No Filter Bubbles? Evidence From an Online Experiment on the News Diversity of Personalizing News Aggregators

Completed Research Paper

Joschka Andreas Hüllmann
University of Twente
7522 NB Enschede, The Netherlands
j.huellmann@utwente.nl

Leonard Wilhelm Sensmeier
University of Münster
48149 Münster, Germany
l_sens01@uni-muenster.de

Abstract

People increasingly use personalizing news aggregators for acquiring news online. Despite benefits such as more relevance, filter bubbles, among other risks, have been elicited. So far, there is little empirical evidence on how personalization affects the news diversity of news aggregators. Furthermore, it remains unclear how the news diversity of news aggregators compares to edited newspaper websites. This study investigates the effect of personalization on news diversity and reports the results of an online experiment. Based on browser instrumentation, news articles were queried using personalized and non-personalized profiles. Using a fixed effects model, the news diversity's change due to personalization was estimated at 8-12%. However, the absolute news diversity of the personalizing news aggregators is comparable to edited newspaper websites with a difference of <5%. Our results contribute empirical insights to the debate on news personalization, finding that filter bubbles stemming from low news diversity are unlikely.

Keywords: news, curation, diversity, personalization, filter bubble, online experiment

Introduction

The internet has become a central source of news for society, used by 69% of Germans, according to Reuters (Newman et al., 2021). Internationally, more than 80% of Americans use digital devices to acquire news online (Shearer, 2021). One way to find news on the internet is through news aggregators. News aggregators are services that collect news from different sources and display them to the user (Athey & Mobius, 2012). News aggregators have become very popular and reach several million people in Germany and the United States (Athey et al., 2021; Newman et al., 2020). Which collection of news is displayed to the people is of great relevance for society because news shape a society's opinion and general education (Czepek et al., 2009). News keep citizens up to date with events, inform about political discourse, sports, and media and thus contribute to a society's culture, pluralism, and values (Oster, 2015).

News aggregators promise to connect news from many sources and display them together. For example, Google News and Apple News cover a broad array of content providers ranging from large journalistic outlets such as the New York Times to smaller regional newspapers. Voakes et al. (1996) argue that the sum of a society's news sources, as opposed to individual news sources, can reflect the diversity of that society. By aggregating news sources, a diverse range of topics can be addressed. Thereby, news aggregators contribute to a higher news diversity and help meet the information needs of a diverse society. Broad coverage of topics is the fundament of an educated society that can form opinions. A lack of diversity may result in a narrowly educated society, which ignores the discourse on important topics required to be functioning and responsible (Czepek et al., 2009).

One component of news aggregators is personalization. As the internet suffers from content overload, consumers cannot possibly consume all the news content that is available (Bodó et al., 2019). Consumers' information needs—that is, the information which consumers deem important for forming opinions—may not be met. Personalization tackles this problem. Personalization is based on recommender systems, that is, algorithmic mechanisms that analyze prior consumer behavior and filter the plethora of content on the internet for news that are relevant to the consumers (Hüllmann, 2021; Hüllmann et al., 2021). Hence, personalization leads to a better fit between an individual's information needs and the curated news collection (McQuail, 1984). For example, one of the co-founders of Flipboard describes it as the company's goal to provide a "deep view into everything from political issues to technology trends to travel inspiration" (Flipboard, 2021a) by bringing together news from a variety of sources. This automated news curation is further improved by recent advances in personalization. More tracking of consumers' behaviors as well as computational and methodological improvements provide better recommender systems for news provision.

Since the early 2010s, however, personalization has been increasingly criticized (Kitchens et al., 2020). The question arises whether too strong personalization overrepresents topics and conveys a biased picture of the news situation. Common subjects of discussion are filter bubbles, which describe intellectual isolation, resulting from a feedback loop of selective content presentation catering to consumers' interests. Filter bubbles divide users and create separate internet realities for individual users (Pariser, 2011). The consequence of such separate internet realities is selective exposure, that is, one-sided news consumption. One-sided news consumption distorts the political discourse and leads to a divided society (Stroud, 2010). The degree of personalization could be the determining characteristic that decides whether a news aggregator shows a diverse news offering that is interesting and relevant to the person; or whether it limits the news collection to the user's personal bubble, creating a feedback loop and reinforcing the division of society.

Algorithmic news curation has been of recent interest to the information systems community (Kitchens et al., 2020; Riemer & Peter, 2021). Although previous research has looked into biased representations of opinions and sentiments when looking at filter bubbles, the actual content presented has received little attention so far (Möller et al., 2018). The content, however, is a fundamental component of how news shape society. Despite some studies on news content diversity in social media, little is known about the personalization of online news aggregators, as the vendors are opaque about the mechanisms of their personalization algorithms (Le et al., 2019). Furthermore, the effect of personalization on the diversity of news collections has been neglected (Kitchens et al., 2020). As a result, empirical evidence for the extent of personalization and its impact on news diversity are lacking. In a recent MIS Quarterly publication, Kitchens et al. (2020) call for more research into "how [...] platforms shape the consumption of information and may foster the creation of information-limiting environments". Therefore, this paper raises the research question:

How does the personalization by online news aggregators affect the diversity of their displayed news collection?

We address this research question by conducting an online experiment, measuring the diversity of two news aggregator websites: Google News and Flipboard. Based on previous research, Google News is examined as a weakly personalizing news aggregator (Haim et al., 2018) and compared to Flipboard, which is examined as a strongly personalizing news aggregator (Flipboard, 2021c). Furthermore, we compare the two news aggregators' diversities against a baseline from an edited national and a local newspaper website from Germany, respectively. The national newspaper is one of the five biggest in Germany. We programmed a bot and scraper based on browser instrumentation that accesses the websites and extracts news articles with simulated profiles. The retrieved news articles (n=3229) are manually coded for their topic. Based on the coded topics, we compute the news diversity by using the Shannon diversity index. We show that personalization changes the news diversity by 8%-12% for Google News and Flipboard. However, the absolute news diversity of the personalizing news aggregators is similar to editorially curated news websites with a difference of <5%. We conclude that filter bubbles stemming from low news diversity are not a problem, thereby contributing important empirical insights to the debate on news personalization. Consumers can use news aggregators to acquire news online and do not need to worry about ending up in a filter bubble. A balance of individual information needs and societal relevance is achieved—similar to edited newspaper websites. Hence, no urgent policy intervention is required for now, although the mechanisms may change any other day and, thus, should be monitored.

The remainder of the paper is structured as follows: First, the basics of personalized news and news diversity are explained. The methods section describes the used browser profiles, their personalization process, the coding of news articles, and the statistical analysis. Then, the results of the survey are presented. In the end, the paper discusses the impact of the effects identified, addresses the study's limitations, and provides avenues for future research.

Background and Hypotheses

Personalized News and Filter Bubbles

The idea of personalized news emerged as early as the 1990s. Nicholas Negroponte describes a newspaper that is personalized for the user in his 1995 book "Being Digital". Instead of optimizing news for the broad mass of people, this newspaper is reduced to relevant information for the individual user or for a small number of users. Negroponte refers to this news experience as the "Daily Me" (Negroponte, 1995). Thurman and Schifferes (2012, p. 776) describe personalization today as follows: "*A form of user-to-system interactivity that uses a set of technological features to adapt the content, delivery, and arrangement of a communication to individual users' explicitly registered and/or implicitly determined preferences*." A distinction is made between two types of personalization. We speak of explicit personalization when the indication of preferences used for personalization originates from the user (Thurman & Schifferes, 2012). An example of explicit personalization is subscribing to a newsletter, that is, explicitly selecting preferences (Sundar & Marathe, 2010; Zuiderveen Borgesius et al., 2016). Contrary to explicit personalization, implicit personalization describes a process in which recommender systems personalize content based on collected data about a user (Thurman & Schifferes, 2012)—for example, tracking a user's internet activity and personalizing the content based on inferred interests (Zuiderveen Borgesius et al., 2016). From the inferred interests relevant news content is suggested (Bodó et al., 2019).

