UvA-DARE (Digital Academic Repository)

Explaining Online News Engagement Based on Browsing Behavior: Creatures of Habit?

Möller, J.; van de Velde, R.N.; Merten, L.; Puschmann, C.

Explaining Online News Engagement Based on Browsing Behavior: Creatures of Habit?

Judith Möller[1], Robbert Nicolai van de Velde[1], Lisa Merten[2], and Cornelius Puschmann[2]

Abstract
Understanding how citizens keep themselves informed about current affairs is crucial for a functioning democracy. Extant research suggests that in an increasingly fragmented digital news environment, search engines and social media platforms promote more incidental, but potentially more shallow modes of engagement with news compared to the act of routinely accessing a news organization's website. In this study, we examine classic predictors of news consumption to explain the preference for three modes of news engagement in online tracking data: routine news use, news use triggered by social media, and news use as part of a general search for information. In pursuit of this aim, we make use of a unique data set that combines tracking data with survey data. Our findings show differences in predictors between preference for regular (direct) engagement, general search-driven, and social media–driven modes of news engagement. In describing behavioral differences in news consumption patterns, we demonstrate a clear need for further analysis of behavioral tracking data in relation to self-reported measures in order to further qualify differences in modes of news engagement.

This article is part of the SSCR special issue on "*Integrating Survey Data and Digital Trace Data*", guest edited by Sebastian Stier, Johannes Breuer, Pascal Siegers (GESIS - Leibniz Institute for the Social Sciences) & Kjerstin Thorson (Michigan State University).

Mapping patterns of news consumption among different users has a long tradition in the social sciences (for an overview, see Mitchelstein & Boczkowski, 2010). The pathway to news conditions the ways in which citizens inform themselves about current events and form political attitudes.

[1] University of Amsterdam, Amsterdam, the Netherlands
[2] Leibniz Institute for Media Research | Hans-Bredow-Institut, Hamburg, Germany

Corresponding Author:
Judith Möller, University of Amsterdam, 1012 WX Amsterdam, the Netherlands.
Email: j.e.moller1@uva.nl

Consequently, understanding how, when, and why citizens are exposed to news is of crucial importance for any democratic society.

There is a wide range of methods available to study patterns of news use. Previously, diary studies, surveys, and a range of technical measurement strategies have been employed to capture audiences' news consumption. However, since an increasing number of people consume news online (Newman, Fletcher, Kalogeropoulos, Levy, & Nielsen, 2018), new and more precise avenues for the study of online news consumption have opened up, in particular, tracking user behavior online.

The shift toward online news consumption has not only improved our ability to observe news use, it has also had a fundamental impact on consumption patterns themselves. The affordances of online news create unique conditions compared to other media. Linear media, for example, afford news use as part of daily routine (Diddi & LaRose, 2006). Reading the newspaper during breakfast or watching television news has been and still is a habitualized behavior for many news users (Wonneberger, Schoenbach, & van Meurs, 2011), and it is clear that the temporal and spatial conditions of these legacy media have implications for the way information is accessed and processed (Graber, 1988). Online news use is associated by many scholars with more personalized and audience-led consumption patterns (Tewksbury, 2005); however, it has also been observed that these developments have led to an increased risk of fragmentation and polarization in public discourse (Tewksbury, Hals, & Bibart, 2008). It has been an oft-reiterated concern that the personalization afforded by search engines and social media may inadvertently trap users inside "filter bubbles" (Pariser, 2011) or lead them to assembling "echo chambers" (Sunstein, 2001) that reinforce their views while neglecting diversity. Recent empirical work, though, has consistently found that these risks are frequently overstated (Flaxman, Goel, & Rao, 2016; Haim, Graefe, & Brosius, 2018); yet the debate clearly indicates that the ways in which citizens come across news and what role personalization plays in this process is a pivotal concern (Diakopoulos & Koliska, 2017). The technology-driven change in habitualized behavior, therefore, has implications beyond the affordances of the specific device that is used to access news. As consuming news becomes embedded into casual browsing sessions, we need to develop a new understanding of news consumption as part of the wider online experience. Do users cultivate new routines that mirror the patterns of offline news consumption or does news use become more incidental as part of general information searches or social media experiences? We aim to study and explain patterns and modes of news use online, focusing on both motivational as well as demographic factors.

Measuring Online News Use

Past studies of news audiences have used a wide range of methods to explore how, when, and why people expose themselves to the news. The most common method to assess news diets are self-reports collected from survey responses (de Vreese & Neijens, 2016) and media diaries (Kubey, Larson, & Csikszentmihalyi, 1996), the advantage of which is that an individual's entire media repertoire can be captured. However, the limitation of any study that relies on self-reports is that participants need to be able to recall their exposure to media channels and content. Alternatively, some research methods directly measure media use through devices such as set-top boxes, or "people meters," embedded in televisions that collect viewing data; these kinds of methods allow for a passive and precise measurement of exposure to news on a single device (e.g., Wonneberger et al., 2011). When comparing these two methods with each other, the extent of measurement error revealed is surprisingly severe. Prior (2009) used people meter data to evaluate the validity of self-report measures and concluded that news users structurally overestimate the time that they spend consuming news.

We expect that it is even more challenging for respondents to reliably self-assess their exposure to news online, especially if it is not part of a news routine. As news exposure can also become ingrained into a browsing session that primarily serves a different purpose, at a cognitive level, it

becomes challenging to remember these very short episodes of exposure. Prior (2013) suggests that the solution to flawed self-reports of new exposure lies in monitoring news use online or using digital trace data to infer it. Recent advances in news audience research make use of these techniques (e.g., Mukerjee, Majó-Vázquez, & González-Bailón, 2018; Taneja & Webster, 2016), but so far there are few studies (e.g., Guess, Nyhan, & Reifler, 2018, on the consumption of the so-called fake news; Wieland et al., 2018, on the relevance of information intermediaries as new sources) that combine the precision of digital traces or tracking data with the rich nature of survey data to measure background and motivational variables that can then be used to explain specific patterns of news use.

In this study, we make use of a unique data set that includes both tracking and survey data. This means that we can not only provide exact measures of exposure to news use online but can also contextualize these episodes of exposure by tracing how users discover news content during their browsing sessions. By adding survey data, we are able to link use patterns to the users' characteristics. This has two distinct advantages. First, we can account for news exposure offline. Second, we can learn whether or not certain patterns are more prevalent in specific sections of the population. Considering that accessing news as part of a routine means that information is processed differently than when users come across news incidentally (Bode, 2016), or not at all, this has implications for potential knowledge gaps among the population.

