototypical models of the public sphere [6, 12], proportional, impartial, or majority visibility information environments can be constructed by NRSs on a long-term basis. By affecting the visibility of parties in this way, NRSs may increase or balance existing news biases, as suggested by the reinforcement-orientation model or the concept of selective exposure [7, 17]. In this paper, we therefore conduct a user study that operationalizes the model of deliberative diversity [12], providing news articles from different sources and perspectives, to focus on party visibility for raising awareness of minority parties.

Against this background, we ask the following research question: *How does varying party visibility in news recommendations affect users' party preferences?* Our contributions include (1) the operationalization of deliberative diversity as a model for NRSs for varying party visibility; (2) a user study on the impact of NRSs on political knowledge; and (3) the impact of NRSs on voting behavior. We approach this question from a *social science perspective.*

We first provide an overview of the related work on the effects of news exposure and NRSs. We then outline the methods applied in our user study, implementing an algorithm for deliberative diversity for contribution (1). This is followed by the outline of the field study, its results, and the discussion of the insights for contribution (2)

and (3). We conclude with our recommendations on media policy for NRS designers.

2 RELATED WORK

In the realm of political news, visibility refers to the relative amount of coverage political parties and their actors receive in the news [5]. Visibility thereby indicates political importance and political power to audiences [20]. News coverage during election campaigns provides the ruling political party with a visibility bonus [3, 5, 14], suggesting an incumbency bias in news supply. Biased party visibility can affect party preferences [14], especially for volatile voters [8] and in interaction with news tonality [5]. Beyond individual news usage, the bias of the overall news information environment is found to affect voting decisions [14]. Exposing U.S. Republican voters to a "liberal" information environment shows short-term effects on candidate and party evaluations [2]. Likewise, being exposed to a *diverse* news environment has a positive effect on tolerance towards opposing views for politically conservative users [11]. And yet, users avoid a highly diverse news environment, suggesting that there is an optimal level of news diversity [29].

While previous work focused on adopting diversity as metrics to measure NRS output [31, 32], we propose to operationalize deliberative diversity (cf. democratic models in [12]) as an algorithm to generate recommendations instead. By proposing such a diversity model, we plan to create a *normative* NRS, putting the needs of the users first[10, 29]. We understand normative as incorporating traditional professional values of journalism [13] and democratic principles [12] to both guide and empower users by offering a diversity of narratives and content [26, p. 61].

Research on the effects of NRS-created news environments reveals mixed results. When exposed to like-minded or opposing arguments for, e.g., COVID-19 restrictions by a collaborative filtering or a random NRS, attitudes towards restrictions and affective polarization did not differ in an experimental design [23]. In contrast, NRSs that recommend politically diversity-optimized news instead of accuracy-optimized news are related to a higher tolerance for opposing views [11]. Likewise, it was shown that spending more time with a content-based NRS accentuating negative sentiment increases affective polarization and that an NRS with balanced sentiment is related to decreasing ideological polarization the longer participants used the NRS [18].

NRS effects can be amplified or diminished based on individuals' pre-existing attitudinal, cognitive, social, and socio-demographic characteristics. Notably, effect sizes of NRS-created news environments on user attitudes are small to moderate, which is largely in line with media effect studies in general [30]. However, the recent literature suggests that too little is known about the effects of NRSs on political attitudes, behavior, and studies integrating the output and outcome stages [21]. Further research is needed to better understand the impact of recommender systems [26, p. 73], particularly in the domain of NRS [1]. By directly manipulating the news supply and real-time monitoring of news consumption, our findings can contribute towards an improved understanding of the impact on NRSs.

Table 1: Study outline of how news was used in the NRS and dimensions measured in the user study.