Implicit personalization of news that a user consumes can lead to filter bubbles that are optimized to the user's topical interests (Pariser, 2011). According to Pariser (2011), filter bubbles have three central characteristics. First, filter bubbles are unique to each user. One user's filter bubble is different from that of other users. Second, filter bubbles are invisible. It is unclear by what process and based on what data a prediction is created. Third, a user is not aware that their experience is personalized (Pariser, 2011). For news services, in particular, the degree of personalization and thus uniqueness is relevant, as it determines the positive or negative effects of a filter bubble. Contrary to filter bubbles, selective exposure is primarily associated with explicit personalization (Zuiderveen Borgesius et al., 2016) when users prefer news resembling or matching their topical interests (Stroud, 2011). Selective exposure and filter bubbles can have various effects on users' opinion-forming. They could cause users' education to suffer, as users are only exposed to content that follows their topical interests. Since they are not exposed to other topics, they are unable to form educated and diverse opinions (Trilling et al., 2017). This may affect the public discourse and lead to polarization. As recommender systems reinforce the users' topical interests, one-sided news content may prevail and impede a diverse public discourse (Kepplinger, 2016; Noelle-Neumann, 1974). Distorting the public perception by selective exposure or filter bubbles can lead to the division of society. Maintaining a sufficient news diversity can help to counteract these problems.

News Diversity

Journalistic actors aim to disseminate diverse news as widely as possible to achieve societal goals, such as cultural pluralism or a functioning society. According to Napoli (1999), news diversity consists of diversity in sources or content. Source diversity refers to the number of different owners of news sources. News services can influence the news reality by including or excluding sources. Content diversity describes the number of formats, program types, representation of the demographics of the recipients, and the number of available topics. News services should show opposing ideas so that diverse opinions are equally represented, and relevant topics are examined holistically (Napoli, 1999). Conversely, journalistic actors may try to steer a debate in their favor by focusing on or ignoring selected topics (called "false balance", Boykoff & Boykoff, 2004). Another type of news diversity is normative diversity (Napoli, 1999). Normative diversity inquiries about opinions and ideological sentiments and has seen extensive research interest recently (Bakshy et al., 2015; Kitchens et al., 2020). In recent research, news diversity has often been operationalized as "cross-cutting exposure of ideological leanings" (Möller et al., 2018). However, news consumption is more complex than exposure to opinions, sentiments, or political leanings. Content diversity, particularly exposure to topics, has been neglected, although it is relevant for a functioning and responsibly informed society (Helberger et al., 2018; Möller et al., 2018). As a result, we focus on the topic dimension of content diversity, that is, the diversity of topics shown. The topics should be equally distributed in prominence, or the distribution should be guided by the relevance of the topics to society (McQuail, 1984). The diversity of a news service can thus be evaluated by assessing the distribution of topics within that service (Voakes et al., 1996). Topics should be relevant to the user but also be relevant to broader society so that balance is achieved (Czepek et al., 2009).

Previous research has investigated news diversity and personalization since the 2010s, with opinion-forming content being examined multiple times, particularly focusing on social media and search engines. Bakshy et al. (2015) examined the political bias of news displayed to users on Facebook concerning the 2014 US election campaign, finding that users are shown more news that correspond to their political views. Robertson et al. (2018) analyzed Google search results of 187 volunteers from the USA and identified no evidence of filter bubbles, except for a slight political bias. Krafft et al. (2018) investigated personalized search results and filter bubbles on names and parties concerning the 2014 federal election in Germany. They used a data donation approach and concluded that established media houses with origins in print dominate search results. However, little personalization of search results was found, with the language setting and location being relevant factors. Other studies have found a reduction of normative news diversity (i.e., sentiments) on social media, inquiring about ideological dimensions of opinion-forming content (Beam, 2014; Dylko, 2016; O'Callaghan et al., 2015).

Le et al. (2019) looked at the political personalization of Google search with a focus on browsing history. They created and simulated browser profiles and concluded that the Google News search results are personalized, with a disproportionate number of articles confirming the user's opinion. Athey and Mobius (2012) investigated the impact of activating the local news feature in Google News on news consumption. Their research finds that users who turned on localization in Google News consumed more news in general and more local news. Haim et al. (2018) conducted two experiments on Google News. First, they explicitly personalized profiles on Google News by setting up Google accounts with distinct interests. Second, they implicitly personalized profiles by simulating internet user activity (similar to our approach). Both experiments showed minimal personalization. Approximately 2.5% of articles showed deviations from the baseline profile.

In summary, there is a strong research interest in personalization and news diversity. However, this interest is primarily related to Google search and social networks instead of news aggregator websites. For example, no publication could be found on personalization or news diversity for the leading news aggregator Flipboard. Furthermore, the current body of research focuses on normative diversity, opinions, and sentiments. These studies often show balanced sentiments and little personalization of news (Bakshy et al., 2015; Kitchens et al., 2020). On the other hand, the content and topic diversity as the foundation of building informed opinions are neglected (Möller et al., 2018). Since news aggregators acquire their content from established newspaper websites with little to no opinionated content, the normative diversity is less relevant for news aggregator research. Instead, the personalization's effect on the content's diversity of news aggregators is considered highly relevant, as it affects the education and opinion-forming of society. Although news aggregator websites are increasingly used for news consumption, it remains unclear how strong personalization affects the diversity of a news aggregator's content.

News Aggregators

The traditional news medium is the newspaper. Edited newspaper websites are the online presence of established newspaper organizations, for example, The New York Times. These organizations create their news through journalistic content production and investigation. Typically, their websites are manually curated with the content by human editors. Content curation and news diversity may follow varying logics depending on the goals of the newspaper outlet. For example, national outlets may cover a broad topic range, whereas niche outlets are focused by smaller newspaper websites. The human editors' news curation decisions may be informed by previous click data and professional experience.

News aggregators are news services that combine journalistic content from different third-party sources and make it available to the user (Chowdhury & Landoni, 2006). Sources include established newspaper websites from different domains and potentially countries. Beyond journalistic sources, news aggregators may, in rare cases, provide alternative news sources such as blogs (Schweiger et al., 2019). The aggregated news content is detached from its original context and rearranged in a new form. As a result, the original creators' goals, for example, neutrality or the journalistic quality, may be neglected in the process (Schweiger et al., 2019). Furthermore, the news diversity and topic arrangement from the original context of the news presentation are lost. News aggregators typically do not have a business relationship with the content creators whose content they link to (Athey & Mobius, 2012). The news used by news aggregators either come from an RSS feed deposited by the publisher (Flipboard, 2021b) or are retrieved using web scrapers. Web scrapers automatically extract information from the sourced websites to identify news articles. News aggregators typically follow an ad-based freemium business model. The free version contains advertisements that users can get rid of by buying a premium pass. Sometimes the premium pass includes extra features such as granular control of the displayed news collection.

Given the plethora of available news content on the internet, users are overwhelmed, and news aggregators see personalization as a potential remedy to keep users engaged. Therefore, an increasing number of people consume personalized news on the internet. Yet, there is a dearth of research into how personalization may, or may not, lead to a lacking diversity for news aggregators. Due to personalization and removing the news content from its original context, news diversity could be impaired (Chowdhury & Landoni, 2006). No or little personalization can lead to irrelevant news and overload. The right extent of news personalization has positive effects, such as a better fit between information needs and news content. Too much personalization can stifle news diversity and cause filter bubbles, which polarize society, among other adverse effects (Kitchens et al., 2020; McQuail, 1984). However, most news aggregator companies publish little or no information about how their personalization systems work. This opacity has led to personalized news often being criticized in the literature (Riemer & Peter, 2021). Independent studies are required to assess personalization and its effect on news diversity for the news aggregators' 'secret' algorithms.