Modes of News Use Online

During a web browsing session, users encounter news stories in many different ways. They might consciously seek out a news website, but it might also "find them" while they are searching for information for an upcoming vacation or checking updates from their peers on social media (Gil de Zúñiga, Weeks, & Ardèvol-Abreu, 2017). These different routes to accessing the news matter: Ethnographic research has demonstrated that the context in which news use occurs shapes readers' expectations of new information as well as their willingness to engage with it (Costera Meijer & Groot Kormelink, 2015; Schrøder, 2015). News has become an integral part of the online experience, and this includes news outside of the classic news website. Search engines, for example, integrate and highlight news items as search output and social media facilitate their inclusion in the newsfeed. As a result, news providers have adapted to these different access paths and optimized news items for the environment in which they are encountered (Bright, 2016). To make sense of the patterns of news use online, the context of the exposure to news items—as the user experiences it—must be taken into account.

To this end, we rely on the concept of communication modes (Hasebrink, 2004), which finds its roots in uses and gratifications theory. Communication modes are defined as the situation-specific pattern of functional expectations and actions that users employ to realize a certain information need such as gathering information on current events (cf. Hasebrink, 2004, p. 73). These situation-related communication modes are conditioned by the device used and the imagined co-audience (e.g., a dispersed public, or experts) for the same information. Building on previous research on situational modes of news use (Hasebrink & Hölig, 2017), we distinguish between three modes of online news use: routine news use, news use triggered by social media engagement, and news use as part of a search for specific information. Because of the differential affordances of these three modes, we expect that while all users of online news are generally likely to engage in each of the different modes depending on their specific information needs, they will use them to differing extents. This is because, first, context and user characteristics are important factors in determining the modes users engage in. Second, considering the crucial relevance of information needs in determining the selected mode of news use, motivational factors will also play a role.

Routine Mode: Routine Use of News

The routine use of news online implies a mode of using the news that mirrors news use offline, when people visit the home page of a news outlet as part of their daily routine, for example. The underlying motivations are manifold, for example, to find out whether new relevant news has broken (Diddi & LaRose, 2006) or to comply with social norms (e.g., duty to keep informed, Poindexter & McCombs, 2001). This mode of news use is habitualized and triggered by situational and contextual cues (Wonneberger et al., 2011) even in multi-optional media environments such as an online browsing session (Yadamsuren & Erdelez, 2011).

Search Mode: News Use Triggered by Search

News can be encountered during a general search for information. Searching for information is one of the key functions of Internet use and can be understood as a response to complex informational needs (Choo, Detlor, & Turnbull, 2000). News (as recent and relevant information) can be part of this search for information. News is then presented in information cards or boxes that are proactively offered to the user based on an algorithmic inference on the users' exact information need (Shokouhi & Guo, 2015). However, it is important to keep in mind that the situation-specific information needs of the user are not necessarily related to current events.

Social Media Mode: News Use Triggered by Social Media

News use triggered by social media can be considered as incidental exposure. In most cases, it is not planned but an accepted part of engagement with social media (Schäfer, Sülflow, & Müller, 2017). In recent years, social media has become an integral part of online news distribution and consumption, and media organizations are actively seeking to facilitate the wider dissemination of their content not only by distributing their content on social media themselves but also by enabling users as secondary gatekeepers to re-disseminate their content via social media (Singer, 2014).

Since not all news articles are shared equally—Bright (2016) found that prominent positioning or the existence of an accompanying image can influence an article's social media share as well as the subject matter of the article itself—the agenda of a news diet dominated by social media can differ from the agenda experienced by users that rely on traditional online news (routine mode). Instead of being exposed to an extended selection of news stories to choose from a news home page, single news items are provided to news users as part of their general newsfeed.

News users appreciate the serendipity of news they encounter through social media (Fletcher & Nielsen, 2017), and some users even actively customize the flow of news and political content on their newsfeeds (Wells & Thorson, 2017). Yet news use on social media is still distinct from news use as part of an information search or routine browsing session; this is for three main reasons. First, if information is shared by personal contacts, their endorsement has a significant impact on the motivation to engage with a message that goes beyond merely clicking a link (Bond, Settle, Fariss, Jones, & Fowler, 2017). Second, in this process, the news source loses significance as a selection criterion or is not even recognized at all (Messing & Westwood, 2014). Third, the news on social media is curated by algorithms that prioritize news shared by close peers and rely on algorithmic filtering that has implications for the breadth and depth of the news disseminated through social media news feeds (Bozdag, 2013).

Factors of News Modes

There are several factors that we need to consider when explaining user preference for each of the three different modes. Based on the past and current research on news channel selection (Bleke-saune, Elvestad, & Aalberg, 2010; Knobloch-Westerwick, Brück, & Hastall, 2006), we incorporate

into our study motivational factors and attitudes toward news and politics as well as more traditional demographic variables. It is important to note that the rationale for many of these factors are similar across the modes; however, since we are focusing on relative preference for the modes, we do expect that some mechanisms are more pronounced in specific modes than others.

Political Interest and Political Efficacy

Political interest is the most important variable to consider because it motivates users to engage with news both as part of a daily routine and incidentally. However, we expect it is of even higher importance in the context of algorithmically curated news on social media due to the design of the algorithm that selects news as part of an individual's newsfeed (Wells & Thorson, 2017). Users who frequently use news and follow peers interested in current affairs send signals to the algorithm that they enjoy this specific type of content and would, therefore, be more likely to receive it in the future. As news items are provided in a context that facilitates further engagement with a message, we also expect that news use via social media is associated with higher levels of political efficacy (Moeller & de Vreese, 2015).

To a lesser extent, we also expect political interest to be associated with a preference for news use through search for two reasons. First, it can be argued that users with high levels of political interest are more likely to formulate queries that are related to current events. Second, even if the query was not intended to discover news items, users with higher levels of political interest are more likely to engage with a proposed news item.

Political ideology and ideological extremism. Following the filter bubble argument as described above, that political ideology, and more specifically, ideological extremism is associated with a preference for news accessed through social media, we argue that users with a more extreme political ideology tend to prefer receiving their news through social media by means of social curation via their peers to avoid what they consider dissonant voices in the mainstream media.