Input		News Recommender Pipeline			User Study			
					Supply		Demand	Effects
Real-world news	→ Data scraping and augmentation	→ Named-entity recognition and political scoring	→ NRS 1: News as published NRS 2: Deliberative diversity	→	Control group (n = 65) Treatment group (n = 78)	→	Political news reading intensity	Political knowledge; campaign issue agreement; voting behavior

3 NEWS RECOMMENDER PIPELINE

Our study combines NRS design with social science research methods (see pipeline overview for the study in Table 1). We implemented a NRS that gathers news from a variety of news outlets, processes them to identify (political) party representation in the articles, and recommends them based on a diversity model. We used the research infrastructure and smartphone news app of [11] for conducting an experimental field study.[1] It supplies political news alongside other news articles, mirroring a traditional news app. To implement our two recommendation algorithms, we leveraged the app's news algorithm customization capability. This NRS pipeline of the study follows three basic steps: 1.) scraping news articles together with pre-processing and data augmentation, 2.) named-entity recognition and assigning political scores to news articles, and 3.) calculating a recommendation list for each user based on deliberative diversity.

The news items are recommended by two different algorithms. A *baseline approach* generated a chronological recommendation list (sorted from newest to oldest news article) that provides recommendations to the control group. For the treatment group, we implement a *deliberative condition* providing impartial (i.e., similar) visibility of parties and their actors. This algorithm for deliberative diversity is drawn from the typology of democratic recommenders by [12]. The following section will address each step of the recommendation pipeline in greater detail.

3.1 Scraping Procedure and Data Augmentation

The user study took place between September 19, 2022, and October 9, 2022, in Lower Saxony, a federal state in Germany. During the experiment, news featured in the app was aggregated from six different German newspapers (featuring four regional and two national outlets): Braunschweiger Zeitung, Die Welt, Neue Osnabrücker Zeitung, Nordwest-Zeitung, Hannoversche Allgemeine Zeitung, and taz - die tageszeitung. All articles were accessed through the API of the WISO research database.[2] The automated scraping took place once per day, at 7:00 a.m. from Monday throughout Saturday. Since there were no updates on Sunday, the app kept displaying the news items scraped on the previous day. News items obtained via the WISO database are articles that appear in the printed newspaper of that day.

A total of twelve categories were featured in the app: economy, education, elections, finance, front page, national/international, opinions, panorama, people, politics, science, and sports. We used the categories and subcategory meta-information that we scraped together with the news article to assign all items to one of our predefined categories automatically. Each of the articles features in the app had a minimum length of 150 words.

[1] Official website of the Informfully news app: https://www.informfully.ch
[2] Official website of the WISO research database: https://www.wiso-net.de

The pre-processing steps included removal of duplicate articles. These can occur when multiple news outlets feature the same press agency report. As this has a negative impact on user experience, the system kept only one version and discarded the rest. All news items were white labeled (removing any mentions of outlet name, author, or source, replacing any proper names with a generic placeholder) to avoid any outlet bias from the user side. In total, 2,191 news articles were stored in the database. For each article, we scraped the title, lead, main text, source, category, and publishing date. We added augmented information on the reading duration and the political entities (occurrence counts of parties and politicians, see Subsection 3.2). To present all articles uniformly, we removed any pictures from the articles and generated a generic thumbnail (rendering the spelled-out word of the category of an article as an image).

3.2 Named-Entity Recognition and Political Scoring

We created a list containing the names of all parties, their abbreviations, spelling variants, synonyms, grammatical variations, and names of each politician that took part in the election. In addition to the politicians that registered for the election on the official voting lists, we scraped all websites of regional and national parliaments in Germany to get a list of current members of parliament and linked them to their respective parties. In total, 58 party names and 2,995 politicians were added to our collection. When performing NER, we first checked if an article contains any political party. If there was no mention, the article was flagged as "non-political," and the politician list was skipped. If a party appeared, the article was flagged as "political," and the system then checked if a politician was mentioned. The politician's first and last names had to appear at least once (matches on any middle name are optional).