Since the news aggregators' business model is based on ad revenue, they have a financial interest in maximizing users' length of stay and their click-through rate. To increase the length of stay, some news aggregators rely on personalization to curate the collection of news displayed to the user (Riemer & Peter, 2021). The premise is that more relevant news content increases the users' length of stay and click-through rate. To this end, Claussen et al. (2019) have shown that algorithmic personalization may have better economic returns than manual editing by humans. Recommender systems are used to implement this personalization, tracking and collecting user data to create rich user profiles with inferred interests (Thurman & Schifferes, 2012). Using these inferred interests, the news aggregators try to personalize the news content as much as possible to provide a narrow fit to the users' information needs and maximize their length of stay (Bodó et al., 2019). Consequently, we hypothesize that accessing the news aggregator's website with a personalized browser profile leads to less diverse news content than accessing the news aggregator's website with a baseline or non-personalized profile.

Hypothesis 1: Personalization decreases news diversity on news aggregator websites compared to a non-personalized baseline.

Since there exists no previous research quantifying news topic diversity on news aggregator websites, we decided to use established newspaper websites as a comparison baseline. We compare the news aggregator websites with a national and a local news website where the news content is edited and manually curated by humans. These websites do not personalize their news content for their users[1] (Claussen et al., 2019). Nevertheless, large editorial teams for national newspapers can present a broad diversity of news topics. Like national newspapers, news aggregators cater to a broad audience and present diverse topics for new users due to the cold start problem. Only through personalization do news aggregators achieve a better fit between information needs and the curation of news topics, identifying the narrow personal interests of users (Bodó et al., 2019). Against the theoretical backdrop on filter bubbles and the news aggregator companies' interest in maximizing their users' length of stay, it is likely that the personalization by news aggregator websites leads to less news diversity compared to established newspaper websites. We

[1] We acknowledge that some websites use A/B testing, and editors may use data analysis for deciding on the news curation.

hypothesize that the news diversity on news aggregator websites for personalized browser profiles is lower than those for edited newspaper websites.

Hypothesis 2: Personalization decreases news diversity on news aggregator websites compared to edited newspaper websites.

Methods

We conducted an online experiment that measured and tested news diversity across personalizing news aggregator websites and compared the results to edited newspaper websites. We chose Google News and Flipboard for news aggregator websites, as they are popular websites that represent two extreme cases of news aggregator websites: weakly and strongly personalizing. Figure 1 shows the layout of both websites. We chose two edited newspaper websites that are well established in Germany and known for neutral and broad news coverage, thus making for a good baseline. In the following, we first describe our case and the selected news websites before outlining the experimental setup in more detail. The experimental setup consisted of two phases. The first phase concerned the data collection and included creating the browser profiles and personalizing the simulated browser profiles. Then, the news websites were queried, and the news articles were scraped. The second phase concerned the analysis, which included coding all retrieved news articles for their topics, and lastly, running the statistical analysis on news diversity.

Case Description

Google News is the news aggregator from Google that was released in 2003. Google News includes three offerings: the iOS app, the Android app, and the Google News website. In the literature, the “News” section of Google Search is also referred to as “Google News Search” (Le et al., 2019). However, Google only addresses the search on the Google News website in the Google News Help, which can be used to search for topics, sources, and location (Google, 2021a). This paper focuses on the home page of the news aggregator, titled “Top Stories”, as it is assumed that most users use this to consume their news. In addition to the home page, Google News includes a personalized “for me” section, different subject areas on topics such as news with different local references (world news, national news, and local news), and features to follow news of interest. Google News was chosen as an example of a service with relatively little personalization for several reasons. First, previous research concluded that Google News shows little personalization in various places, specifically the Google News home page (Haim et al., 2018; Le et al., 2019). Furthermore, Google describes that only the for me section is personalized for the logged-in user, and the other sections are shown to all users (Google, 2021b). We, therefore, use Google News as a weakly personalizing news aggregator.

Flipboard is a news aggregator from the US that was first released as an iPad app in 2010 (Luna-Nevarez & McGovern, 2018). Nowadays, Flipboard is available as a website, iOS, Android, and Windows app. Again, the news service’s home page, titled “FOR YOU”, is examined. In addition to this, the “Daily Edition” is available, which displays articles on various topics and areas, such as national news, business, sports, or entertainment. A central feature of Flipboard is the following of subject areas, news channels, individuals, or magazines. Users and news channels can share magazine articles with other users (Flipboard, 2021c). Flipboard was chosen as an example of strongly personalizing aggregators because it is one of the most used news aggregators in Germany, and personalization is at the forefront of Flipboard’s marketing (Flipboard, 2021a). Flipboard continues to develop features for its personalization algorithm (Flipboard, 2018). The greater focus on personalization compared to Google News is evident in the structure of the website. When creating an account, the user is asked to select interests from a set of available topics. Flipboard’s home page is the “For Me” page. We, therefore, use Flipboard as a strongly personalizing news aggregator.

We choose to compare the personalization of the news aggregator websites to a baseline of **two manually editorialized newspaper websites**. The limited sizes of their editorial teams constrain the breadth of topics the edited newspaper website can address. To make baseline comparisons contingent on the size of the editorial team, we choose one national and one local newspaper. Both newspapers are known for their broad news coverage. The **national newspaper** website belongs to a weekly newspaper, with more than 500,000 print copies sold per week. Their online news portal is split from the print edition, and the online team has more than 160 paid editors who provide the news. They cover all topics from politics, sports, culture, arts, travel, and events to science, tech, food, and more. The **local newspaper** website belongs to a daily newspaper, with around 100,000 print copies sold per week. The company employs around 100 editors that curate the news for both the online and print edition. The local newspaper covers both superregional content as well as local content.

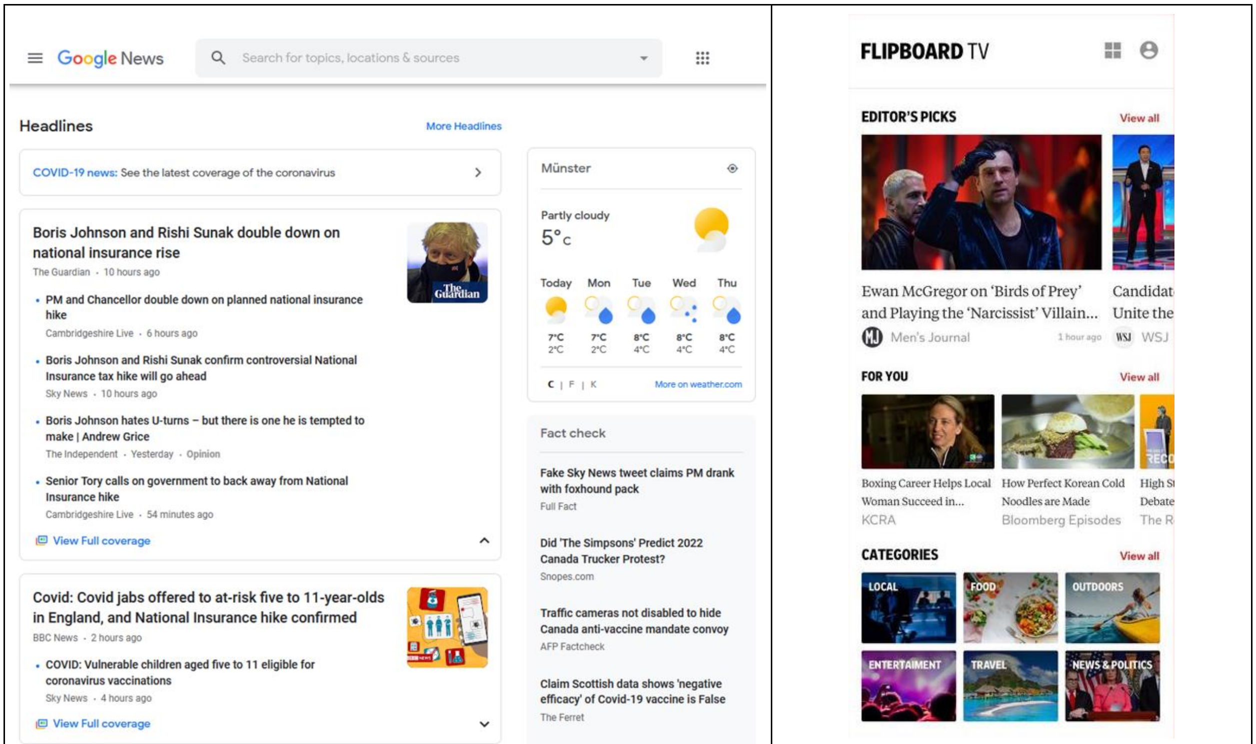

Figure 1. Screenshots of Google News and Flipboard.

