News Won't Find Me? Exploring Inequalities in Social Media News Use With Tracking Data

LISA MERTEN[1]
Leibniz-Institute for Media Research | Hans-Bredow-Institut, Germany

NADIA METOUI
Delft University of Technology, The Netherlands

MYKOLA MAKHORTYKH
University of Bern, Switzerland

DAMIAN TRILLING
JUDITH MOELLER
University of Amsterdam, The Netherlands

The rise of news content on social media has been accompanied by a hope that people with lower socioeconomic status and less interest in political affairs would be "accidentally" exposed to news. By combining tracking and survey data from a Dutch online panel ($N =$ 413), we analyze how political interest, income, and education influence social media news exposure and consumption. Higher levels of political interest are associated with higher amounts of news exposure on Facebook and more news items consumed via social media. Users engage less often in news-related follow-up behavior after consuming news items via social media than after consuming news items referred via news websites. If social

Lisa Merten: l.merten@leibniz-hbi.de
Nadia Metoui: n.metoui@tudelft.nl
Mykola Makhortykh: mykola.makhortykh@unibe.ch
Damian Trilling: d.c.trilling@uva.nl
Judith Moeller: j.e.moller1@uva.nl
Date submitted: 2020-12-20

media news use seems to occur particularly for those who are already interested in current affairs and makes follow-up consumption less likely, the specificities of the social media ecosystems might accelerate rather than level inequalities in news use.

Keywords: social media, news consumption, political interest, computational methods, survey, digital divide

Individuals privileged in terms of education and income and those already interested in current affairs and politics are more likely to consume and learn from traditional news (Lecheler & de Vreese, 2017). The widening of media choice in post-broadcast democracies further increases inequalities in the use of news media, leading to larger inequalities in knowledge of politics and current affairs among different groups in society (Dahlgren, 2019; Prior, 2007).

The rise of social media platforms as a means of news distribution and reception was initially hailed by some as a mitigating factor: They were thought to facilitate ample opportunities for incidental news exposure or trap effects. Seeing news content on social media would allow particularly "less attentive citizens" (Bode, 2016, p. 29) who would not otherwise engage with news to stumble upon information on current affairs without actively searching for it. Through this kind of incidental news exposure, motivational gaps in (online) political participation can be narrowed (Boulianne, 2018), learning opportunities could be facilitated (Bode, 2016), and the diversity of viewpoints seen might be increased (Fletcher & Nielsen, 2018). However, exposure to news on social media is determined by its visibility in a personalized newsfeed, curated by users and algorithms. Kümpel (2020) and Thorson (2020) have pointed to systematic differences in those who receive news incidentally or intentionally in their newsfeeds and those who do not. Both authors link these findings to research on digital use divides.

Most of the scant empirical work on the exact level of visibility and the differential use of news within social platforms relies on self-reported data, which are prone to overestimating news consumption, especially in comparison with other online activities (Guess, 2015). In particular, politically interested individuals overestimate their news-related activities on Facebook (Haenschen, 2019). In this study, we explore whether social media news use is mitigating or further amplifying the divide in news use by combining tracking and survey data. We specifically analyze differences in social media news exposure and consumption related to political interest, education, and income. Our unique approach to data collection and analysis allows us to compare the consumption of news items accessed via social media with news items consumed via news websites. Before proceeding to examine the data, we discuss how the dynamics of digital divides and filtering processes on social media platforms might influence social media news exposure and consumption.

Inequalities in (Online) News Use

With the rise of digital technology, an inequality in digital news consumption has been conceptualized as a "digital divide" (e.g., Hargittai, 2002). Whereas early studies focused on the gap between "those who have access to digital technologies and those who do not" (Hargittai, 2002, p. 822),

more recent work examines differences in practices and literacy: Wei and Hindman (2011) demonstrated that the effect of educational level on information-oriented versus entertainment-oriented Internet use is actually greater than the effect of educational level on access to the Internet itself. These inequalities in access and practices translate into differences in outcomes, in particular differences in political knowledge, as explained by the knowledge gap hypothesis (Tichenor, Donohue, & Olien, 1970).

The antecedents of the widening news usage and knowledge gap online are more controversial than the existence of such disparities. It can be argued that inequality in information exposure is a consequence of factors at the level of the individual, such as political interest and education (Strömback, Djerf-Pierre, & Shehata, 2013), and contextual factors, specifically the availability of media content. If a greater variety of individual use patterns is possible, educational background and motivational status are more relevant for the choice of content leading to a greater potential for digital inequality (Prior, 2007). Conversely, this argument implies that before today's diversified media environment, television and newspapers were able to "trap" media users with low interest in politics into exposure to news and that a high-choice environment such as the Internet may not be as effective in trapping audiences that would rather avoid news (see Mahrt, 2019, for a summary). Empirical research does not unequivocally support this assumption (e.g., Trilling & Schönbach, 2015).

From a sociological perspective, the origins of the news gap can also be understood as a consequence of the expression of social class or "socially inherited taste palettes" (Ohlsson, Lindell, & Arkhede, 2016, p. 3). Referring to Bourdieu's (1984) notion of cultural consumption as structured by social class, Lindell (2018) and Ohlsson and colleagues (2016) argue that class is an important yet largely overlooked variable affecting the fragmentation of online news consumption and that overlaps in online news preferences and practices are strongly linked among citizens with similar economic, cultural, and social capital. Since different groups form altogether different news repertoires, these news practices and preferences have the potential to solidify the positions of groups in the social structure (Bourdieu, 1984; Lindell, 2018). To better understand inequalities in social media news use, we incorporate not only political interest as a motivating factor and education as both an ability-related factor and indicator of cultural capital, but also include income as an indicator of economic capital. Income can also be a relevant factor in online news consumption merely concerning access because of the rise of paywalls.

Regardless of the exact causes of the news gap, inequalities in knowledge tend to lead to exclusion from social resources and inequalities in political participation (Delli Carpini & Keeter, 1996). Major group differences in terms of news use and political participation might pose risks to the very foundations of democratic societies that rely on equal participation and representation among all social groups (Habermas, 1989). The emergence of social media platforms as part of a relatively low choice online environment thanks to their algorithmic filtering functions was seen as an opportunity to level divides in news use (Bode, 2016).