Of the 2,191 news articles, 1,621 were flagged as political articles and 261 of these mentioned at least one politician. In total, the political articles contained 11,199 references to either parties or politicians. Any article containing a politician from Lower Saxony received an additional flag, marking it as local political news potentially relevant for the elections. The information gathered from political entities was used to calculate a *Political Reference Score* (PRS) for each news article (stored with each article as additional attributed). PRS is an array of length n, where n equals the number of parties in the given political landscape. In the scope of this user study, the PRS included occurrence counts for entities from the following parties: Social Democrats (SPD), Greens (Die Grünen), Liberals (FDP), Conservatives (CDU), Left Party (Die Linke), Right-Wing Populists (AfD), and Others.[3]

For the elections in Lower Saxony, PRS contains 7 different counts of political entities. Each count is in a tuple $(name, count)$, where *name* is the name of a political party and *count* is the number of occurrences of this party within the given news article. For example, the *PRS* for an article that mentions the Greens 8 times and the Liberals 4 times would be: *PRS = [(Social Democrats,* 0*), (Greens,* 8*), (Liberals,* 4*), ...]*. When detecting the name of a politician, the score of their affiliated party is increased.

3.3 Operationalizing Deliberative Diversity

After scraping and pre-processing the news articles, the next step in the NRS pipeline is creating the user recommendations lists. For this purpose, we present our deliberative model Algorithm 1, our contribution towards a first operationalization of deliberative diversity (cf. [12]). The goal of deliberative diversity is to achieve a given normative, impartial party visibility (PV) within a news feed. The personalization of the news feeds was done on a group level, to both lower the execution time of the NRS and to allow for a meaningful comparison within and between groups.

Algorithm 1 creates two recommendation lists in the shape of news feeds, one for participants in the control group and one for the participants in the treatment group. To achieve an impartial PV, the ordering of items needed to be considered (as articles on more prominent positions, e.g., on the top of the reading list, receive more attention). And to provide a consistent ratio of news article types, we subdivided the news feeds into smaller slices of six articles (containing three political and three non-political articles). The resulting feed of political and non-political articles is what we refer to as ***F***eed of ***U***ser ***N***ews (FUN).

Every FUN slice is composed of three articles flagged as political (top three articles) and three non-political articles (bottom three articles).[4] As there were daily fluctuations in the number of articles scraped, FUN consisted of eight to twelve slices daily. Algorithm 1 shows the details of FUN for calculating the recommendations for the treatment and control group. The input consists of three different datasets: (1) a set of non-political news, (2) a set of local political news, and (3) a set of (national) political news. The three sets are mutually exclusive and the ordering of items in each set is randomized. For calculating the recommendation list of the control group, set (2) consists of local political news on *majority* parties and set (3) consists of (national) political news on *majority* parties.

For calculating the recommendations for the treatment group, both sets (2) and (3) feature news on *minority* parties instead. SPD, CDU, Greens, and Liberals are defined as majority parties; the remaining ones are minority parties. Note that as news items often mentioned multiple parties, both groups also received news items typically covered in the other condition: i.e., the control group also read about minority parties when they were mentioned in news stories covering majority parties. Respectively the treatment group also read about majority parties when they were co-mentioned in the news about minority parties. The number of political and non-political articles per FUN slice is provided as input to Algorithm 1; the resulting output is a ranked list of item recommendations for the groups.

The rank of an article is determined by its index within the array storing the recommendations. Special consideration is given to local news articles (see Algorithm 1, Line 8). Local articles will get added first to the recommendation list. By controlling for the rank in this

[3] The total of parties added to the 'Others' category was 15. For the complete overview of the participating parties and politicians, please see the official election publication (available only in German): https://landeswahlleiterin.niedersachsen.de/startseite/wahlen/landtagswahl/landtagswahl_2022/endgultiges_amtliches_ergebnis/wahl-zum-19-niedersachsischen-landtag-am-9-oktober-2022-endgultiges-amtliches-ergebnis-216438.html

[4] While our goal was to achieve deliberative diversity by promoting political viewpoints, we complemented the news feed with non-political articles, as this has a positive effect on user engagement and has shown to strengthen social cohesion [28].

way, local news stories appear higher in the recommendation list, have greater prominence, and are more likely to be selected by users [25].[5]

The small parties (Liberals, Left, Extreme right, other parties) had significantly higher visibility shares in the treatment group FUN (29%, 10%, 12%, 4%) than in the control group FUN (13%, 5%, 5%, 2%) ($Z = 9.072, 5.115, 5.525, 3.192, p < .001, n = 1,621$). The large parties (Social Democrats, Conservatives, Greens) had significantly lower visibility in the treatment (22%, 15%, 8%) than in the control FUN (37%, 26%, 12%) ($Z = -7.117, -5.9123, -3.1043, p < .001, n = 1,621$).