Trust

In contrast, we expect that trust in news media is related to a preference for routine news use due to a mutually causal relationship. Users choose to engage with a news medium over a longer time span because they trust the coverage they find there. In addition, due to the continued use of a trusted news medium, further trust is established under the condition that users are satisfied with the coverage (Lee, 2010). Trust plays a significantly less important role in the context of news use on social media, in fact, according to Messing and Westwood (2014), many are not even aware of the source of the articles they engage with.

Control variables. Research on news audiences has coherently demonstrated that gender, age, and level of education correlate with modes of news use (Blekesaune et al., 2010; Knobloch-Westerwick et al., 2006; and in the Dutch context, Trilling & Schoenbach, 2013 offline as well as online. We include these variables as statistical controls to account for potential spurious relationships. We also include a self-reported measure of news use offline or on mobile devices not tracked during our study.

Method

In order to understand the factors present in predicting different news modes, we rely on a research design that combines tracking data collected between January 2017 and February 2018 in the Netherlands with survey data from 302 respondents.[1]

Data Collection

Online news consumption patterns were collected from a representative sample of Dutch respondents obtained through a panel company (CentERdata). Compared to the Dutch population as a whole, the sample is slightly older and has a lower than average income (see Table 1 of the Online Appendix). Respondents were informed about the goals of the project, its methods, and the extent of data collection and its privacy protection protocol, and were asked whether or not they agreed to install a traffic-monitoring plug-in (Bodó et al. 2017). They were allowed to opt out at any time. About 712 participants had installed the plug-in. From those that installed the plug-in, 573 (80.5%) completed the online survey in November 2016. Nearly 302 of the 573 (53%) visited news websites at least once and produced data during the field period. Respondents that agreed to participate installed a Google Chrome or Mozilla Firefox plug-in which, for an a priori determined set of 317 white-listed domain names, tracked all incoming and outgoing traffic. The white list was compiled by the authors, and experts on the Dutch media system using audience reach data as input. An additional blacklist indicated URL patterns of exceptions to the white list such as bank-transaction pages, personal information pages, and so on. The plug-in would route all HTTP/HTTPS traffic related to these domains through a secure VPN proxy that served as a data collection point. HTTPS traffic was decrypted and re-encrypted to allow for analysis. As such, all webpage content, external libraries, images, and banners as well as all user-provided information such as search terms were captured. Before storage, best-effort anonymization scripts would remove sensitive information such as bank account numbers. In this study, we consider only the URLs visited by respondents, rather than the news content itself, as our primary interest in the pathways through which users access news. This is because URL patterns provide the most straightforward operationalization of news consumption modes irrespective of technical differences in webpage construction by individual news outlets. Demographic information, self-reported news consumption, political interest, political efficacy, and political extremism data were collected through an online questionnaire run by the panel company after the plug-in was installed.

Scales Used

Political interest was measured on a 7-point Likert-type scale ranging from fully disagree to fully agree to the item "Generally speaking, how interested are you in politics".

Political efficacy was measured using a 3-item, 7-point scale ($\alpha = .73$, $\lambda_6 = .69$), that asked "to what extent do you agree with the following statements?" The following responses were provided: "I have a good idea about the most pressing political problems in our country," "politics is so complicated most people can't understand what is going on," and "I'm better informed about politics than most people in the Netherlands."

Political position was measured as a single-item scale, asking the question "people generally talk about 'left-wing' and 'right-wing' politics when it comes to political ideas. Where would you place your ideas?" with a 10-point left (0) to right (10) scale. Political extremism was measured by taking the absolute distance from the midpoint political position.

Trust in media was measured using an 8-item, 11-point scale ($\alpha = .91$, $\lambda_6 = .93$), that asked "how much trust do you have in the following news media?" with the following items: "newspapers," "television," "radio," "current affairs shows," "newspaper websites," "blogs," "news on social media such as Facebook," and "other news sites."

Gender, age, and level of education were provided by the market research company. News use was measured using 4 items measured on a 7-point scale using the question "how many days per week do you watch/read": "news broadcast [a] or [b]," "current affairs program or reports," "physical newspaper or news on tablet," and "news on mobile phone." The items formed the basis

for two predictors in analysis "offline news use," covering items 1–3 (TV news, TV reports, newspaper) to measure the use of legacy media mainly offline ($\alpha = .77$, $\lambda_6 = .70$) and item 4, mobile news use. These scales meant to cover media use unobserved by our tool and functions as a control variable in the analysis. Means and standard deviations of all scales can be found in Table 1 in the Online Appendix to this article.

Operationalizing Modes of News Consumption

Visits to a website are operationalized based on each page shown to a user as a result of some navigation action such as clicking a link or entering a new URL into the browser's navigation bar. For the purpose of this study, we define four types of visits relevant to our research questions: (1) news site visit, simply a visit to a news site, either to the front page or another page on the website's the domain; (2) social media visit, a visit to a social media site, such as Facebook, Twitter, LinkedIn, or Instagram, occurs any time a respondent opens a window to these pages; (3) generic search, a visit to a search page, includes the most popular search engines in the Netherlands, such as Google, Yahoo, and DuckDuckGo; and (4) other, which are visits to any other type of site and not of direct interest to this study.

We then define access modes as specific sequences in which these visit types occur. We deliberately focus on how users encounter news rather than on audiences shared between sites (see Mukerjee et al., 2018, for a study that takes a similar approach and conceptualizes digital news consumption in network terms). For instance, a person who would visit two news articles would generate the sequence "news → news." The core interest of this article lies in the modes of news engagement. Each user session consists of a number of visits of different types. A user who browses Facebook, goes to a news article, looks up additional information via Google, and ends up on Wikipedia will generate the sequence "social media → news → search → other." Modes were calculated for all users that had visited at least one news site[2] in the period under study and that had responded to the survey, this included a total of 302 users. We operationalize the three news consumption modes as follows:

Routine mode is a mode in which news is consumed without specific triggers related to the use of intermediaries. Generally, these are sequences in which respondents begin a browsing session with a news site or come from any nonintermediary site to a news site in order to be updated about general news events. We define this mode as any pattern in which respondents land on the main page of a news site (the "index" page). Even though the number of occurrences was limited (only 0.53% of all browsing sessions), this operationalization also includes participants that employ a news site as their browser starting page since a general willingness to get an overview of news at the start of a session can be assumed by settings one's home page to a news site.