Data Collection (Phase 1)

Previous research into personalization has made use of three prominent data collection approaches. First, researchers have used virtual machines with different profiles and then manually downloaded and coded the website content (e.g., Haim et al., 2018). Second, researchers call for data donations and crowdsource the data, resulting in a set of distinct, real profiles (e.g., Krafft et al., 2018; Robertson et al., 2018). Third, researchers make use of automated browser instrumentation (e.g., Le et al., 2019). We chose automated browser instrumentation for our study because it is highly controlled and scalable.

We developed four theoretically grounded, idealistic user descriptions based on studies about German society to ensure that the simulated profiles are comparable to real users. These descriptions inform how the browser profiles are personalized (Brenke & Kritikos, 2017; Mahrt & Begenat, 2013). The relevant attributes of these user descriptions include the sociodemographic attributes of gender, age, income, education level, political party, news preferences, and topical interests, and the technical attributes include operating system and browser. These attributes have previously shown to be relevant for personalization in the context of news (e.g., Haim et al., 2018; Le et al., 2019) and in more general contexts such as price discrimination (e.g., Badmaeva & Hüllmann, 2019; Klein & Hüllmann, 2018). The profiles are depicted in Table 1.

The browser profiles were personalized both explicitly and implicitly to capture both effects. For the explicit personalization, we manually selected the respective topical interests referencing the idealistic user descriptions on the Google News and Flipboard websites for each profile. Implicit personalization infers user interests by tracking the user's internet activities across multiple websites, for example, via cookies or browser fingerprinting (Kristol, 2001; Röttgen, 2018). Hence, we programmed our bot to simulate user activity by accessing the websites Google, YouTube, Amazon, eBay, and Flipboard. The access was designed to randomly select and interact with content (reading articles, viewing youtube videos, scrolling through websites, and following links) based on keywords derived from the defined attributes in the idealistic user descriptions. We included realistic delays, mouse movements, and devised a time schedule depending on the idealistic user description (some interactions in the morning, during the day, at night) to mimic human behavior. We also changed the user-agent in the browser profile to reflect the two technical attributes.

	Profile 1: „Teacher"	**Profile 2: „Economist"**	**Profile 3: „Student"**	**Profile 4: „Nurse"**
Gender	Male	Male	Female	Female
Age	30	50	18	40
Location	Germany	Germany	Germany	Germany
Income	above average	above average	average	below average
Education	master diploma	master diploma	bachelor diploma	high school
Political leaning	left, social	center, conservative	no interest	no interest
News orientation	local	international	none	boulevard
Interests	environment and sustainbility, social	finance, money	fashion, fitness	cooking, covid
Operation system	Windows, Chrome 89	Mac OS X, Chrome 89	Windows, Firefox 86	Windows, Chrome 80

Table 1. The user descriptions.

As illustrated in Figure 2, we have four simulated profiles in total for Google News and Flipboard, which are personalized according to the previous explanation. We have four separate baseline profiles, one for each website, which are not personalized. The non-personalized profiles are created empty without any cookies, internet activity history, or any existing state, except for creating a new Flipboard or Google News account, respectively. Our bot loads the correct browser profile and then queries the home page of each website under study. Then the scraper identifies all news articles on each home page, follows the links, downloads all news articles, and stores them on our server. The experiment was conducted over five days in March and April 2021. It started on the first day with personalizing the profiles (t=1) and continued with accessing and extracting the news articles from day three until the last day (t=3 until t=5, cf. Figure 2). Afterward, the coding and analysis were conducted.

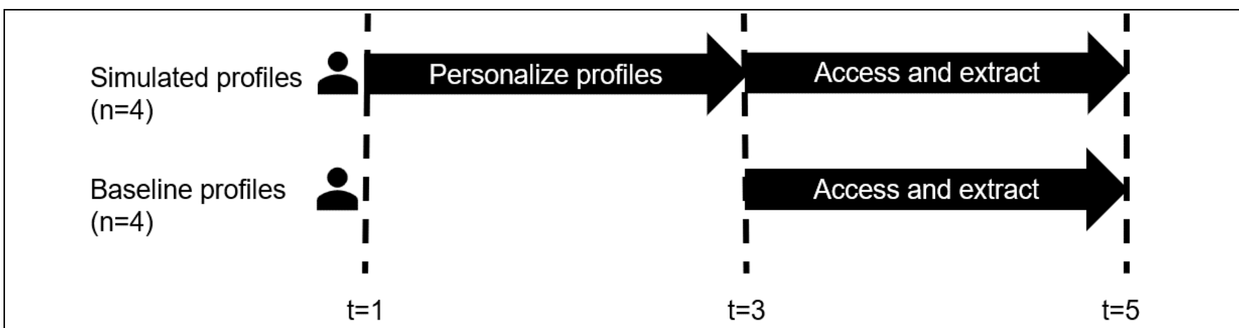

Figure 2. Timeline of data collection. The simulated profiles are personalized, the baseline profiles are not personalized.

The bot and the scraper are developed in python based on the selenium browser instrumentation framework. It uses the firefox geckodriver, as firefox is a common browser. Selenium executes all javascript and stores all cookies. The news articles are stored in mongodb. Each query is run within a docker container that only contains the browser, the selenium instrumentation, and the respective firefox browser profile. In the European Union, you must accept cookies, which we automated by bundling each firefox instance with a browser plugin called "I don't care about cookies".

Analysis (Phase 2)

Manipulating the level of personalization is achieved by selecting the respective browser profile. Hence, the intervention is binary, that is, using a simulated profile (=personalized), or using a baseline profile (=not personalized). All retrieved articles were then coded for their topic by the authors of this manuscript. We coded for a primary topic (n=12, cf. Table 2) and additionally for subtopics (n=60) to test the robustness of our analysis. Coding of the topics was performed by either adopting the news category of the news website – if available – or by determining a code intuitively through the research team. The coding procedure followed the guidelines by Kuckartz (2014) and the coded topics represented common news topics in the media (Newman et al., 2021).

For the subsequent statistical analysis, we pooled the different personalized profiles to calculate an average treatment effect of personalization. We removed all COVID articles as they biased the diversity due to being overly represented. We also removed articles that were blocked by a paywall because only a fraction of users in Germany pay for online news (Newman et al., 2021). From the remaining news articles, we calculated the news topic diversity. We chose the Shannon index as diversity measure because it is a widely used measure for assessing news diversity and the relative occurrence of topics (Möller et al., 2018). The Shannon diversity is widely used because it measures the entropy of a set of news topics, that is, how "surprising" the average occurrence of a topic is. It accounts for how evenly distributed the topics are in the sample (Morris et al., 2014). The Shannon diversity index $diversity$ for each home page is calculated as follows:

$$diversity = -\sum_i p_i \, ln(p_i) \text{ with } p_i = \frac{n_i}{N}$$

where, p_i denotes the proportional frequency of topic i relative to the total number of topics across all websites N. The index is then normalized by dividing through $ln(N)$ (Oksanen et al., 2020; Shannon, 1948).