News Divides on Social Media

Information intermediaries such as social networking platforms are considered main drivers for incidental news exposure (Newman et al., 2021). Yet recent studies indicate that the influence of incidental exposure is not universal. Maybe only those already interested in politics benefit from accidentally

encountering news on social media. Woo-Yoo and Gil de Zúñiga (2014) demonstrated that highly educated U.S. citizens who often used Facebook to read political content knew more about current affairs and political structures than similarly educated users with less Facebook usage. However, lower educated users who used Facebook for news were frequently found to be even less knowledgeable than similarly educated users who did not use Facebook for news very often. For users with low political interest, exposure to incidental news on Facebook may even lead to an illusion of knowledge rather than factual knowledge (Feezell & Ortiz, 2021).

These results are especially disconcerting in light of research related to source replacement effects. The belief in being informed via social media without having to actively search for news ("news-finds-me" perception) can decrease news consumption outside of social media (Gil de Zúñiga et al., 2017, p. 105). Overall, social media do not appear to be particularly useful tools for compensating inequalities in news use via incidental exposure.

Because of the hyperlinked architecture of social media, being exposed to a news story does not equate to actually consuming, processing, or retaining its content. To examine social media news environments in terms of their gap leveling potential, a distinction between the mere exposure and the actual consumption of news content is necessary and possible with digital trace data. We distinguish between (1) mere social media news exposure as in seeing a link to news content in the Facebook newsfeed; (2) the actual consumption of news content as in clicking on a link to a specific news item; and then (3) explore follow-up news consumption after the first piece of news content. We argue that motivational and sociodemographic factors can contribute to news gaps at all three stages and ask how these factors might play out differently from other forms of news use because of the specific characteristics of social media platforms.

Inequalities in Social Media News Exposure

The emergence of social media was seen as a chance to overcome a gap in news use because they are online news spaces that are attended more equally than news websites by citizens with differing levels of education, income, and political interest. Kalogeropoulos and Nielsen (2018) found that individuals in the United Kingdom with lower socioeconomic status are significantly less likely to access news via news websites. However, they found no differences regarding the likelihood of relying on distributed forms of discovery for news on social media. But, even though access to social media might be more balanced than it is through news websites, the algorithmically filtered news streams on those platforms differ fundamentally from news websites curated by editors where a (largely) uniform newsfeed is still the norm. Not everybody is afforded equal opportunity to come across news in algorithmic media environments.

Since the multitude of algorithms shaping social media newsfeeds is not made public, identifying factors determining the inclusion of news within the newsfeed is challenging. However, some basic algorithmic functionalities are known from instances of professional communication and patents related to the platforms (DeVito, 2017). First, users create deliberate signals themselves through *personal curation or customization* by following specific (news) accounts or adjusting their newsfeed settings to see more or less content from one source and, therefore, predetermining the chance of news exposure independent from their social contacts (Merten, 2020a).

Second, previous *interactions with news content* on platforms such as clicking on articles or sharing items can determine a categorization as interested in news by the newsfeed algorithms and, therefore, news exposure. Thorson, Cotter, Medeiros, and Pak (2019) found that respondents with high levels of education and political interest and income were more likely to be algorithmically classified as interested in politics or news for potential advertisers. Such users are then more likely to self-report exposure to content related to news and politics in their feeds (Thorson et al., 2019). These results suggest that the amount of news exposure on Facebook is also dependent on inferences made by the platform on user behavior and characteristics. The inferences and classifications based on previous behavior cannot only be used by the social media platforms to customize the newsfeed and to suggest additional news-poor or news-rich accounts to follow and groups to join. But news organizations can also directly promote individual articles or target users with other content based on their information on user interests that they get from the platform that could increase an individual's social media news exposure.

Third, we know that, at least for Facebook, the interactions with news content by the *social contacts* of the user (sharing, liking, commenting on news content) are prioritized in the newsfeed (Mosseri, 2016). Previous research has shown that individuals tend to associate and maintain connections with those with similar socioeconomic status (McPherson, Smith-Lovin, & Cook, 2001), political interest, and activity (Knoke, 1990). Social media users who exhibit certain characteristics are more likely to be in contact with similarly oriented individuals on the platforms (Aiello et al., 2012) and should therefore display similar behavior related to news content.

Finally, content on social media is prioritized based on interactions with news content by the overall platform population (*collaborative filtering*, see e.g., Kabiljo & Ilic, 2015). This implies that users are recommended content that other users who share similar characteristics and online behavior have engaged with. For example, if a user engages with a lot of content about a television show and other fans of that show rarely click on news, that user may also see less news.

These dynamics of personal, algorithmic, and social news curation (Thorson & Wells, 2016) can lead to an accumulation of signals that signify either news interest or a lack thereof on the social media platform. We assume that users with higher SES and political interest are more likely to display news-related behavior on social media, have social contacts with interest in news, and are categorized as news-savvy by the site's algorithms and therefore hypothesize:

H1:　　*Users with higher levels of (a) political interest, (b) education, and (c) income have higher levels of news exposure on Facebook than users with lower levels of political interest, education, and income.*

Inequalities in Social Media News Consumption

Even though exposure to news is a prerequisite, most benefits of social media news exposure rely on the actual consumption and processing of news (Oeldorf-Hirsch, 2018). Even if all users had the same opportunity to consume news on social media because of equal news exposure, the factors that prompt

individuals to engage with the content in their social media newsfeeds are highly dependent on thematic interest and social cues.