Social media mode occurs when interest in news is the result of content encountered on social media. Clear examples are news articles linked to by social media or news events discussed by friends. We define this mode as a sequence in which a visit to a news article is directly preceded by a visit to a social media page. Note that it excludes patterns in which the visited news site is the main page of the news domain. In such sequences, a relation between previous social media use and news use can simply be a switch between social media use and routine news use.

Search mode denotes the information-seeking mode of news use in which respondents arrive at news articles through search engines. This does not include the use of search engines as simple navigational tools, as we exclude patterns where search to news sequences first lead respondents to the index page of the news website. For example, a respondent who searches Google for influenza and encounters an article on a flu epidemic performs a search mode pattern. By contrast, a respondent who enters "new york times" as a search term into a search engine instead of typing nytimes.com into the browser's address bar does not perform a search mode pattern.

Table 1. Results of Zero-Inflated Poisson Model of Relative Share of Mode Occurrence.

	Routine Mode		Social Media Mode		Search Mode	
	Zero	Count	Zero	Count	Zero	Count
Intercept	−0.95 (0.46)*	3.99 (0.03)***	0.57 (0.39)	3.32 (0.06)***	−0.73 (0.46)	3.93 (0.03)***
Political interest	−0.10 (0.21)	−0.08 (0.01)***	0.21 (0.18)	0.11 (0.03)***	−0.08 (0.22)	0.05 (0.01)***
Political efficacy	−0.05 (0.20)	0.06 (0.01)***	−0.21 (0.17)	−0.05 (0.02)*	0.14 (0.22)	−0.06 (0.01)***
Political position (1 = left)	−0.16 (0.16)	−0.02 (0.01)*	−0.04 (0.13)	−0.11 (0.02)***	−0.21 (0.16)	0.02 (0.01)
Political extremism	−0.04 (0.16)	0.01 (0.01)	0.11 (0.13)	0.01 (0.02)	−0.07 (0.16)	0.00 (0.01)
Trust in media	−0.13 (0.15)	0.04 (0.01)***	0.10 (0.13)	−0.20 (0.02)***	0.08 (0.16)	−0.03 (0.01)**
Gender (2 = female)	−0.66 (0.32)*	0.16 (0.02)***	−0.17 (0.26)	−0.33 (0.04)***	−0.71 (0.34)*	−0.18 (0.02)***
Age	0.18 (0.21)	0.01 (0.01)	0.17 (0.17)	0.13 (0.03)***	0.20 (0.22)	0.03 (0.01)*
Education level	−0.02 (0.10)	−0.02 (0.01)**	−0.11 (0.08)	0.00 (0.01)	−0.10 (0.10)	−0.02 (0.01)*
Offline news use	−0.19 (0.21)	−0.00 (0.01)	−0.03 (0.17)	−0.03 (0.02)	0.11 (0.21)	0.01 (0.01)
Mobile news use	−0.08 (0.16)	0.09 (0.01)***	−0.16 (0.13)	−0.01 (0.02)	0.09 (0.16)	−0.12 (0.01)***
AIC	5,646.65		3,832.10		6,817.03	
Log likelihood	−2,801.33		−1,894.05		−3,386.51	
Number of observations	302		302		302	

Note. *p < .05. **p < .01. ***p < .001.

Figure 1 shows typical navigation sequences for the three modes described. In the routine mode, the user accesses refdag.nl, a news site, and browses articles there for around 4 min before subsequently navigating to various other Dutch news websites. In the social media mode, the user begins with Facebook.com followed by De Correspondent, an online only news website, revisits Facebook, and then the websites of the Guardian and Vice. In the search sequence, one can observe a user alternating between google search queries and various news websites.

Analysis

These three different modes serve as the main dependent variables for our regression models. Because we were interested in the relative engagement of each mode, the variables were constructed by extracting all relevant news-related clickstream sequences in accordance with the operationalization principles stated above and calculating the relative share of each mode in comparison to the other two. The resulting relative percentage for each mode of news consumption can be characterized as count data with strong zero inflation, since many respondents don't engage in a particular pattern, whereas others engage at varying levels following a Poisson-like distribution. To account for the zero-inflated nature of the observations, we analyze sequence occurrences using a zero-inflated Poisson model. This model combines two predictive models, one to do a binary zero or nonzero prediction (the "zero model") and a second estimating the number of >0 occurrences (the "count model").[3] The analysis was conducted in R; key packages are cited in the software information section.

Results

The findings are based on a diverse sample (for a detailed information regarding descriptive statistics, see the Online Appendix). Of all 302 respondents, 245 (81%) in one or more search patterns, 239 (79%) engaged in at least one routine pattern, and 151 (50%) in one or more social media patterns, all consumption patterns show a zero-inflated Poisson-like distribution as shown as boxplot in Figure 2, with medians ($median_{routine} = 3$, $median_{search} = 4$, $median_{social\ media} = 0$) all below the means ($M_{routine} = 53.08$, $M_{search} = 13.79$, $M_{social\ media} = 8.83$).

To illustrate the different patterns of news use, we also carried out a k means cluster analysis ($k = 4$, between_SS/total_SS $= 84.5\%$), resulting in relatively distinct clusters of users grouped by similar mode distributions. Figure 3 displays the results. The composition of the four clusters ($n = 87, 108, 112, 38$) suggests clusters with particularly high shares of search (2), routine (3), and social media mode (4), as well as one cluster in which search and routine modes (but not social media mode) are strongly represented (1).

We ran three zero-inflated Poisson models on the rate of the modes to explain the occurrence and frequency in which they occurred (e.g., the routine mode, social media mode, and search mode). The results are displayed in Table 1 and include the parameters that model the zero or nonzero occurrence of patterns (listed as "zero model" parameters) and the estimates for the greater-than-zero parameters (listed as "count model" parameters). Predicting whether or not users engage in particular modes is challenging. Only gender significantly predicts routine and search modes, but not social media mode.

The count model estimates how often respondents engage in modes provided they engage at least once (the outcome of the zero model). Here, we see several significant predictors of preference for specific modes, implying that engagement in modes follows a number of different logics.