To test our two hypotheses, we computed four fixed effects regression models, estimating the effects of personalization on news diversity for both news aggregators (Angrist & Pischke, 2009). Since the data collection occurred over multiple days, we controlled for time as a confounding effect and modeled a fixed effect for the variable days (Hanck et al., 2021). The model specification is:

$$diversity_{it} = \beta_0 + \beta_1 personalization_{it} + \beta_2 day_t + \epsilon_{it}$$

where, $diversity_{it}$ is the Shannon diversity index, $personalization_{it}$ is a dummy variable whether the profile is personalized or not, day_t is the unobserved effect of time, and ϵ_{it} is the error term.

The four models compare:

1. Google News personalized profile versus Google News baseline (hypothesis 1),
2. Flipboard personalized profile versus Flipboard baseline (hypothesis 1),
3. Google News personalized profile versus national and local news baselines (hypothesis 2),
4. Flipboard with personalized profile versus national and local news baselines (hypothesis 2).

Prior to each estimation, the data was filtered to include only the relevant news articles for comparison. The nonrelevant news articles were simply excluded for the respective estimations. For example, model (1) includes only data from Google News; or for example, model (3) only includes data from the personalized Google News profile, the national, and the local news baselines, but no data from the baseline Google profile or any of the Flipboard profiles.

All models were estimated using base R v4.1.1. The source code for both the bot and the statistical analysis and the data are available for reproducing the results on our GitHub (https://github.com/johuellm/news-aggregators).

Results

In total, we retrieved 3229 news articles from four websites (Table 2). The news articles were identified from 77 different home pages for which diversity scores have been calculated (Table 3). Google News shows about twice as many articles on the home page as Flipboard. Therefore, about twice as many articles from Google News were retrieved. The numbers for the simulated profiles in Table 2 are roughly double the numbers of the baseline profiles because the number for simulated profiles includes the news articles for both Google News and Flipboard, respectively. The distribution of coded topics shows that the news websites are dominated by news content related to the COVID pandemic.

Variable	Levels	n	%	Cum.%	Variable	Levels	n	%	Cum.%
Source	Flipboard	826	25.60	25.60	Topic	Other	373	11.60	11.60
	Google News	1898	58.80	84.40		COVID	1017	31.50	43.10
	National News	265	8.20	92.60		Economics	263	8.10	51.20
	Local News	240	7.40	100.00		Health	45	1.40	52.60
		3229	100.00			Leisure	429	13.30	65.90
Profile	Simulated 1	562	17.40	17.40		Culture	120	3.70	69.60
	Simulated 2	536	16.60	34.00		Social	114	3.50	73.10
	Simulated 3	541	16.80	50.80		Sports	158	4.90	78.00
	Simulated 4	508	15.70	66.50		Crime & Accidents	77	2.40	80.40
	Baseline (Flipboard)	178	5.50	72.00					
	Baseline (Google News)	399	12.40	84.40		Tech	134	4.20	84.60
	Baseline (National News)	265	8.20	92.60		Politics	469	14.50	99.10
	Baseline (Local News)	240	7.40	100.00		Science	30	0.90	100.00
		3229	100.00				3229	100.00	

Table 2. Summary of collected data.

	n	Mean	St.dev.	Variance	Median	Min	Max
Flipboard	28	0.62	0.15	0.023	0.66	0.29	0.85
Google News	25	0.62	0.13	0.017	0.59	0.45	0.89
National News	12	0.77	0.06	0.004	0.77	0.70	0.88
Local News	12	0.60	0.08	0.006	0.64	0.43	0.69

Table 3. Summary statistics of diversity.

	Google News	Flipboard	National News	Local News
Google News	1.00	-0.16	-0.24	-0.11
Flipboard	-0.16	1.00	0.25	0.02
National News	-0.24	0.25	1.00	-0.01
Local News	-0.11	0.02	-0.01	1.00

Table 4. Bivariate correlations of diversity.

The distributions of the Shannon diversity indices for each website are displayed in Table 3. The edited news websites vary only a little in diversity. The two personalizing news aggregator websites have more variance in diversity because the statistic includes both the baseline and the personalized profiles. The bivariate correlations of diversity indices across the four websites are depicted in Table 4. The table shows no strong correlation, as all values are below or equal to 0.25. None of the correlations is significant.

	Google News	**Google News**	**Flipboard**	**Flipboard**
	Shannon-Diversity	**Shannon-Diversity**	**Shannon-Diversity**	**Shannon-Diversity**
Baseline	0.4961 (0.0266)**	0.4734 (0.0225)**	0.7193 (0.0662)†	0.6938 (0.0703)†
Simulated	0.5849 (0.0119)***	0.5622 (0.0186)***	0.5944 (0.0307)***	0.5762 (0.0693)***
Day Fixed Effect	no	yes	no	yes
R^2	0.38	0.63	0.12	0.21
F Statistic	11.18 (df=1;18)**	6.40 (df=4;15)**	3.56 (df=1;26)	0.92 (df=6;21)
(Values are unstandardized coefficients; standard errors are in parentheses; †p<0.1; *p<0.05; **p<0.01; ***p<0.001)				
Table 5. Fixed effects model specification for hypothesis 1.				

We report the results of the two model fits (1, 2) for hypothesis 1 in Table 5. The results show that the simulated personalized Google News profile reaches 8.88% higher diversity than the Google News baseline profile. A Welch two-sample t-test shows that the difference is statistically significant ($t(4.57) = 3.31$, $p < 0.05$). Conversely, the simulated Flipboard profile has 12.49% lower diversity compared to the Flipboard baseline profile. Although the regression coefficient for the mean is only significant at the $p<0.10$ level, a Welch two-sample t-test shows that the difference is statistically significant ($t(23.81) = -3.53$, $p < 0.01$). It is interesting to note that, despite the inverted direction of change in means for the two websites, both personalized profiles for Google News and Flipboard end up at a mean diversity with a difference of <1%. Controlling for day fixed effects did not lead to a meaningful change in the estimated coefficients but improved the quality of the model fit.

	Google News	**Google News**	**Flipboard**	**Flipboard**
	Shannon-Diversity	**Shannon-Diversity**	**Shannon-Diversity**	**Shannon-Diversity**
Baseline (Local News)	0.5973 (0.0240)	0.5834 (0.0260)	0.5973 (0.0440)	0.5895 (0.0449)
Baseline (National News)	0.7730 (0.0240)***	0.7590 (0.0260)***	0.7730 (0.0440)***	0.7651 (0.0449)***
Simulated	0.5849 (0.0157)***	0.5782 (0.0251)***	0.5944 (0.0261)***	0.5918 (0.0477)***
Day Fixed Effect	no	yes	no	yes
R^2	0.66	0.70	0.30	0.37
F Statistic	35.61 (df=2;37)***	10.54 (df=7;32)***	9.28 (df=2;43)***	3.12 (df=7;38)*
(Values are unstandardized coefficients; standard errors are in parentheses; *p<0.05; **p<0.01; ***p<0.001)				
Table 6. Fixed effects model specification for hypothesis 2.				

The results of the two model fits (3, 4) for hypothesis 2 are reported in Table 6. Comparing the simulated personalized Google News profile diversity to the national news website shows that the national news website has 18% higher diversity than the simulated Google News website. Comparing the simulated Google News profile diversity to the local news website shows that the diversity between the two is similar (<2% difference). A Welch two-sample t-test shows that the comparison with the national news website is statistically significant ($t(21.40) = -9.36$, $p < 0.001$). The difference to the local news website is not significant. The comparison results are similar for Flipboard. Compared to the national news website, Flipboard has 17% less diversity, and again no meaningful difference from the local news website diversity (<2%). A Welch two-sample t-test shows that the comparison with the national news website is statistically significant ($t\ (28.79) = -4.74$, $p < 0.001$), while the difference with the local news website is not.