Algorithm 1: Deliberative diversity algorithm based on FUN slices for item re-ranking.

Input : $newsLocal, newsPol, newsNonPol, sliceCount_{political}, sliceCount_{non-political}$
Output: $recList$

```
recList = []
countPol = length(newsLocal) + length(newsPol)
countNPol = length(newsNonPol)
countMax = countNPol + countPol
while length(recList) < countMax do
    i = 0
    while i < sliceCount_political do
        if length(newsLocal) > 0 then
            recList.append(newsLocal.PopFirstElement())
        else if length(newsPol) > 0 then
            recList.append(newsPol.PopFirstElement())
        i = i + 1
    j = 0
    while (j < sliceCount_non-political) do
        if length(newsNonPol) > 0 then
            recList.append(newsNonPol.PopFirstElement())
        j = j + 1
return recList
```

4 EXPERIMENTAL DESIGN

We conduct a field experiment prior to federal state elections in Lower Saxony, Germany (elections on October 9, 2022). While voter decision-making has become more volatile in Germany [27], the two centrist parties (Social Democrats and Conservatives) account for the largest parliament share in the selected state compared to any other German state. Lower Saxony thus represents a stable multi-party system with a strong centrist majority. The case can therefore be regarded as a conservative research environment for testing visibility effects of party diversity with the reasoning that "if it happens there, it will happen anywhere" [24, p. 236].

[5]Algorithm 1 will create group-level recommendations (same set of articles for each user within a given group). Using the historical interaction data recorded during the experiment, however, our algorithm would also allow for personalization at the level of the individual. For example, the selection of political entities can be made dependent on the place of residency of a user, or the pool of articles to recommend can be limited to articles published within a given time window of the user opening the app (as opposed to calculating one list in the morning that will not change during the day).

$N = 143$ participants were recruited from an ISO 26362:2009 certified online panel to use the news app during the state election campaigns, i.e., for three weeks before election day. Participants were informed about the goals and potential risks of the study. Participants were randomly distributed to the control group ($n = 65$) or diversity-NRS treatment condition ($n = 78$). In an ex-ante survey, participants were asked about their party preferences, governance ability evaluation for each party, and attitudes toward campaign issues per party [5, 8]. These items were measured on an 11-point Likert scale from −5 (not at all) to +5 (very much). Participants were on average 44 years old ($SD = 12.49, Min = 19, Max = 73$), 51% were female, two thirds (67%) indicated vocational training or lower, and one third (33%) a university degree as highest level of education. All participants owned a smart phone, lived in Lower Saxony, and were eligible to vote in the upcoming state elections.

A post hoc survey after the election day measured political attitudes on three dimensions. For the cognitive dimension of political knowledge, participants were asked to ascribe the party affiliation to the name of the top candidate per party [16]. For the affective dimension, we measured party preferences, governance ability evaluation for each party, and attitudes toward campaign issues identical to the ex-ante survey [5, 8]. For the behavioral dimension, we asked participants to indicate whether they voted and which party they voted for. Control variables were the usage of other news sources and demographics. Finally, we tracked the reading behavior of each participant with a political-news-reading intensity score (PIS) accounting for the number of PRS articles, reading duration (RD), and scrolling depth (SD) in the form $PIS = PRS * (RD * SD)$.

5 RESULTS

To analyze the user experiment results, we compare the demand side with political news usage as well as the effect side with cognitive, affective, and behavioral dimensions of political attitudes between the control and diversity NRS groups. We list separately the results on news usage and the observed effects on the users.

5.1 Political News Usage

Overall, users in the diversity-NRS treatment condition read political news more intensively ($PIS\ M = 302.52, SD = 244.04, n = 75$) than users in the control condition ($PIS\ M = 216.46, SD = 212.69, n = 63$) ($t(136) = -2.187, p = .015$). Specifically, users in the treatment condition read news about the Social Democrats, Conservatives, Left, and Far Right similarly and about the Greens, Liberals, and other parties more intensively than users in the control condition ($t(136) < -1.96, p < .05$). Users in the treatment condition used news about the Social Democrats and the Liberals most intensely ($PIS\ M = 91.31, SD = 75.05; M = 71.06, SD = 68.55$), followed by news about the Far-Right and Conservative party ($PIS\ M = 45.75, SD = 63.11; M = 42.26, SD = 37.07$).