Similar to traditional media environments, there is evidence that individuals with high levels of political interest are more likely to consume news on social media in general. Karnowski, Kümpel, Leonhard, and Leiner (2017) and Kümpel (2019) identify topical interest as one of the strongest determinants on the intention to read news items and to seek out further information after seeing a news item on Facebook. Just like news exposure, the probabilities of actual news consumption on social media are in favor of those already interested in news who are also predominantly users with higher education, income, and political interest. If a news item on social media is recommended by close contacts, the likelihood of consumption also increases (Kaiser, Keller, & Kleinen-von-Königslöw, 2018). As an important heuristic for content selection, social cues from friends and family can even outweigh partisan selectivity (Anspach, 2017). Since users are more likely to interact and form strong ties with peers with similar socioeconomic backgrounds and interests, users with higher levels of education, income, and political interest are also more likely to receive social cues that motivate them to engage in news consumption.

H2:   Users with higher levels of (a) political interest, (b) education, and (c) income exhibit higher levels of news consumption on social media than users with lower levels of political interest, education, and income.

One feature of social media platforms makes the influence of motivational factors on news consumption there more likely than on news websites: Social media news content is not exclusive but mixed with personal content from friends and family, sponsored posts, and other branded content. This fundamentally differs from the content offered within the website of a news organization, where the choice of non-news-related content is smaller. Individuals who visit news websites are already motivated for news consumption. The decision to engage with news content is happening at the stage of access to the website. Therefore, factors for news consumption such as political interest, education, and income should be relevant when measuring the probability of accessing news websites. On intermediaries such as social media, news content is debundled from its original sources and rebundled together with other content (Schmidt, Merten, Hasebrink, Petrich, & Rolfs, 2019). In general, news use is rather a by-product than the main objective of social media use (e.g., Newman et al., 2021). Motivational factors such as political interest become more relevant at a later stage in the pathway toward news. In a way, social media create the high-choice media environment dilemma as described by Prior (2007) for the mass media system in one newsfeed, with comparable consequences. Those with less motivation to engage with news are more likely to invest their time in other content. One could assume, therefore, that higher interest is necessary to actually choose the news items within the content mix of social media newsfeeds. We test whether political interest is not only an important predictor for news consumption on social media but also if the effect is larger than news websites.

H3:   The gap between users with higher levels of (a) political interest, (b) education, and (c) income is higher for news consumption after article exposure via social media than via a news website.

Inequalities in Extended Engagement With News

To understand possible news divides on social media, it is important to also examine how news exposure serves as a starting point for engagement beyond the consumption of a single news item. If news consumption via social media triggers a cascade of news-related behavior on the social media platform, further news consumption, or a more detailed information search, the information obtained during these more elaborate patterns of news consumption is more likely to be better retained and contextualized and stimulate knowledge gain (Eveland, 2001). Oeldorf-Hirsch (2018) found that participants who report engagement with news snippets on social media beyond mere exposure, such as liking or commenting and sharing, also report greater cognitive elaboration on news stories than those who do not engage with news content on Facebook and Twitter.

Our rich data set also allows us to test not only the actual reading or viewing of news content referred to via social media or a website but also behavior after the consumption of a single news item. To test Karnowski and colleagues' (2017) survey results on the influence of thematic interest on the intention to follow up news consumption with tracking data, we extend the model and measure not only which factors and pathways make consumption more likely but also which of those determine further engagement such as follow-up news consumption. Since, so far, few empirical studies have investigated determinants of information search and follow-up news reading after social media news exposure, our expectations are derived deductively. Political interest, education, and higher SES are positively related to news consumption. We therefore test if the inequalities observed in exposure and consumption extend to follow-up consumption.

H4:   *Users with higher levels of (a) political interest, (b) education, and (c) income have higher levels of follow-up news consumption on social media than users with lower levels of political interest, education, and income.*

H5:   *The gap between users with higher levels of (a) political interest, (b) education, and (c) income is greater for follow-up consumption after article exposure via social media than via a news website.*

Method

Data Collection

Individuals tend to overreport their use of online news websites (Guess, 2015) and cannot accurately recall the frequency of website visits (Revilla, Ochoa, & Loewe, 2017). To address this limitation, we combined survey data on individual characteristics with tracking data. Respondents were recruited via the Longitudinal Internet Studies for the Social Sciences (LISS) panel of CentERdata, which is based on a true probability sample of the Dutch population and were informed about the goals of the project, the extent of data collection and its privacy protection protocol (as well as the possibility to opt out at any time).[2] Seven hundred and twelve panel members were contacted, and 573 responded to the initial survey request.

---

[2] The sample composition is by and large representative of the Dutch population in terms of age, gender, and education level. See Table 4 in the supplemental material on OSF (https://osf.io/e8rf7/) for a

The respondents installed a custom-built Chrome or Firefox plug-in that monitored their online activity on the device they used most, typically their main computer at home. However, it is perfectly possible that participants use multiple devices (e.g., a desktop and a laptop). Our estimations therefore present a lower bound, and it is possible that there are some services that participants use only on a device we did not track. The plug-in tracked all incoming and outgoing traffic for a predetermined set of 317 domain names (i.e., the allowlist). This allowlist was compiled by experts on the Dutch media system using audience reach data and covers all news websites that were featured among the 300 most used websites according to ALEXA data. The list also includes local, regional, international, and alternative news websites not listed by ALEXA, but known to the researchers.

All HTTP/HTTPS traffic was routed via a secure proxy, where traffic not related to the allowlisted domains was filtered out. The content from allowlisted domains went through a set of custom filters to hash sensitive information included in http(s) requests such as e-mail addresses or any Facebook account names associated with the content in the newsfeed. Remaining content was ingested into an ElasticSearch cluster for storage and subsequent analyses by the researchers, who followed a strict security protocol to prevent potential data breaches. Raw HTML was then extracted from the response content field of the captured traffic. For more information on the tracking tool and process, see Vermeer, Trilling, Kruikemeier, and de Vreese (2020) and Moeller, van de Velde, Merten, and Puschmann (2020), and see Bodo et al. (2018) for a detailed account of the ethical and legal aspects that guided the development of the tracking tool. To further protect respondents' privacy, we also used a denylist including nontrackable URLs (for instance, online banking and personal information pages). Participants were asked at the beginning of the tracking period to fill out a survey that included data on their sociodemographic characteristics, media use, information habits, and political interest.