Since the comparison to the local news baseline is non-significant, we cannot make robust causal inferences about the differences between local news and personalized profiles. It is worth noting that the two baselines, national and local news, respectively, have different diversity means, with local news being closer to the simulated profiles rather than the national news diversity.

Robustness

We performed multiple robustness checks to increase confidence in the results and ensure that the results were not driven by the design choices of our experiment. We checked if the results were driven by (a) the chosen similarity metric, (b) violated assumptions of the regression analysis, (c) our approach to coding and filtering the topics, or (d) by single simulated profiles.

To check if the choice of our diversity metric caused the results, we repeated all calculations with the Simpsons diversity index instead of the Shannon diversity. The Simpson diversity index is similar to the Shannon index but puts more weight on dominant topics (Morris et al., 2014). The results for both indices remained consistent. Furthermore, we included a day-fixed effect to control for the effect of time. The inclusion of this fixed effect did not change the results. We checked whether there is a linear decrease towards more personalization and less diversity with increasing days. To check this linear relationship, a simple linear regression was fitted, but it showed an insignificant coefficient for the variable days. We also modeled an interaction effect between days and personalization to investigate days as a moderating variable, but the results were inconclusive.

Looking at the assumptions of our regression model estimation, a Shapiro Wilk test showed no evidence for a non-normal distribution of our diversity variable ($W = 0.976$, p-value $= 0.157$). We conducted a Durbin-Watson test, which showed no autocorrelation ($d = 0.045$, $p = 0.656$). The Breusch-Pagan test showed no evidence for heteroscedasticity ($bp = 0.07$, $p = 0.789$). Thus, we assume homogeneity of variance. We used Welch's two-sample t-test to avoid any further issues with heteroscedasticity.

To check the robustness of our coding approach, we repeated all calculations with the coding of subtopics instead of the primary topic. Instead of 12 topics, we used 60 different subtopics. This check eliminated the high diversity of the national news website because this news website was considerably less diverse across subtopics. For example, the subtopics for sports are "skiing, soccer, formula 1, canoeing", and the frequency of subtopics for the national news website is driven only by soccer. The effect is analogous for other topics, which are driven by a few dominant subtopics. Google and Flipboard are more equally distributed across subtopics, thus catering to niche topics more equally. The local news website is somewhere between the national news website and the two news aggregator websites.

We further coded more data for each news article based on the news diversity attributes mentioned in the background section, such as form (information, education, opinion piece), place (international, national, local news), or paywall (yes, no). We explored personalization for these attributes but found no meaningful differences.

To check the influence of COVID articles, we repeated all calculations with COVID articles included. Leaving the COVID articles in the data set reduced the diversity of national news and local news baselines because they focused heavily on COVID. In contrast, Google News and Flipboard – while also having news on COVID – did not focus less on other topics. Including the COVID articles, the difference from simulated to baseline diversity increased consistently in the direction of the effect in the original estimation. Google with the simulated profile was now 12% more diverse than the baseline, whereas Flipboard with the simulated profile was 15% less diverse than the baseline profile.

We tested the single profiles instead of pooling them for identifying the treatment effect per profile. However, there were no meaningful differences between the profiles, so we stuck with the average treatment effect across all profiles.

Discussion

Our results for Flipboard are intuitive because the personalization leads to less news diversity, as hypothesized. On the contrary, it was surprising that the personalized Google News profile was more diverse than the baseline, conflicting with our hypothesis 1. These effects can be explained by the different mechanisms of how the personalization works for both news aggregator websites (Möller et al., 2018). Due to these websites' secrecy about the mechanisms, these mechanisms were unknown before. The Google News baseline focuses on very few topics. For example, checking the home page of Google News right now (2021-11-12), it only shows COVID-19 and politics as topics. After the personalization, more topics of interest are added, and the news diversity increases. So, for Google News, the mechanism of personalization is additive in that it adds personalized topics to the news website. For Flipboard, it is the opposite. The Flipboard baseline shows topics from across the board (e.g., news, science, entertainment, politics), and the personalization effect is subtractive, that is, it removes the topics irrelevant to the target user. In the end, both simulated profiles end up at a similar level of diversity (<2% difference). As a result, hypothesis 1 is not supported. Although personalization changes the news diversity significantly, it does not necessarily decrease diversity. Instead, personalization may also increase diversity depending on how the baseline home page looks like, as seen for Google News.

Contrary to our hypothesis 2, the diversity of the news aggregator websites with personalized profiles is actually closer to the news diversity of the local newspaper website than compared to the non-personalized baseline profiles of the news aggregators. It remains unclear if this level of diversity is a designed goal by the news aggregator websites or by chance. The national news website has the highest diversity, which is expected as the local newspaper does not have as many staff members as the national news website. They cannot serve such a diverse range of topics. Looking at the subtopics and niche topics, the news aggregator websites may provide more news diversity than the local news website and facilitate broader opinion-forming. Hypothesis 2 is not supported, as the effect is more complicated than a simple linear relationship.

Contribution to Theory

We find evidence that personalization affects the diversity of the displayed news collection (Zuiderveen Borgesius et al., 2016). Despite widespread belief, the effect is not as clear as people think. Personalization can lead to both higher and lower news diversity, depending on the baseline and context, for example, the different personalization mechanisms that we identified. Similar suggestions have been put forward by Möller et al. (2018), for which we find empirical evidence. The personalization for Google News increases the diversity of the news collection, whereas for Flipboard, it decreases news diversity.

Beyond the relative change of news diversity, also the absolute level of diversity is relevant. Personalization is a double-edged sword. It can "expand an individual's information consumption by connecting them with a broad range of relevant information", but it can also reinforce filter bubbles and limit the user's access to diverse topics of interest (Kitchens et al., 2020). The result would be exposure to narrow news content, which may polarize a user's opinion-forming. For our sample, filter bubbles through personalization have a negligible effect on users' opinion-forming and education because the level of news diversity after personalization is comparable to that of edited newspaper websites. Since the news diversity after personalization is similar to the national and local news baselines, Google News and Flipboard do not induce more selective bias and distorted display of news than edited newspaper websites. Compared to the small-staffed local newspaper website, which cannot cover as many different topics, a news aggregator website may be a viable alternative and actually be a more diverse news source. Looking at the high diversity of the national news website compared to all other websites is a testament to the national news website and its successful effort to address a broad range of diverse topics.

Ultimately, we argue that the degree of personalization should lead to a **balanced news diversity** (Kitchens et al., 2020; McQuail, 1984). This balance ensures that, on the one hand, personalization should offer interesting and relevant news content to ensure a fit between the information needs and the available news content. On the other hand, personalization should not limit the news collection too strictly that it polarizes and divides the society. Again, our results show that this balance is achieved for Google News and Flipboard because their absolute diversity measures are similar to edited newspaper websites (assuming that the edited newspapers serve as a good example; the two edited newspaper websites from the sample are known to have a broad and neutral news coverage).

We conclude that people who primarily consume their news from news aggregator websites such as Flipboard and Google News do not have to worry about filter bubbles or a lacking news diversity. News aggregator websites do not lead to filter bubbles due to low news diversity in terms of topics. However, we emphasize that this conclusion only holds for the content dimension of news diversity and not for polarizing or ideology-laden news sources. We call the latter "normative diversity" (also referred to as "sentiment" or "tonality"). For example, politics may be the topic, but within the topic of politics, news aggregators might curate and display ideologically polarizing news articles (Bakshy et al., 2015). As a result, people using online news websites will not be ill-informed or in a filter bubble due to a lack of news diversity. However, they might be ill-informed or in a filter bubble due to polarizing sources or opinions and a lack of normative diversity. Although we argued that normative diversity is less useful for researching news aggregators due to the curation of non-opinionated content, this curation may change in the future. Given recent developments in state-sponsored news, future research may reconsider including normative diversity.