5.2 Effects on Political Knowledge

Regarding cognitive effects, there was no difference between the two groups in the proportion of participants who indicated to not know the party affiliation of the candidates. However, there was a difference in the proportion of participants who correctly identified the party compared to the proportion of participants who

incorrectly identified the party of a candidate. Participants in the treatment condition correctly identified the Liberal and Far-Right candidates more often (98%; 95%, $n = 39$) than the control group (82%; 78%, $n = 32$) ($Z > 1.96, p < .05$). That is, the small party condition is associated with increased knowledge about the small party candidates.

To investigate the effects of party news usage, knowledge about the candidate, and preferences towards the party on the affective ex-post agreement with the campaign issue between groups, we conducted univariate ANOVAs. The model for the Liberal party campaign issue (no state support for companies) was significant, $F(7, 127) = 3.742, p = 0.001$, with a corrected R^2 value of 0.125, indicating that the included factors explain 12.5% of the variance in the dependent variable. Significant parameter estimates were found for the liberal party preference ($B = .321, SE = 0.103, t = 3.118, p = .002$, partial $\eta^2 = .071$) and the interaction between the experimental conditions and news usage of the liberal party (when treatment condition: $B = -.020, SE = 0.009, t = -2.082, p = 0.039$, partial $\eta^2 = 0.033$). The other parameter estimates were not statistically significant. More usage of news about the Liberals in the treatment condition is related to a greater disagreement with their campaign issues. Agreement to the campaign issue of the Far-Right party (minimize financial support for immigrants) was also only related to the Far-Right party preference ($B = .316, SE = .130, t = 2.436, p = .016$) with a partial $\eta^2 = 0.045$, indicating a small effect size. That is, group differences in news usage or candidate knowledge are not found to be related to the level of agreement with the Far-Right campaign issues.

5.3 Effects on Voting Behavior

Participants did not differ between groups in their voting participation. In the treatment (control) group, 93.6% (89.2%) reported having voted, which is a similar share according to a Pearson chi-square test ($\chi^2(1) = .876, n.s., n = 143$). Participants, however, differed between groups in their reported voting decision for the incumbent party (Social Democrats). More participants in the treatment group reported voting for the Social Democrats (28.2%, $n = 22$) than in the control group (12.3%, $n = 8$) ($Z = -1.9354, p = .053$). Voting decisions for the remaining parties did not differ between groups. The Social-Democrat voters in the treatment group used political news about all parties significantly more intense ($M = 413.5, SD = 263.5, n = 22$) than Social-Democrat voters in the control group ($M = 142.0, SD = 103.3, n = 8$) ($t(28) = -2.810, p = .009$).

To investigate a possible effect of news usage on political behavior between groups, we conducted a logistic regression analysis. We included a dummy variable representing the experimental groups, ex-ante and ex-post party preferences and governance ability evaluations, news usage, and candidate knowledge of the Social Democrats to predict the likelihood of voting for Social Democrats or not ($n = 143$). The model was found to be a good fit for the data, as indicated by the Hosmer-Lemeshow test ($p = .954$). Only ex-ante party preference ($B = .909, SE = 0.316, p = .004$) and ex-post party preference for the Social Democrats ($B = .685, SE = .323, p = .034$) were significant predictors of voting for the Social Democrats. Hence, the diversity-NRS treatment condition cannot be ascribed to changing voting behavior.

6 DISCUSSION

Our diversity NRS exposed participants to significantly more minority party news. Participants in the diversity NRS condition used political news more intensively, which contrasts with previous research suggesting that users avoid highly diverse news [29]. More intensive news usage translates into greater candidate knowledge for two minority parties, but usage and knowledge are not positively related to campaign issue agreement and unrelated to voting behavior. While diverse news exposure may enhance awareness and knowledge about political candidates, it might not necessarily sway people's decisions regarding policy issues or candidate selection. The capacity of political news to impact party preferences and eventually political actions is low, as suggested previously [5]. For a more volatile electorate, however, effects may become more pronounced, as individuals may be more open to changing their party preferences based on the political news they consume [8].