Between June 1, 2017, and June 2, 2018, 413 respondents used the plug-in and took part in the online survey. Only desktop or laptop devices were tracked, and for each given machine, the plug-in was installed on a single browser, so we would not detect usage if participants used other browsers during the tracking period. A major challenge for our analysis was the different forms of (tracked) user activity within the sample. Not all participants used their tracked devices to the same extent or over the whole field period. We therefore calculate news exposure and consumption relative to overall online activity tracked. The variance in observations for the different exposure, consumption, and follow-up models presented in the results section is because of differences in numbers of participants being exposed to news or engaging in different practices of news consumption.

---

comparison of all respondents participating in the survey (N = 573), those who also installed and used the browser plug-in for tracking (n = 413), and those for whom we have data on their Facebook exposure (n = 115) during the field period in terms of all survey variables used plus a comparison with the Dutch population in terms of gender, income, and age.

Operationalization

Independent Variables (Survey Data)

Education was measured in the categories used by the Dutch Office for Statistics: primary education, prevocational education, selective secondary education, middle-level vocational education, university of applied sciences, and university (M = 4.13, SD = 1.43). Respondents were asked to name their numerical pretax income (M = 2541 euro, SD = 2132) and their level of political interest in a single measure ranging from *not at all* (1) to *very much* (7) (M = 4.28, SD = 1.67). We included age (M = 48.17, SD = 17.9) and gender (50.8% female) in our models as they have shown to be consistent predictors of news consumption (Strömback et al., 2013). Compared with the Dutch population (Centraal Bureau voor de Statistiek database, 2017a, 2017b), our sample is slightly older (population 41.6) and similar in gender distribution (population 50.4) and educational level (population 3.95). Beyond our main variables of interest, we also included a self-reported measure of news use on mobile devices not tracked in our study, measures of internal efficacy as a proxy of perceived political knowledge, political extremism, and technical literacy in our models to account for potential interplay discussed in the literature review. An overview of operationalization of all survey variables can be found in Table 4 in the additional online material (https://osf.io/e8rf7/), which also includes a comparison of our survey and tracking samples for each independent variable.

Dependent Variables (Tracking Data)

Facebook News Exposure. News exposure refers to the amount of news content an individual sees in their Facebook newsfeed. During the field period, Facebook was the social media platform most used for news consumption in the Netherlands. In the Reuters Institute Digital News Report 2018, 65% of the participants in the Dutch sample reported using Facebook, and 29% used the platform for news (Costera Meijer & Groot Kormelink, 2018). The URLs and text fields from the Facebook newsfeed were extracted and cross-checked with the list of the 101 allowlisted news domains. The amount of exposure is measured as the percentage of newsfeed entries that included the names of or links to these news domains, which allows us to compare the share of news in each individual newsfeed.

News Consumption. All sites visited by the participants were labeled as either news outlets, social media, or other websites. We rely on a source-based approach and define news as all content originating from a predefined set of 101 allowlisted news websites.[3] Instagram, Pinterest, Tumblr, Twitter, Facebook, and LinkedIn were labeled as social media platforms. The amount of overall news consumption is operationalized as the percentage of news items (e.g., specific news articles or videos) and news websites (e.g., www.nytimes.com) each user clicked through in relation to total online activity measured. To identify different pathways to news consumption, we analyzed the clicking sequences of each individual (Figure 1). If a social media visit was followed by accessing a specific news item, we identified this sequence as social-

---

[3] Table 5 in the supplementary material (https://osf.io/e8rf7/) lists and categorizes all allowlisted news websites in our sample with distribution across URLs clicked and the Facebook newsfeed.

media-led news consumption.[4] If a news item was accessed after a visit to a news website, the online interaction counted as news-website-led news consumption. Time spent on each website (time stamp difference between successive clicks) was also tracked and operationalized along the same principles as time spent with overall news consumption, time spent with social-media-led news items, and news-website-led news items. While our operationalization of personalized newsfeed content measured as *news exposure* includes data from Facebook only because of the technical complexity of data collection for dynamic social media site content, our categorization of clicking data as *social media news consumption* combines data also from the other social media platforms mentioned above.
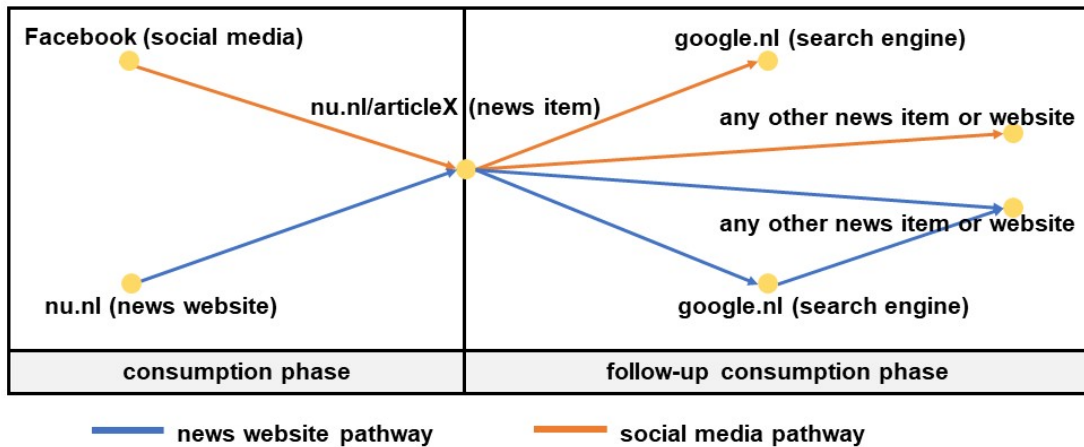

Figure 1. Exemplary operationalization of clicking sequences as news consumption and follow-up news consumption.

Follow-up News Consumption. The amount of follow-up consumption is calculated as the percentage of occasions in which the consumption of a news item led to additional news-related behavior such as visiting another news article or video, the website of a news outlet, or carrying out a news-related search. A news-related search was operationalized as a visit to a search engine followed by accessing any news outlet site after consuming a news item. Again, we differentiate between follow-up consumption after a news item that was social-media-led or a news item that was news-website-led. Follow-up consumption is contingent on initial news consumption. We can therefore measure the variability of follow-up consumption only in terms of political interest, income, and education within the group of users who actually consume news. If H2 is supported, this group might already have higher levels of political interest and income than those who do not consume news often.