News aggregator websites are qualified for forming a general opinion and education across a diverse range of topics, fulfilling the information needs of a diverse society (Claussen et al., 2019; Möller et al., 2018; Voakes et al., 1996). Future research may look into how the personalization of news aggregators affects content and normative diversity jointly, building on top of our study and extending the works by Bakshy et al. (2015) and Kitchens et al. (2020) to the context of news aggregators.

Contribution to Practice and Policy

Consumers can benefit from using personalizing news aggregators for acquiring news because these platforms cover both broad topics and niche topics. The news diversity is comparable to that of edited newspaper websites. Consumers relying on personalizing news aggregators will be served a range of topics, enabling broad informing and reducing the risk of filter bubbles. Both implicit personalization through liking and viewing content, or explicit personalization through selecting topics of interest, may help consumers to balance topics of societal relevance and personal interests.

We uncover insights about the previously unknown mechanisms of personalization on news aggregator websites, that is, additive and subtractive personalization. Although the news diversity was comparable to that of edited newspaper websites at the time of the study, the news aggregators remain opaque about their algorithms. These algorithms can change anytime. Since the news aggregators' revenue is based on maximizing the users' length of stay, the public should remain wary of future changes to the algorithms. As a result, we do not have explicit policy recommendations at this time, however, continuous monitoring and assessment of the algorithmic logics are encouraged (cf. Riemer & Peter, 2021).

Generalizability and Limitations

This study inquired about two prominent and leading news aggregators and compared them to representative edited newspaper websites in Germany. The structure of the German news industry is similar to other European and Western countries such as the United States, the Netherlands, Denmark, and France (Claussen et al., 2019). Therefore, the results may generalize to these countries. However, for other countries where the news industry operates differently, the results may differ. More specialized news aggregator websites may follow other recommendation logics and thus personalize the news curation in different ways than inquired here.

Despite our best efforts, our study naturally has limitations. We only collected five days of data with our bot—the data collection results in 3229 articles and 77 home pages. Collecting data for an extended period might produce further interesting insights. The manual coding of the news articles was our bottleneck. Nevertheless, the collected sample is big enough to generate relevant and significant results. We performed the coding of the topics ourselves, which introduces subjectivity into the analysis. However, we checked the results by repeating the analysis with the coded sub-topics, and we stuck to the assigned categories by the websites as closely as possible. We evaluated the success of manipulating the degree of personalization by verifying the inferred interests for Google and Flipboard within the *ads settings* on the respective website.

However, we acknowledge that the profile simulation using automated browser instrumentation is not identical to real-world profiles that are used in other studies with volunteers (Robertson et al., 2018).

Furthermore, automated browser instrumentation can suffer from bot detection. We added various activities to mimic real human behavior and prevent bot detection, for example, random delays for user interactions with the websites. We have not received any "403 forbidden" errors, which are typically used when bots are blocked. Our experimental setup is a black-box test and does not elucidate the inner workings of the personalization algorithms. The interpretation of the results is our own. As outlined in the discussion, we focused on the content dimension of news diversity, not sentiments or normative diversity. Future studies can follow up and address the limitations of this study. Such studies will provide even more insights into the effects of personalization on news diversity, overcoming the limitations of this work.

Conclusion

We conducted an online experiment to quantify the effect of personalization by online news aggregators on the diversity of their displayed news collection. After collecting data through automated browser instrumentation and simulated browser profiles, we fit four fixed effects models. The estimated change of news diversity due to personalization is between 8% and 12%. Remarkably, this effect is not unidirectional. While the personalization of Flipboard decreases news diversity, as hypothesized, the personalization of Google News increases news diversity. We explain how this effect occurs due to the observed differences in personalization mechanisms. These mechanisms warrant further research. Comparing the absolute values for news diversity, we find negligible differences between personalizing news aggregators and edited newspaper websites. Hence, we find no filter bubble effect for news diversity in terms of topics. News aggregators are a viable medium for fulfilling diverse information needs, especially since they also cater to niche topics. Through our study, we contribute theoretical and empirical insights on the news diversity of personalizing news aggregators, addressing the neglected context of news aggregators and content diversity. Future research should extend our study into the context of other news industries and investigate the joint effects of personalization on normative and content diversity in news aggregators.

Acknowledgments

We declare no potential conflicts of interest with respect to the research, authorship and/or publication of this article. We are unaffiliated with all websites that were part of this online experiment. We thank Stefan Klein, Dominik Gutt, Jessica Ulke, and the anonymous reviewers for feedback on this work.

References

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics*. Princeton University Press. https://doi.org/10.1515/9781400829828

Athey, S., & Mobius, M. (2012). The Impact of News Aggregators on Internet News Consumption: The Case of Localization. In *TNIT Working Paper*. http://www.markusmobius.org/sites/default/files/localnews.pdf

Athey, S., Mobius, M. M., & Pál, J. (2021). The Impact of Aggregators on Internet News Consumption. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3838513

Badmaeva, T., & Hüllmann, J. A. (2019). Investigating Personalized Price Discrimination of Textile- , Electronics- and General Stores in German Online Retail. *Proceedings of the 14th International Conference on Wirtschaftsinformatik (WI).*

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Beam, M. A. (2014). Automating the News: How Personalized News Recommender System Design Choices Impact News Reception. *Communication Research*, *41*(8), 1019–1041. https://doi.org/10.1177/0093650213497979

Bodó, B., Helberger, N., Eskens, S., & Möller, J. (2019). Interested in Diversity. *Digital Journalism*, *7*(2), 206–229. https://doi.org/10.1080/21670811.2018.1521292

Boykoff, M. T., & Boykoff, J. M. (2004). Balance as bias: global warming and the US prestige press. *Global Environmental Change*, *14*(2), 125–136. https://doi.org/10.1016/j.gloenvcha.2003.10.001

Brenke, K., & Kritikos, A. (2017). Wählerstruktur im Wandel. *DIW Wochenbericht*, *84*(29), 595–606.