The result that exposure to small parties is associated with an increase in knowledge about their candidates would support a regulatory and normative goal of ensuring that people are familiar with a breadth of existing viewpoints. By providing minimum quotas of exposure for all parties, such an NRS would allow for making a more informed voting decision. This finding therefore supports using a model of deliberative diversity to curate the selection of articles in an NRS as a viable approach to achieving the social goal of a well-informed citizenry.

We recognize that amplifying right-wing extremist positions prior to elections could have harmful consequences. However, our study did not directly amplify the positions of the Right-Wing Populists. Instead, we artificially increased the visibility of the *journalistic coverage* related to this party, which often offers critical insights into party positions. For future user experiments, we urge researchers to conduct studies with utmost sensitivity to ethical concerns and societal implications as well as carefully consider the risks and benefits of featuring extreme political viewpoints.

7 LIMITATIONS AND FUTURE WORK

When using the approach to deliberative diversity outlined in this paper, there are three main limitations that need to be considered. First, the selection of diversity dimension is rooted in the local political landscape. When creating any visibility quotas, identifying parties, their proximity, and dominance is required to establish the categories over which the NRS diversifies. We acknowledge that such a setup is always going to be highly time-dependent and not directly applicable to other political landscapes and election campaigns.

Second, the success of deliberative diversity requires access to a broad selection of news sources, as the diversity of the NRS is ultimately limited by the supply side of news articles. For this reason, it is important to feature a selection of news outlets that cover all the relevant political entities identified within the local political landscape. In the context of news supply, a key consideration for future experiments is to also control for framing and sentiment of articles. Having this information in addition to the counts of political entities would allow to, e.g., create a more balanced reporting, by having an equal number of articles that mention a political entity in a positive and a negative context).

Finally, this study relies on self-reported responses by self-selected participants, which introduces the possibility of sampling, recall, and social desirability biases, potentially impacting the generalizability and reliability of the results. The experimental design only captures short-term effects of diverse news exposure on political attitudes and behavior. Future work can address the limitations by using larger and more diverse samples, conducting long-term longitudinal studies, and analyzing news consumption across multiple platforms to enhance the understanding of the impact of diverse news exposure on political attitudes and behaviors.

8 CONCLUSION

This study is among the first to operationalize the typology of democratic NRSs using the concept of deliberative diversity in a field experiment during elections. We operationalized deliberative diversity to serve as a model for NRSs and enable us to manipulate party visibility within news feeds based on real-life news content. By implementing the FUN pattern, we achieved a deliberative diversity distribution of party visibility, offering a valuable approach to studying the effects of NRSs in the political domain. Our findings offer insights into evaluating political actors, voting decisions, and democratic societies impacted by automated news distribution.

We saw the capacity of NRSs to enhance knowledgeability about minority parties within the treatment group. However, it is essential to consider that increased exposure and knowledge do not automatically translate to a change in voting behavior. Results suggest the possibility of an inverse relationship between exposure and party support, wherein heightened exposure and knowledge might even lead to opposing positions toward the parties under consideration.

Our findings highlight the nuanced and multifaceted nature of NRSs in the context of political news. These systems undoubtedly have a measurable and significant impact on their target audience, but their influence on cognitive processes, attitudes, and ultimately voting behavior requires further research. Designers of NRSs need to approach their development with caution and awareness, understanding that the relations between news recommendation, usage, cognition, attitudes, and voting behavior are not linear or one-dimensional. As an interdisciplinary team, we want to stress that collaboration between engineers and social scientists is essential to approach these challenges.

This collaboration allowed for a comprehensive approach to NRS research, integrating technical considerations with social and ethical factors. The interdisciplinary dialogue can identify potential biases and ethical concerns in NRSs, leading to more informed design decisions. Making the principles guiding the recommendation process explicit enables designers and engineers to recognize how we "willingly or unwittingly, infuse technical objects with values" [9], allowing them to consciously opt for specific normative principles. Such a holistic perspective ensures that NRSs align with democratic principles, empower users, and positively contribute to the public sphere.