---

[4] Clicking sequences where the visit of a social media site was followed by visiting the home page of a news organization were not labeled as social-media-led news consumption since no clear connection can be assumed. Practices of "news snacking" (e.g., scrolling down a home page without clicking) are not taken into account. However, of all 330 users who engaged with news content, just nine visited only the main page of a news outlet but not any news items.

Results

Facebook News Exposure

To test H1, we analyzed the data gathered from our participants' Facebook newsfeeds. For 115 users, we were able to match newsfeed and survey data and identify (news-related) newsfeed content via our allowlist.[5] On average, 4% of all items in their feeds contained the name of or a link to a news outlet. This finding is consistent with numbers published by the company itself. CEO Zuckerberg stated that news content makes up about 5% of users' newsfeeds (Zuckerberg, 2018). Figure 2 in the supplementary online material (https://osf.io/e8rf7/) depicts the rather skewed distribution of news exposure and consumption in our data across participants.

Model A in Table 1 depicts the results of the linear regression model on the percentage of news items in the Facebook newsfeed ($n = 110$). For users with higher levels of political interest, we observe significantly higher levels of news exposure than for users with lower levels of political interest. For users who reported the highest levels of political interest, we identified on average 10% of their newsfeed content as news-related ($M = 10.12$, $SD = 9.89$, $Median = 6.8$), whereas we measured less than 3% news content on average in the newsfeeds of those who reported the lowest interest in politics ($M = 2.84$, $SD = 2.19$, $Median = 1.8$). While H1a is supported, the hypothesized relation between higher levels of education (H1b) and income and higher levels of news exposure was not significant (H1c).

---

[5] Many participants in our original sample ($N = 413$) had accounts but barely used the platform on the device that we tracked in the yearlong field period even though a Facebook task was given. Therefore, the numbers are limited because of non-Facebook use (−178 participants), the absence of items from the allowlist (−67) or certain survey variables (−53).

Table 1. Results of OLS Models for (a) Percentage of News Items in the Facebook Feed, (b) Percentage of Overall News Consumption (News Websites and News Items), (c) Percentage of News Items Accessed via Social Media, and (d) Percentage of News Items Accessed via the Main Page of a News Website.

	Model (A) Exposure on Facebook	Model (B) Overall Consumption	Model (C) Consumption via Social Media	Model (D) Consumption via News Website
Intercept	0.75 (0.60)	1.21* (0.55)	−0.43 (0.28)	0.94* (0.46)
Education	0.09 (0.05)	0.07 (0.05)	−0.01 (0.02)	0.05 (0.04)
Income	−0.05 (0.05)	0.04 (0.04)	0.01 (0.02)	0.04 (0.03)
Political Interest	0.18** (0.06)	0.12* (0.05)	0.07** (0.03)	0.06 (0.04)
Political Extremism	−0.12* (0.05)	−0.04 (0.05)	−0.05* (0.02)	−0.05 (0.04)
Internal Efficacy	−0.13 (0.10)	−0.02 (0.09)	0.04 (0.05)	0.01 (0.08)
Technical Literacy	0.13 (0.09)	−0.01 (0.08)	0.08 (0.04)	−0.15* (0.07)
Mobile News Use	0.03 (0.03)	0.03 (0.02)	0.00 (0.01)	0.02 (0.02)
Age	−0.01 (0.00)	0.00 (0.00)	0.00 (0.00)	−0.00 (0.00)
Gender Female	−0.06 (0.16)	−0.48*** (0.14)	0.02 (0.07)	−0.26* (0.12)
Observations	110	317	177	229
R^2/R^2 adjusted	0.19/0.12	0.15/0.12	0.12/0.08	0.13/0.09

*Note. *$p < .05$, **$p < .01$, ***$p < .001$, VIF for all factors < 2.0. Dependent variables calculated in relation to overall online activity measured and log-transformed. Variances in observations are results of the differences in the number of participants ($N = 413$) engaging in these practices.*

News Consumption

We analyzed the click sequences for all 413 users who completed tracking and survey to identify the amount of news consumption by comparing accessed URLs to our allowlist of news organizations. On average, 10% of all clicks were related to news websites and 79% of all users visited at least one news website (main page or specific news item) during the tracking period.

Table 1 also depicts the results of the three linear regression models for news consumption. In the overall news consumption model (model B), a higher level of political interest significantly predicts a higher number of all news-related websites clicked, a relation also discussed in previous research. In model C, predicting the amount of news items accessed via social media, the data also show that levels of political interest have a significant effect. Users with higher levels of political interest consume higher amounts of news items via social media than users with lower levels of political interest. H2a is supported by these results. A significant influence of income and education cannot be identified in any of the models (H2b–c, H3b–c).

While political interest is a significant factor in model C (news items consumed via social media), it is not a significant predictor for model D (news items consumed via a news website). This could point toward a higher importance of the motivational variable political interest for a social media news environment than for news pathways via traditional online news websites (H3).

The reason for this difference might be that users who visit a news website are already interested in politics and, therefore, no effect of political interest at the stage of the actual article selection is measurable. Interest in politics is not a common motivator for visiting a social media website. Therefore, the decision to consume news while on a social media website could be taken at the stage of concrete exposure to a news item. To test this post hoc interpretation, we additionally examined the influence of political interest on the probability of news home page and social media use (Table 2).

Table 2. Results of the Logistic Regression Models for News Home Page Use and Social Media Use.