Chowdhury, S., & Landoni, M. (2006). News aggregator services: User expectations and experience. *Online Information Review*, *30*(2), 100–115. https://doi.org/10.1108/14684520610659157
Claussen, J., Peukert, C., & Sen, A. (2019). The Editor vs. The Algorithm: Returns to Data and Externalities in Online News. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3518959
Czepek, A., Hellwig, M., & Nowak, E. (2009). *Press Freedom and Pluralism in Europe: Concepts and Conditions*. Intellect Books.
Dylko, I. B. (2016). How Technology Encourages Political Selective Exposure. *Communication Theory*, *26*(4), 389–409. https://doi.org/10.1111/comt.12089
Flipboard. (2018). *Personalisiere dein Flipboard und bekomme nur was dir wichtig ist*. https://de-de.about.flipboard.com/flipboard-deutschland/personalisiere-dein-flipboard-und-bekomme-nur-was-dir-wichtig-ist/
Flipboard. (2021a). *Home Flipboard*. https://about.flipboard.com/
Flipboard. (2021b). *Publishing Q&A*. https://about.flipboard.com/publishers-qa/?noredirect=en_US
Flipboard. (2021c). *What is Flipboard?* https://flipboard.helpshift.com/a/flipboard/?s=rules-reporting&f=what-is-flipboard&p=all
Google. (2021a). *Get started with Google News*. https://support.google.com/googlenews/answer/9005669?hl=en&ref_topic=7688381
Google. (2021b). *How Google News stories are selected*. https://support.google.com/googlenews/answer/9005749?hl=en&ref_topic=7688382
Haim, M., Graefe, A., & Brosius, H. B. (2018). Burst of the Filter Bubble? Effects of personalization on the diversity of Google News. *Digital Journalism*, *6*(3), 330–343. https://doi.org/10.1080/21670811.2017.1338145
Hanck, C., Arnold, M., Gerber, A., & Schmelzer, M. (2021). *Introduction to Econometrics with R*. Self-Published.
Helberger, N., Karppinen, K., & D'Acunto, L. (2018). Exposure diversity as a design principle for recommender systems. *Information, Communication & Society*, *21*(2), 191–207. https://doi.org/10.1080/1369118X.2016.1271900
Hüllmann, J. A. (2021). *Smarter Work? Promises and Perils of Algorithmic Management in the Workplace Using Digital Traces*. Self-Published.
Hüllmann, J. A., Sivakumar, A., & Krebber, S. (2021). Data Management Platforms: An Empirical Taxonomy. *Proceedings of the 34th Bled EConference*.
Kepplinger, H. M. (2016). Die Schweigespirale. Öffentliche Meinung – unsere soziale Haut. In *Schlüsselwerke der Medienwirkungsforschung* (pp. 173–182). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-09923-7_15
Kitchens, B., Johnson, S. L., & Gray, P. (2020). Understanding Echo Chambers and Filter Bubbles: The Impact of Social Media on Diversification and Partisan Shifts in News Consumption. *MIS Quarterly*, *44*(4), 1619–1649. https://doi.org/10.25300/MISQ/2020/16371
Klein, S., & Hüllmann, J. A. (2018). Datenkapitalismus akademischer Wissenschaftsverlage. *Wirtschaftsdienst*, *98*(7), 477–480. https://doi.org/https://doi.org/10.1007/s10273-018-2318-3
Krafft, T. D., Gamer, M., & Zweig, K. A. (2018). What did you see? Personalization, regionalization and the question of the filter bubble in Google's search engine. *CEUR Workshop Proceedings*, *2657*, 1–9. https://doi.org/1812.10943
Kristol, D. M. (2001). HTTP Cookies: Standards, Privacy, and Politics. *ACM Transactions on Internet Technology*, *1*(2), 151–198. https://doi.org/10.1145/502152.502153
Kuckartz, U. (2014). *Qualitative Text Analysis* (K. Metzler (ed.); English ed). SAGE Publications Ltd.
Le, H., High, A., Maragh, R., Havens, T., Ekdale, B., & Shafiq, Z. (2019). Measuring political personalization of Google news search. *Proceedings of the World Wide Web Conference, WWW 2019*, 2957–2963. https://doi.org/10.1145/3308558.3312504
Luna-Nevarez, C., & McGovern, E. (2018). On the Use of Mobile Apps in Education: The Impact of Digital Magazines on Student Learning. *Journal of Educational Technology Systems*, *47*(1), 17–31. https://doi.org/10.1177/0047239518778514
Mahrt, M., & Begenat, M. (2013). Von Lebenswelten und ihren Horizonten. Mediennutzung und Themenwahrnehmung in politischen Milieus. *Medien & Kommunikationswissenschaft*, *61*(1), 21–37. https://doi.org/10.5771/1615-634x-2013-1-21
McQuail, D. (1984). McQuail´s Mass Communication Theory. In *Social Science Information Studies* (6th ed.). SAGE Publications Ltd.
Möller, J., Trilling, D., Helberger, N., & van Es, B. (2018). Do not blame it on the algorithm: an empirical

assessment of multiple recommender systems and their impact on content diversity. *Information, Communication & Society*, *21*(7), 959–977. https://doi.org/10.1080/1369118X.2018.1444076

Morris, E. K., Caruso, T., Buscot, F., Fischer, M., Hancock, C., Maier, T. S., Meiners, T., Müller, C., Obermaier, E., Prati, D., Socher, S. A., Sonnemann, I., Wäschke, N., Wubet, T., Wurst, S., & Rillig, M. C. (2014). Choosing and using diversity indices: insights for ecological applications from the German Biodiversity Exploratories. *Ecology and Evolution*, *4*(18), 3514–3524. https://doi.org/10.1002/ece3.1155

Napoli, P. M. (1999). Deconstructing the diversity principle. *Journal of Communication*, *49*(4), 7–34. https://doi.org/10.1111/j.1460-2466.1999.tb02815.x

Negroponte, N. (1995). *BEING DIGITAL*. Vintage Books.

Newman, N., Fletcher, R., Schulz, A., Andı, S., Robertson, C., & Kleis Nielsen, R. (2021). *Reuters Institute Digital News Report 2021*. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021

Newman, N., Richard Fletcher, W., Schulz, A., Andı, S., & Kleis Nielsen, R. (2020). *Reuters Institute Digital News Report 2020*. Reuters Institute for the Study of Journalism.

Noelle-Neumann, E. (1974). The Spiral of Silence A Theory of Public Opinion. *Journal of Communication*, *24*(2), 43–51. https://doi.org/10.1111/j.1460-2466.1974.tb00367.x

O'Callaghan, D., Greene, D., Conway, M., Carthy, J., & Cunningham, P. (2015). Down the (White) Rabbit Hole: The Extreme Right and Online Recommender Systems. *Social Science Computer Review*, *33*(4), 459–478. https://doi.org/10.1177/0894439314555329

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2020). *vegan: Community Ecology Package*. https://cran.r-project.org/package=vegan

Oster, J. (2015). *Media Freedom as a Fundamental Right*. Cambridge University Press. https://doi.org/10.1017/CBO9781316162736

Pariser, E. (2011). *The Filter Bubble*. Penguin Bookes.

Riemer, K., & Peter, S. (2021). Algorithmic audiencing: Why we need to rethink free speech on social media. *Journal of Information Technology*, *36*(4), 409–426. https://doi.org/10.1177/02683962211013358

Robertson, R. E., Jiang, S., Joseph, K., Friedland, L., Lazer, D., & Wilson, C. (2018). Auditing Partisan Audience Bias within Google Search. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 1–22. https://doi.org/10.1145/3274417

Röttgen, C. (2018). Like or Dislike—Web Tracking. In T. Hoeren & B. Kolany-Raiser (Eds.), *Big Data in Context. SpringerBriefs in Law.* (pp. 73–80). Springer. https://doi.org/10.1007/978-3-319-62461-7_9

Schweiger, W., Weber, P., Prochazka, F., & Brückner, L. (2019). *Algorithmisch personalisierte Nachrichtenkanäle*. Springer Fachmedien. https://doi.org/10.1007/978-3-658-24062-2

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shearer, E. (2021). *More than eight-in-ten Americans get news from digital devices*. Pew Research Center.

Stroud, N. J. (2010). Polarization and partisan selective exposure. *Journal of Communication*, *60*(3), 556–576. https://doi.org/10.1111/j.1460-2466.2010.01497.x

Stroud, N. J. (2011). *Niche News: The Politics of News Choice*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199755509.001.0001

Sundar, S. S., & Marathe, S. S. (2010). Personalization versus customization: The importance of agency, privacy, and power usage. *Human Communication Research*, *36*(3), 298–322. https://doi.org/10.1111/j.1468-2958.2010.01377.x

Thurman, N., & Schifferes, S. (2012). The Future of Personalization at News Websites. *Journalism Studies*, *13*(5–6), 775–790. https://doi.org/10.1080/1461670X.2012.664341

Trilling, D., Van Klingeren, M., & Tsfati, Y. (2017). Selective exposure, political polarization, and possible mediators: Evidence from the Netherlands. *International Journal of Public Opinion Research*, *29*(2), 189–213. https://doi.org/10.1093/ijpor/edw003

Voakes, P. S., Kapfer, J., Kurpius, D., & Chern, D. S.-Y. (1996). Diversity in the News: A Conceptual and Methodological Framework. *Journalism & Mass Communication Quarterly*, *73*(3), 582–593. https://doi.org/10.1177/107769909607300306

Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review*, *5*(1), 1–16. https://doi.org/10.14763/2016.1.401