Our findings support our expectations with regard to the effect of political interest. News users with higher levels of political interest are more likely to encounter news through social media and to a lesser extent triggered by an information search. Interestingly, and contrary to our expectations,
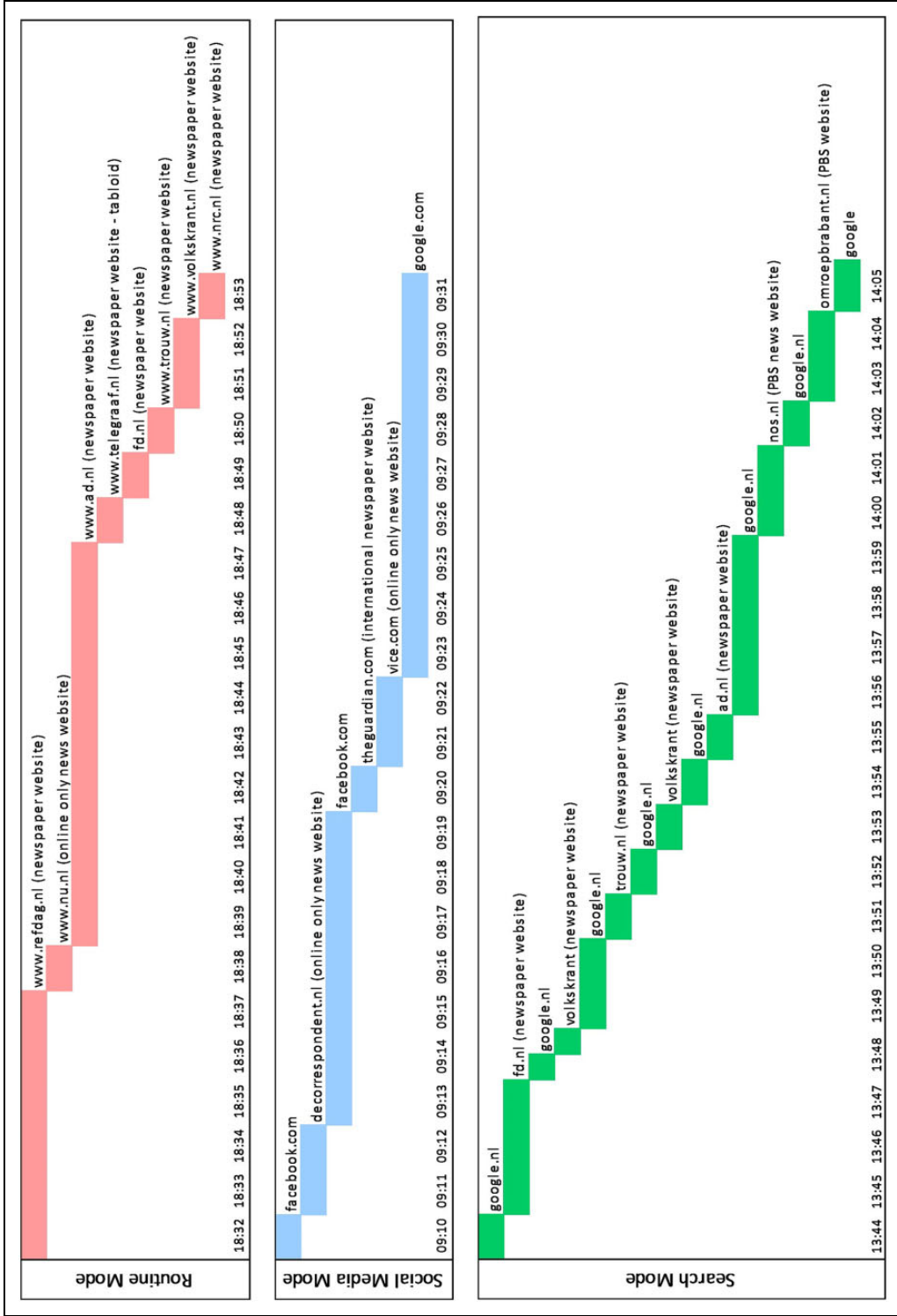
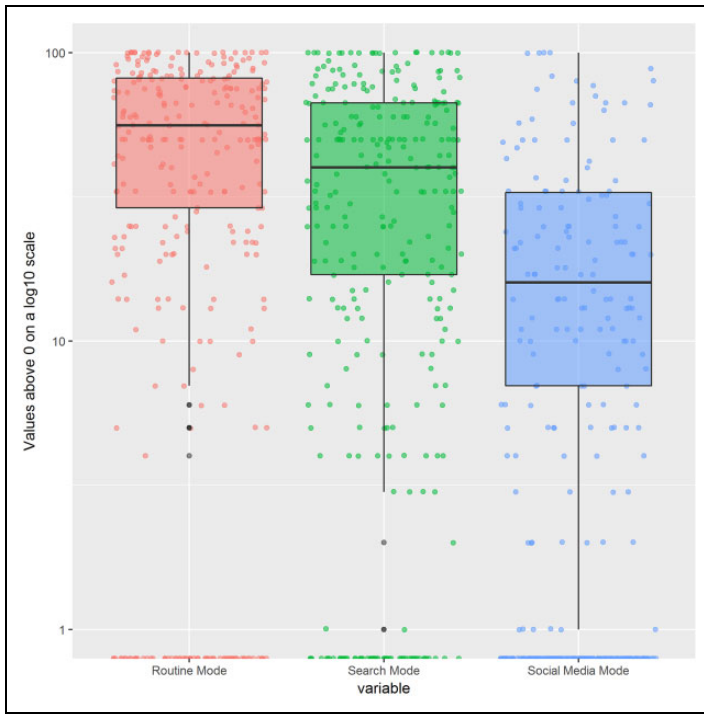**Figure 1.** Exemplary navigation sequences for the modes of online news consumption.