	News Home Page Use	Social Media Use
(Intercept)	−1.03 (0.93)	0.84 (1.13)
Education	0.04 (0.08)	−0.03 (0.10)
Income	−0.06 (0.06)	−0.04 (0.06)
Political Interest	0.11 (0.09)	−0.05 (0.11)
Political Extremism	0.08 (0.08)	0.04 (0.10)
Internal Efficacy	0.13 (0.16)	0.30 (0.19)
Technical Literacy	0.09 (0.14)	0.12 (0.17)
Mobile News Use	0.13 (0.04)**	0.02 (0.05)
Age	0.00 (0.01)	−0.01 (0.01)
Gender	−0.43 (0.24)	0.09 (0.30)
Observations	397	397

Note. $*p < .05$, $**p < .01$, $***p < .001$, VIF for all factors < 2.0.

In the logistic regression models in Table 2, we do see a higher level of political interest having a positive effect on the use of news home pages and can even see a negative effect on the use of social media platforms among our sample. Even though this effect is not significant, it could indicate that political interest becomes a relevant factor at different steps of the information-seeking process in the respective pathways to news.

We did not measure only clicks but also the time spent on each URL. A high amount of news-related URLs accessed correlates strongly with a high amount of time spent on news content ($r = .92$, $p < .001$). We also compared how much time users spend with articles accessed via social media. Users who accessed the news item via social media spent on average 10 seconds longer with the news item.

Follow-up News Consumption

In the linear regression models predicting the amount of follow-up news consumption (E, F, G) we did not identify any significant effects of political interest and education and observe low levels of explained variance with our factors (Table 3). H4 and H5 are, therefore, rejected. Only the level of income served as a significant factor for the amount of news items accessed via main pages; this could possibly be related to the paywall systems on different Dutch news websites.

Users were much more likely to employ follow-up behavior after the consumption of a news item via the news website than via social media. On average, only 22% of all news items referred via social media led to follow-up consumption, while in 41% of all cases when news items were accessed through a news-related website, follow-up consumption occurred.

Table 3. Results of OLS Models for (e) Percentage of Follow-up Consumption After any News Consumption, (f) Percentage of Follow-up Consumption After Consumption via Social Media, and (g) Percentage of Follow-up Consumption After Consumption via a News Website.

	Model (E) Overall Follow- up Consumption	Model (F) Follow-up Consumption via Social Media	Model (G) Follow-up Consumption via News Website
Intercept	4.02*** (0.20)	4.72*** (0.35)	4.43*** (0.17)
Education	0.03 (0.02)	0.06* (0.03)	−0.02 (0.01)
Income	0.01 (0.01)	0.02 (0.03)	0.03* (0.01)
Political Interest	0.01 (0.02)	−0.04 (0.03)	−0.01 (0.02)
Political Extremism	−0.00 (0.02)	0.04 (0.03)	0.01 (0.01)
Internal Efficacy	0.02 (0.03)	−0.00 (0.05)	0.03 (0.03)
Technical Literacy	−0.02 (0.03)	−0.18** (0.05)	−0.01 (0.03)
Mobile News Use	0.00 (0.01)	0.01 (0.01)	−0.01 (0.01)
Age	−0.00 (0.00)	−0.00 (0.00)	−0.00 (0.00)
Gender Female	−0.03 (0.05)	−0.22** (0.08)	−0.03 (0.04)
Observations	288	138	203
R^2/R^2 adjusted	0.04/0.01	0.17/0.11	0.05/0.00

Note. *$p < .05$, **$p < .01$, ***$p < .001$, VIF for all factors < 2.0. Dependent variables calculated in relation to overall online activity measured and log-transformed. Variances in observations are results of the differences in the number of participants ($N = 413$) engaging in these practices.

Conclusion and Discussion

The combination of survey and tracking data allowed us to reconstruct the user's journey along news exposure, consumption, and follow-up engagement on social media and how these practices are influenced by education, income, and political interest.

Users with higher levels of political interest were not only more likely to actually see news content on Facebook but they were also more often likely to actually click on news items on Facebook or other social media platforms. Once a user had accessed a news item via social media, they spent more time with the specific content but engaged less often in follow-up behavior such as searching for/reading more news items than when discovering news via a news website.

While social media filtering algorithms remain opaque, one can assume that their impact on news selection depends on previously demonstrated interest in news and politics. Our findings align with recent results from Thorson and colleagues (2019) that Facebook users who are algorithmically categorized as interested in news or politics are more likely to report such content in their newsfeeds. These results also

indicate that the increasing importance of political interest in news consumption (Strömbäck et al., 2013) shown for other offline and online media is also applicable for social media.

News exposure on social media has often been characterized as incidental, as users may "encounter current affairs information when they had not been actively seeking it" (Tewksbury & Maddex, 2001, p. 534). However, in our analysis of the data regarding the most used social media platform in the Netherlands, we find that (a) more than half of all Facebook users in our sample did not see news on Facebook on their desktop devices, (b) for those exposed to news, the amount of news items was less than 5% compared with other content, and (c) the amount of news exposure was dependent on preexisting political interest. Our results align with the recent work, arguing for a reconsideration of the common conceptualization of social media exposure as incidental or accidental (Thorson, 2020) but rather propose the existence of a "Matthew Effect" (Kümpel, 2020, p. 1083) in social media news use, suggesting (relative) enrichment among users already interested in news and (relative) impoverishment among those with little or no interest in current affairs. In our analysis of consumption after news exposure, the effect of political interest was more prevalent for news consumption via social media compared with news consumption via a news website. This dynamic also substantiates the potential for the acceleration of news gaps related to political interest rather than a leveling effect of social media.

However, the result that the motivational variable of political interest is more important for clicking onto a news item in a social media news environment than for news selection on a news website could, in co-occurrence with the positive of political interest on the visit of a new website but not social media platforms, also indicate that political interest might be a relevant factor at different steps on the respective pathways to news. Users interested in politics are already more likely to use news websites than users who do not have this interest. These results should encourage a differentiation in future research on personalized news environments beyond news exposure, news consumption, and news engagement to additionally explore probabilities of individuals engaging in information-seeking behavior by accessing news-poor or news-rich environments online.