ACKNOWLEDGMENTS

This user study was funded by the Digital Society Initiative (DSI) of the University of Zurich under a grant of the DSI Excellence Program, the Federal Ministry of Education and Research (BMBF), and the Free and Hanseatic City of Hamburg under the Excellence Strategy of the Federal Government and the Länder.

REFERENCES

[1] Abraham Bernstein, Claes De Vreese, Natali Helberger, Wolfgang Schulz, Katharina Zweig, Lucien Heitz, Suzanne Tolmeijer, et al. 2021. Diversity in News Recommendation. *Dagstuhl Manifestos* 9, 1 (2021), 43–61.

[2] David Broockman and Joshua Kalla. 2022. *The manifold effects of partisan media on viewers' beliefs and attitudes: A field experiment with Fox News viewers.* https://doi.org/10.31219/osf.io/jrw26

[3] Laia Castro. 2021. Measuring Partisan Media Bias Cross–Nationally. 27, 2 (2021), 412–433. https://doi.org/10.1111/spsr.12459

[4] Saul Cunow, Scott Desposato, Andrew Janusz, and Cameron Sells. 2021. Less is more: The paradox of choice in voting behavior. 69 (2021), 102230. https://doi.org/10.1016/j.electstud.2020.102230 PII: S0261379420301104.

[5] Jakob-Moritz Eberl, Hajo G Boomgaarden, and Markus Wagner. 2017. One bias fits all? Three types of media bias and their effects on party preferences. *Communication Research* 44, 8 (2017), 1125–1148.

[6] Myra Marx Ferree, William A Gamson, Jürgen Gerhards, and Dieter Rucht. 2002. Four models of the public sphere in modern democracies. *Theory and society* 31, 3 (2002), 289–324.

[7] R. Kelly Garrett and Natalie Jomini Stroud. 2014. Partisan paths to exposure diversity. 64, 4 (2014), 680–701. https://doi.org/10.1111/jcom.12105

[8] Sabine Geers and Linda Bos. 2017. Priming Issues, Party Visibility, and Party Evaluations: The Impact on Vote Switching. 34, 3 (2017), 344–366. https://doi.org/10.1080/10584609.2016.1201179

[9] Jeremy Grosman and Tyler Reigeluth. 2019. Perspectives on algorithmic normativities: engineers, objects, activities. *Big Data & Society* 6, 2 (2019), 2053951719858742.

[10] Jaron Harambam, Dimitrios Bountouridis, Mykola Makhortykh, and Joris van Hoboken. 2019. Designing for the Better by Taking Users into Account: A Qualitative Evaluation of User Control Mechanisms in (News) Recommender Systems. In *Proceedings of the 13th ACM Conference on Recommender Systems* (Copenhagen, Denmark) *(RecSys '19)*. Association for Computing Machinery, New York, NY, USA, 69–77. https://doi.org/10.1145/3298689.3347014

[11] Lucien Heitz, Juliane A Lischka, Alena Birrer, Bibek Paudel, Suzanne Tolmeijer, Laura Laugwitz, and Abraham Bernstein. 2022. Benefits of diverse news recommendations for democracy: A user study. *Digital Journalism* 10, 10 (2022), 1710–1730.

[12] Natali Helberger. 2019. On the democratic role of news recommenders. *Digital Journalism* 7, 8 (2019), 993–1012.

[13] Natali Helberger, Max van Drunen, Judith Moeller, Sanne Vrijenhoek, and Sarah Eskens. 2022. Towards a Normative Perspective on Journalistic AI: Embracing the Messy Reality of Normative Ideals. , 1605–1626 pages.

[14] David Nicolas Hopmann, Rens Vliegenthart, Claes H. de Vreese, and Erik Albæk. 2010. Effects of Election News Coverage: How Visibility and Tone Influence Party Choice. 27, 4 (2010), 389–405. https://doi.org/10.1080/10584609.2010.516798

[15] Mozhgan Karimi, Dietmar Jannach, and Michael Jugovac. 2018. News recommender systems–Survey and roads ahead. *Information Processing & Management* 54, 6 (2018), 1203–1227.