Figure 2. Boxplots of the distribution of nonzero mode engagement on a log 10 scale.
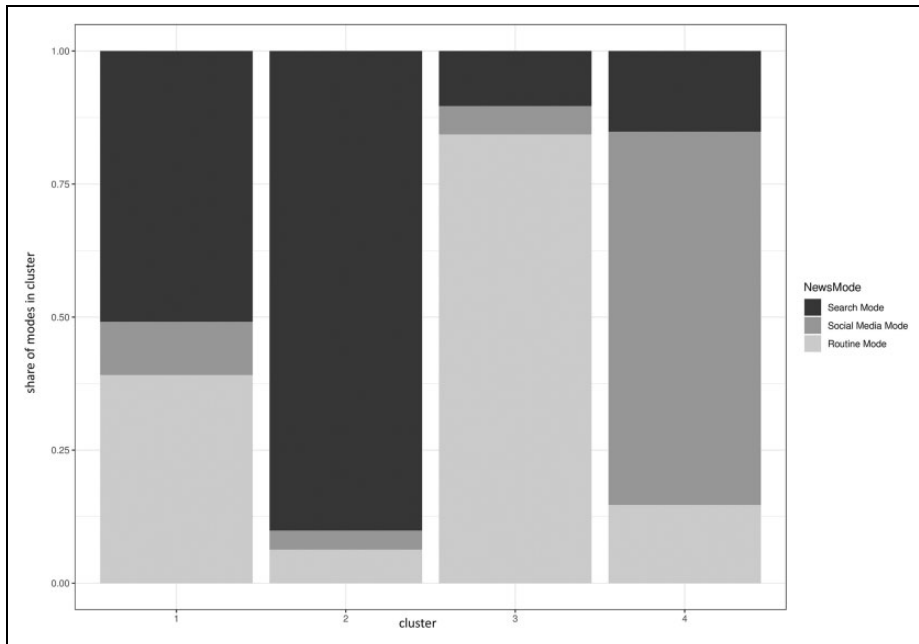

Figure 3. Four user clusters based on mode distributions ($n = 87, 108, 112, 38$).

with regard to political efficacy, we find the reverse, we find that a preference for routine news use is associated with higher levels of political efficacy, whereas the effects of efficacy on preference for news use stimulated by social media and search are negative.

We expected ideological extremism to be associated with a preference for the social media mode. This is not what we find in our data. In fact, we find that political extremism is not associated with any preference for news modes. However, we do find that people on the left of the political spectrum have a preference for news gathered through social media.

In line with our expectations, we find a positive relationship between trust in news and preference for the routine mode. We find a relatively strong negative effect of trust in media on preference for news gathered through social media, implying that the more users distrust media the more likely they are to receive news through social media.

We find a negative relationship between (albeit mostly insignificant) offline and mobile news use and exposure to news in all modes. While this result seems counterintuitive, we interpret it as a replacement effect. If users get their news through other sources, in particular their mobile phone, they are less likely to do so on the laptop on which our tracking tool was installed.

Overall, far fewer predictors seem to significantly predict whether respondents engage with modes at all (2 of 30 potential relations in the zero model) than they do when trying to predict how often modes occur (20 of 30 potential predictors in the count model).

Discussion

In this study, we relied on a unique data set combining the tracking data of browsing sessions with survey data to gain a better understanding of different modes of news use online. To put it more specifically, we were interested in three different modes: routine news use, news use stimulated by social media use, and news use as part of search. Overall, news use triggered by search is the most common mode of news use in our sample, both in terms of general reach and if we consider the median. However, the arithmetic mean for routine news use is much higher which implies that the extent to which users engage in routine news use online is fairly dispersed. Some users visited the home page of news organizations very frequently, while more than half of our participants visited those sites less often than they engaged with news they received through search engines. We can conclude, then, that only few users develop news routines mirroring patterns of offline news consumption. According to our data, the majority of news use has become more incidental as part of their general information searches and, to a lesser extent, their social media experiences.

Social media has been heralded as a conduit to incidental exposure to news. When it comes to news consumption resulting from social media streams, our findings show that social media drives news consumption particularly for those with existing political interest. This suggests that those with low political interest are, in general, not using social media as a substitute entry point to news. Either these users are unlikely to read beyond quoted headlines and teasers, or users with low political interest have less exposure to news due to a less news-rich social circle. In both cases, social media does not provide access to news.

Another interesting finding of our study was the effects political ideology and political extremism have on news use. According to the echo chamber or filter bubble argument, news users that have an extreme political ideology prefer to access news through algorithmically curated services such as social media to avoid challenging information. In harmony with other recent research (Beam, Hutchens, & Hmielowski, 2018; Flaxman et al., 2016; Haim et al., 2018), we find no evidence of this behavior. While news use through social media did not actively contribute to extreme political positions, we conclude that it is not related to less extreme political ideology either. Relatedly, our results indicate that news users inhabiting the left of the political spectrum are more likely to get their news through social media or as part of their general search.

It is important to note that the direction of the causality between preference for a certain mode and political ideology or extremism is not entirely clear-cut. In fact, it can be expected that the underlying process is characterized by mutual causality (Slater, 2007). A similar argument can be made for the effects of political efficacy and interest. For example, political interest may drive knowledge search that increases self-perceived efficacy, but efficacy may also drive interest because the background knowledge aids the interpretation of new events.

Trust in news is an important factor in establishing the routine of using news directly on the home pages of news organizations. Conversely, this finding indicates that those who have less trust in journalism and news organizations opt out of accessing news through the platform itself and make use of algorithmic selection (Thurman, Moeller, Helberger, & Trilling, 2018).

It is important to note that our findings should be interpreted with caution due to several limitations that came about as a consequence of the way our data were collected and analyzed. First, we assume that patterns of behavior are aligned with specific modes. However, some types of user behavior may not reflect modes in the way that we anticipate. A simple but important issue is that opening multiple browser tabs reflects a sequence in our conceptualization even when multiple browsing threads are pursued in parallel. Furthermore, users may generally rely on social media to find news so that their approach does not reflect serendipitous encounters but the default way of finding news. Also patterns may be generated by errors or a lack of computer literacy rather than through deliberate choice. Second, due to limited resources and technical constraints, for this study, it was only feasible to track online behavior on desktop or laptop computers. This means that consumption patterns on mobile devices such as phones and tablets are excluded. In our analyses, we have included the use of mobile applications as a covariate to partially control for this omission. Third, respondents might have adapted their behavior to the fact that they were under observation because of reasons of social desirability. To account for this issue, we tracked over a relatively long period of time to allow participants' time to adjust. Fourth, the relatively small sample of users limits the statistical power of our models.

Future work that leverages larger samples would enable a more granular analysis based on clusters of users based on recurring activity. In its current iteration, this study does show significant differences in the predictors of different news consumption modes that warrant further exploration. Finally, some of our findings may be specific to the media and political context of the Netherlands. As both news consumption patterns and adoption of social media differ widely across Europe, other countries may also differ with regard to common pathways to news.