Inequality in news exposure and consumption on social media is especially problematic because of the potential consequences for overall information seeking. Individuals who trust that "news will find them" on social media (Gil de Zúñiga et al., 2017, p. 105) might lack motivation to engage with other sources of news. Schäfer (2020) found that users who were exposed to news only in their feeds felt as knowledgeable as individuals who had actually consumed items. An illusion of knowledge and news use can hinder actual engagement with news on social media.

The fact that in our study follow-up news consumption was more prevalent when the original news item was accessed via a news website than via social media can be interpreted in different ways: On the one hand, consumption via a news website is generally more closely related to intentional information seeking and thereby more likely to lead to attention and cognitive elaboration, which would warrant further news consumption (Eveland, 2001). On the other hand, because of the personalized news selection on social media users could already satisfy their informational needs adequately with relevant content, so further news consumption is unnecessary. The last interpretation connects to the fact that users spend more time with personalized news items accessed via social media than with items via news websites. Even though

this study already combines tracking data with survey data, more explicit situational self-reporting via experience sampling or reconstruction interviews (Merten, 2020b) could help in interpreting these digital traces further.

The positive effects of education and income on news exposure and consumption, which were demonstrated in previous research, were not confirmed. This could mean that, even though news gaps related to these sociodemographic factors might not be leveled, they are at least not widened through social media news exposure and consumption. But these effects just might not be identifiable with our limited sample and statistical power. Larger representative studies could provide more definitive evidence in that regard.

Our findings should be interpreted with caution because of several limitations that came as a consequence of the way our data were collected and analyzed. We designed a long tracking period to combat potential behavior adjustments because of the measurement; nevertheless, there might have been effects of social desirability. The small sample of participants limits the statistical power of our models and the representativeness of the results. The latter problem, however, is common for the tracking research in general (Stier, Breuer, Siegers, & Thorson, 2020; Thorson, 2020), because it relies on the participants' decision to opt in to the tracking. It results in the tracking samples often diverging from the general/online population for such parameters as age, education, or gender (Stier et al., 2020).

In the Reuters Institute Digital News Report data on Facebook use in the Netherlands in the field period, 65% of participants reported using Facebook (Costera Meijer & Groot Kormelink, 2018). We measured Facebook use for 56% of our tracking participants (235 of 413). A reason for this discrepancy beyond different sample composition in the two online access panels could be an overreporting of Facebook use. In a recent meta-analysis of 66 studies comparing survey and tracking data on digital media use, Parry and colleagues (2021) demonstrate that self-reports do not exhibit convergent validity with tracked measures of usage.

Another reason for the divergence of Reuters Institute Digital News Report results and our tracking data could be the lack of mobile measurement data and reliance on data from one browser in our data collection. Individuals might have been exposed to news on Facebook in digital environments where our tracking software was not installed. Because of the closed app and system software on mobile devices, only the tracking of desktop devices was possible in this study and obtaining comprehensive tracking data from mobile devices is a largely unsolved challenge.

Instant messaging services such as WhatsApp were also frequently used social media applications for news in 2018 in the Netherlands (Costera Meijer & Groot Kormelink, 2018). Because of the even more private nature of posts and messages on these chat apps and accompanying privacy concerns plus the fact that these services are mainly used via mobile devices, including the use of the instant messaging services was beyond the scope of our methodological and legal setup. Hopefully, future studies can add analysis on instant-messaging platforms and also include more social networking platforms beyond Facebook when collecting dynamic social media news exposure content.

Nevertheless, this study is one of the first attempts to thoroughly examine individual news use via social media combining tracking and survey data. As the accuracy of self-reported measures of news use is diminished, digital traces provide opportunities for the measurement of online news use in a nonintrusive way and with high precision and granularity (Stier et al., 2020). Alongside the traces, relevant information about individuals' offline activities, sociodemographic characteristics, and motivational factors can be gathered as demonstrated in this study. With our operationalization of news use on social media, we have established a methodological framework for future tracking studies that reflect the multidimensionality of online news exposure, consumption, and engagement.

References

Aiello, L. M., Barrat, A., Schifanella, R., Cattuto, C., Markines, B., & Menczer, F. (2012). Friendship prediction and homophily in social media. *ACM Transactions on the Web, 6*(2), Article 9, 1–33. doi:10.1145/2180861.2180866

Anspach, N. M. (2017). The new personal influence: How our Facebook friends influence the news we read. *Political Communication, 34*(4), 590–606. doi:10.1080/10584609.2017.1316329

Bode, L. (2016). Political news in the news feed: Learning politics from social media. *Mass Communication and Society, 19*(1), 24–48. doi:10.1080/15205436.2015.1045149

Bodo, B., Helberger, N., Irion, K., Zuiderveen Borgesius, F., Moeller, 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 & Technology*, *19*, 133–180. Retrieved from http://www.yjolt.org/tackling-algorithmic-control-crisis-technical-legal-and-ethical-challenges-research-algorithmic

Boulianne, S. (2018). Twenty years of digital media effects on civic and political participation. *Communication Research, 47*(7), 947–966. doi:10.1177/0093650218808186

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. Cambridge, MA: Harvard University Press.

Central Bureau voor de Statistiek. (2017a). *Statline (responsive open data portal): Population distribution: Age and gender*. Retrieved from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37296ned/table?dl=3DF05

Central Bureau voor de Statistiek. (2017b). *Statline (responsive open data portal): Population distribution: Level of education*. Retrieved from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82275NED/table?dl=3DF07

Costera Meijer, I., & Groot Kormelink, T. (2018). *Reuters Institute digital news report. Country page: Netherlands*. Retrieved from http://www.digitalnewsreport.org/survey/2018/netherlands-2018/

Dahlgren, P. M. (2019). Selective exposure to public service news over thirty years: The role of ideological leaning, party support, and political interest. *International Journal of Press/Politics, 24*(3) 293–314. doi:10.1177/1940161219836223

Delli Carpini, M. X., & Keeter, S. (1996). *What Americans know about politics and why it matters*. New Haven, CT: Yale University Press.