[16] Bumsoo Kim, Ryan Broussard, and Matthew Barnidge. 2020. Testing political knowledge as a mediator of the relationship between news use and affective polarization. (2020), 1–13. https://doi.org/10.1080/03623319.2020.1750845

[17] Silvia Knobloch-Westerwick. 2015. *Choice and preference in media use.* Routledge.

[18] Katharina Ludwig, Alexander Grote, Andreea Iana, Mehwish Alam, Heiko Paulheim, Harald Sack, Christof Weinhardt, and Philipp Müller. 2023. Divided by the Algorithm? The (Limited) Effects of Content- and Sentiment-Based News Recommendation on Affective, Ideological, and Perceived Polarization. (2023). https://doi.org/10.1177/08944393221149290

[19] Maxwell E McCombs and Donald L Shaw. 1972. The agenda-setting function of mass media. *Public opinion quarterly* 36, 2 (1972), 176–187.

[20] Joanne M Miller and Jon A Krosnick. 2000. News media impact on the ingredients of presidential evaluations: Politically knowledgeable citizens are guided by a trusted source. *American Journal of Political Science* (2000), 301–315.

[21] Eliza Mitova, Sina Blassnig, Edina Strikovic, Aleksandra Urman, Aniko Hannak, Claes H. de Vreese, and Frank Esser. 2022. News recommender systems: a programmatic research review. (2022), 1–30. https://doi.org/10.1080/23808985.2022.2142149

[22] Diana C Mutz. 2006. *Hearing the other side: Deliberative versus participatory democracy.* Cambridge University Press.

[23] Tim Neumann, Ole Kelm, and Marco Dohle. 2021. Polarisation and silencing others during the covid-19 pandemic in Germany: An experimental study using algorithmically curated online environments. *Javnost-The Public* 28, 3 (2021), 323–339.

[24] Michael Quinn Patton. 2009. *Qualitative research & evaluation methods* (3. ed., [nachdr.] ed.). Sage.

[25] Luca Rossetto, Matthias Baumgartner, Ralph Gasser, Lucien Heitz, Ruijie Wang, and Abraham Bernstein. 2021. Exploring Graph-querying approaches in LifeGraph. In *Proceedings of the 4th Annual on Lifelog Search Challenge*. 7–10.

[26] Holli Sargeant, Eliska Pirkova, Matthias C Kettemann, Marlena Wisniak, Martin Scheinin, Emmi Bevensee, Katie Pentney, Lorna Woods, Lucien Heitz, Bojana Kostic, et al. 2022. Spotlight on Artificial Intelligence and Freedom of Expression: A Policy Manual. *Organization for Security and Co-operation in Europe* (2022).

[27] Rüdiger Schmitt-Beck, Sigrid Roßteutscher, Harald Schoen, Bernhard Weßels, and Christof Wolf. 2022. A new era of electoral instability. In *The Changing German Voter*. Oxford: Oxford University Press, 3–24.

[28] Morten Skovsgaard and Kim Andersen. 2020. Conceptualizing news avoidance: Towards a shared understanding of different causes and potential solutions. *Journalism studies* 21, 4 (2020), 459–476.

[29] Celina Treuillier, Sylvain Castagnos, Evan Dufraisse, and Armelle Brun. 2022. Being Diverse is Not Enough: Rethinking Diversity Evaluation to Meet Challenges of News Recommender Systems. In *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*. 222–233.

[30] Patti M Valkenburg, Jochen Peter, and Joseph B Walther. 2016. Media effects: Theory and research. *Annual review of psychology* 67 (2016), 315–338.

[31] Saúl Vargas. 2011. New approaches to diversity and novelty in recommender systems. In *Fourth BCS-IRSG Symposium on Future Directions in Information Access (FDIA 2011) 4*. 8–13.

[32] Sanne Vrijenhoek, Gabriel Bénédict, Mateo Gutierrez Granada, Daan Odijk, and Maarten De Rijke. 2022. RADio–Rank-Aware Divergence Metrics to Measure Normative Diversity in News Recommendations. In *Proceedings of the 16th ACM Conference on Recommender Systems*. 208–219.