Taking these limitations into account, the results presented in this article provide ample ground for future work on the routes users take when consuming news online. Each mode of news consumption seems to be just as driven by different logics as it is by the significant predictors of preference for the different modes. In future research, the effects of different news consumption modes can be examined for similar dissimilarities such as different propensities to share news, diversity in content consumed, and differences in levels of engagement. It would also be worthwhile to gain insight into the interrelation between news modes and news content, for example, hard news and infotainment. In addition, the changes in consumption modes and their relation to political efficacy, interest, and ideological position should be further explored in a longitudinal setting that can point to probable directions of effects. Future work is also needed if we are to shed light on the role news portals play in the routine use of news, which should be embedded in a more conceptual discussion of routine behavior online.

Our findings provide an important first step toward a more systematic research program on the interrelation of news consumption and sociodemographic variables in concert with social and psychological factors such as attitudes, beliefs, and political leanings. Observing news use through digital trace data, as we have done in this study, has three core advantages: (1) It enables precise exposure measurement beyond what is traditionally possible, (2) it allows us to understand news use

as a part of general web browsing, rather than as an isolated activity, and (3) it facilitates the study of usage patterns in tandem with survey research to achieve a holistic understanding of news engagement grounded in a rich social context.

Authors' Note

Data Availability

Replication files can be obtained and replication analyses carried out on a secure University of Amsterdam server by applying to the directors of the Personalised Communication project, Prof. Dr. Natali Helberger and Prof. Dr. Claes de Vreese.

Declaration of Conflicting Interests

Funding

Supplemental Material

Supplemental material for this article is available online.

Notes

1. This article makes use of data from the Longitudinal Internet Studies for the Social sciences (LISS) panel administered by CentERdata (Tilburg University, the Netherlands). Details regarding the recruitment of the participants can be obtained at https://www.lissdata.nl/about-panel/sample-and-recruitment
2. We did not include those that had not visited a single news site to account for measurement error. The measurement error could occur if users were not able to use the browser plug-in sufficiently for technical reasons, or if they chose not use the plug-in for news for privacy or other reasons.
3. To identify potential multicollinearity, we inspected the Variance Inflation Factor (VIF). The VIF for all variables in all models were smaller than 3.

References

Beam, M. A., Hutchens, M. J., & Hmielowski, J. D. (2018). Facebook news and (de) polarization: Reinforcing spirals in the 2016 US election. *Information, Communication & Society*, *21*, 940–958.

Blekesaune, A., Elvestad, E., & Aalberg, T. (2010). Tuning out the world of news and current affairs—An empirical study of Europe's disconnected citizens. *European Sociological Review*, *28*, 110–126.

Bode, L. (2016). Political news in the news feed: Learning politics from social media. *Mass Communication and Society*, *19*, 24–48.

Bodó, B., Helberger, N., Irion, K., Zuiderveen Borgesius, F., Moller, J., van de Velde, B., . . . de Vreese, C. (2017). Tackling the algorithmic control crisis - the technical, legal, and ethical challenges of research into algorithmic agents. *Yale Journal of Law and Technology*, *19*, 133–180.

Bond, R. M., Settle, J. E., Fariss, C. J., Jones, J. J., & Fowler, J. H. (2017). Social endorsement cues and political participation. *Political Communication*, *34*, 261–281.

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, *15*, 209–227.

Bright, J. (2016). The social news gap: How news reading and news sharing diverge. *Journal of Communication*, *66*, 343–365.

Choo, C. W., Detlor, B., & Turnbull, D. (2000). Information seeking on the web: An integrated model of browsing and searching. *First Monday*, *5*.

Costera Meijer, I., & Groot Kormelink, T. (2015). Checking, sharing, clicking and linking: Changing patterns of news use between 2004 and 2014. *Digital Journalism*, *3*, 664–679.

de Vreese, C. H., & Neijens, P. (2016). Measuring media exposure in a changing communications environment. *Communication Methods and Measures*, *10*, 69–80.

Diakopoulos, N., & Koliska, M. (2017). Algorithmic transparency in the news media. *Digital Journalism*, *5*, 809–828.

Diddi, A., & LaRose, R. (2006). Getting hooked on news: Uses and gratifications and the formation of news habits among college students in an Internet environment. *Journal of Broadcasting & Electronic Media*, *50*, 193–210.

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, *80*, 298–320.

Fletcher, R., & Nielsen, R. K. (2017). Are people incidentally exposed to news on social media? A comparative analysis. *New Media & Society*, *20*, 2450–2468. [Online first]

Gil de Zúñiga, H., Weeks, B., & Ardèvol-Abreu, A. (2017). Effects of the news-finds-me perception in communication: Social media use implications for news seeking and learning about politics. *Journal of Computer-Mediated Communication*, *22*, 105–123.

Graber, D. A. (1988). *Processing the news: How people tame the information tide*. New York, NY: Longman.

Guess, A., Nyhan, B., & Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. Retrieved from https://www.dartmouth.edu/~nyhan/fake-news-2016.pdf

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*, 330–343.

Hasebrink, U. (2004). Konvergenz aus Nutzerperspektive: Das Konzept der Kommunikationsmodi [Convergence from an user perspective: The concept of communication modes]. In U. Hasebrink, L. Mikos, & E. Prommer (Eds.), *Mediennutzung in konvergierenden Medienumgebungen* [Media use in converging media environments][In German] (pp. 67–88). Munich, Germany: Reinhard Fischer.

Hasebrink, U., & Hölig, S. (2017). Deconstructing audiences in converging media environments. In S. Sparviero, C. Peil, & G. Balbi (Eds.), *Media convergence and deconvergence* (pp. 113–133). Cham, Switzerland: Palgrave Macmillan.

Jackman, S. (2017). *pscl: Classes and methods for R developed in the Political Science Computational Laboratory*. Retrieved from https://cran.r-project.org/web/packages/pscl

Keyes, O., Jacobs, J., Schmidt, D., Greenway, M., Rudis, B., Pinto, A., . . . Jiang, X. (2018). *urltools: Vectorised tools for URL handling and parsing*. Retrieved from https://cran-r-project.org/package=urltools

Knobloch-Westerwick, S., Brück, J., & Hastall, M. R. (2006). The gender news use divide: Impacts of sex, gender, self-esteem, achievement, and affiliation motive on German newsreaders' exposure to news topics. *Communications*, *31*, 329–345.

Kubey, R., Larson, R., & Csikszentmihalyi, M. (1996). Experience sampling method applications to communication research questions. *Journal of Communication*, *46*, 99–120.

Lee, T. T. (2010). Why they don't trust the media: An examination of factors predicting trust. *American Behavioral Scientist*, *54*, 8–21.

Messing, S., & Westwood, S. J. (2014). Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication Research*, *41*, 1042–1063.

Mitchelstein, E., & Boczkowski, P. J. (2010). Online news consumption research: An assessment of past work and an agenda for the future. *New Media & Society*, *12*, 1085–1102.

Moeller, J., & de Vreese, C. (2015). Spiral of political learning: The reciprocal relationship of news media uses and political knowledge among adolescents. *Communication Research*. doi:0093650215605148

Mukerjee, S., Majó-Vázquez, S., & González-Bailón, S. (2018). Networks of audience overlap in the consumption of digital news. *Journal of Communication*, *68*, 26–50. [Online first]

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D., & Nielsen, R. (2018). *The Reuters institute digital news report 2018*. Oxford, England: Reuters Institute for the Study of Journalism. Retrieved from http://www.digitalnewsreport.org/

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. London, England: Penguin.

Poindexter, P. M., & McCombs, M. E. (2001). Revisiting the civic duty to keep informed in the new media environment. *Journalism & Mass Communication Quarterly*, *78*, 113–126.

Prior, M. (2009). The immensely inflated news audience: Assessing bias in self-reported news exposure. *Public Opinion Quarterly*, *73*, 130–143.

Prior, M. (2013). The challenge of measuring media exposure: Reply to Dilliplane, Goldman, and Mutz. *Political Communication*, *30*, 620–634.

R Core Team. (2017). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria*. Retrieved from https://www.R-project.org/

Revelle, W. (2017). *psych: Procedures for personality and psychological research*. Retrieved from https://CRAN.R-project.org/package=psych

Schäfer, S., Sülflow, M., & Müller, P. (2017). The special taste of snack news: An application of niche theory to understand the appeal of Facebook as a source for political news. *First Monday*, *22*. doi:10.5210/fmv22i4.7431

Schrøder, K. C. (2015). News media old and new: Fluctuating audiences, news repertoires and locations of consumption. *Journalism Studies*, *16*, 60–78.

Shokouhi, M., & Guo, Q. (2015). *From queries to cards: Re-ranking proactive card recommendations based on reactive search history*. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 695–704) Santiago, Chile: ACM.

Slater, M. D. (2007). Reinforcing spirals: The mutual influence of media selectivity and media effects and their impact on individual behavior and social identity. *Communication Theory*, *17*, 281–303.

Singer, J. (2014). User-generated visibility: Secondary gatekeeping in a shared media space. *New Media & Society*, *16*, 55–73.

Sunstein, C. R. (2001). *Echo chambers: Bush v. Gore, impeachment, and beyond*. Princeton, NJ: Princeton University Press.

Taneja, H., & Webster, J. G. (2016). How do global audiences take shape? The role of institutions and culture in patterns of web use. *Journal of Communication*, *66*, 161–182.

Tewksbury, D. (2005). The seeds of audience fragmentation: Specialization in the use of online news sites. *Journal of Broadcasting & Electronic Media*, *49*, 332–348.

Tewksbury, D., Hals, M. L., & Bibart, A. (2008). The efficacy of news browsing: The relationship of news consumption style to social and political efficacy. *Journalism & Mass Communication Quarterly*, *85*, 257–272.

Thurman, N., Moeller, J., Helberger, N., & Trilling, D. (2018). My friends, editors, algorithms, and I: Examining audience attitudes to news selection. *Digital Journalism*, Advance online publication. 1–23. doi: https://www.tandfonline.com/doi/full/10.1080/21670811.2018.1493936

Trilling, D., & Schoenbach, K. (2013). Skipping current affairs: The non-users of online and offline news. *European Journal of Communication*, *28*, 35–51.

632 Social Science Computer Review 38(5)

Wells, C., & Thorson, K. (2017). Combining big data and survey techniques to model effects of political content flows in Facebook. *Social Science Computer Review*, *35*, 33–52.

Wickham, H. (2017). *tidyverse*. Retrieved from https://cran.r-project.org/web/packages/tidyverse

Wieland, M., In der Au, A. M., Keller, C., Brunk, S., Bettermann, T., Hagen, L. M., & Schlegel, T. (2018). Online behavior tracking in social sciences: Quality criteria and technical implementation. In C. Stützer, M. Welker, & M. Egger (Eds.), *Computational social science in the age of big data. Concepts, methodologies, tools, and applications*. Cologne, Germany: Herbert von Halem.

Wonneberger, A., Schoenbach, K., & van Meurs, L. (2011). Interest in news and politics - or situational determinants? Why people watch the news. *Journal of Broadcasting & Electronic Media*, *55*, 325–343.

Yadamsuren, B., & Erdelez, S. (2011). Online news reading behavior: From habitual reading to stumbling upon news. *Proceedings of the Association for Information Science and Technology*, *48*, 1–10.

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software*, *27*(8). Retrieved from http://www.jstatsoft.org/v27/i08/

Author Biographies

Judith Möller is a postdoctoral researcher in the Amsterdam School of Communication Research at the University of Amsterdam. Her research interests include media effects, political socialization, the affordances of political information use online, and their consequences for the public sphere.

Robbert Nicolai van de Velde is a data scientist with research experience in the fields of communication science, computer science, and artificial intelligence. He is primarily interested in the intersection of technology and society in the forms of pervasive monitoring, problematic content, and access to information.

Lisa Merten is a junior researcher and a PhD student at the Leibniz Institute for Media Research | Hans-Bredow-Institut in Hamburg. She is interested in patterns of digital news consumption, customization practices in current digital environments and media effects.

Cornelius Puschmann is a senior researcher at the Leibniz Institute for Media Research | Hans-Bredow-Institut in Hamburg where he coordinates the international research network Algorithmed Public Spheres. His interests include online hate speech, the role of algorithms for the selection of media content, and methodological aspects of computational social science.