DeVito, M. A. (2017). From editors to algorithms. *Digital Journalism, 5*(6), 753–773. doi:10.1080/21670811.2016.1178592

Eveland, W. P. (2001). The cognitive mediation model of learning from the news: Evidence from nonyear election, off-year election, and presidential election contexts. *Communication Research, 28*(5), 571–601. doi:10.1177/009365001028005001

Feezell, J. T., & Ortiz, B. (2021). 'I saw it on Facebook': An experimental analysis of political learning through social media. *Information, Communication & Society*, *24*(9), 1283–1302. https://doi.org/10.1080/1369118X.2019.1697340

Fletcher, R., & Nielsen, R. K. (2018). Are people incidentally exposed to news on social media? A comparative analysis. *New Media & Society, 20*(7), 2450–2468. doi:10.1177/1461444817724170

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*(3), 105–123. doi:10.1111/jcc4.12185

Guess, A. M. (2015). Measure for measure: An experimental test of online political media exposure. *Political Analysis, 23*(1), 59–75. doi:10.1093/pan/mpu010

Habermas, J. (1989). *The structural transformation of the public sphere*. Cambridge, MA: MIT Press.

Haenschen, K. (2019). Self-reported versus digitally recorded: Measuring political activity on Facebook. *Social Science Computer Review, 38*(5), 567–583, doi:10.1177/0894439318813586

Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday, 7*(4). doi:10.5210/fm.v7i4.942

Kabiljo, M., & Ilic, A. (2015, June 2). *Recommending items to more than a billion people.* Facebook engineering. Retrieved from https://engineering.fb.com/2015/06/02/core-data/recommending-items-to-more-than-a-billion-people/

Kaiser, J., Keller, T. R., & Kleinen-von Königslöw, K. (2018). Incidental news exposure on Facebook as a social experience: The influence of recommender and media cues on news selection. *Communication Research, 48*(1), 77–99. doi:10.1177/0093650218803529

Kalogeropoulos, A., & Nielsen, R. K. (2018). *Factsheet: Social inequalities in news consumption.* Oxford, UK: Reuters Institute for the Study of Journalism. Retrieved from https://reutersinstitute.politics.ox.ac.uk/our-research/social-inequalities-news-consumption

Karnowski, V., Kümpel, A. S., Leonhard, L., & Leiner, D. J. (2017). From incidental news exposure to news engagement: How perceptions of the news post and news usage patterns influence engagement with news articles encountered on Facebook. *Computers in Human Behavior, 76*, 42–50. doi:10.1016/j.chb.2017.06.041

Knoke, D. (1990). Networks of political action: Toward theory construction. *Social Forces, 68*(4), 1041–1063. doi:10.2307/2579133

Kümpel, A. S. (2019). The issue takes it all?: Incidental news exposure and news engagement on Facebook. *Journal of Communication, 69*(4), 373–395. doi:10.1080/21670811.2018.1465831

Kümpel, A. S. (2020). The Matthew effect in social media news use: Assessing inequalities in news exposure and news engagement on social network sites (SNS). *Journalism, 21*(8), 1083–1098. doi:10.1177/1464884920915374

Lecheler, S., & de Vreese, C. H. (2017). News media, knowledge, and political interest: Evidence of a dual role from a field experiment. *Journal of Communication, 67*(4), 545–564. doi:10.1111/jcom.12314

Lindell, J. (2018). Distinction recapped: Digital news repertoires in the class structure. *New Media & Society, 20*(8), 3029–3049. doi:10.1177/1461444817739622

Mahrt, M. (2019). *Beyond filter bubbles and echo chambers: The integrative potential of the Internet.* Berlin, Germany: Digital Communication Research.

McPherson, M., Smith-Lovin, L., & Cook, J. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27*(1), 415–444. doi:10.1146/annurev.soc.27.1.415

Merten, L. (2020a). Block, hide, or follow: Personal news curation practices on social media. *Digital Journalism, 9*(8), 1018–1039. doi:10.1080/21670811.2020.1829978

Merten, L. (2020b). Contextualized repertoire maps: Exploring the role of social media in news-related media repertoires. *Forum Qualitative Sozialforschung*, *21*(2). doi:10.17169/fqs-21.2.3235

Moeller, J., van de Velde, R. N., Merten, L., & Puschmann, C. (2019). Explaining online news engagement based on browsing behavior: Creatures of habit? *Social Science Computer Review*, *38*(5), 616–632. doi:10.1177/0894439319828012

Mosseri, A. (2016, June 29). *Building a better news feed for you*. Facebook Newsroom. Retrieved from https://newsroom.fb.com/news/2016/06/building-a-better-news-feed-for-you/

Newman, N., Fletcher, R., Schulz, A., Andi, S., Robertson, C. T., & Nielsen, R. K. (2021). *Reuters Institute digital news report 2021*. Reuters Institute for the Study of Journalism. Retrieved from https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021

Oeldorf-Hirsch, A. (2018). The role of engagement in learning from active and incidental news exposure on social media. *Mass Communication and Society, 21*(2), 225–247. doi:10.1080/15205436.2017.1384022

Ohlsson, J., Lindell, J., & Arkhede, S. (2016). A matter of cultural distinction: News consumption in the online media landscape. *European Journal of Communication, 32*(2), 116–130. doi:10.1177/0267323116680131

Parry, D. A., Davidson, B. I., Sewall, C. J. R., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour, 5*, 1535–1547. doi:10.1038/s41562-021-01117-5

Prior, M. (2007). *Postbroadcast democracy: How media choice increases inequality in political involvement and polarizes elections*. New York, NY: Cambridge University Press.

Revilla, M., Ochoa, C., & Loewe, G. (2017). Using passive data from a meter to complement survey data in order to study online behavior. *Social Science Computer Review, 35*(4), 521–536. doi:10.1177/0894439316638457

Schäfer, S. (2020). Illusion of knowledge through Facebook news? Effects of snack news in a news feed on perceived knowledge, attitude strength, and willingness for discussions. *Computers in Human Behavior, 103*, 1–12. doi:10.1016/j.chb.2019.08.031

Schmidt, J.-H., Merten, L., Hasebrink, U., Petrich, I., & Rolfs, A. (2019). How do intermediaries shape news-related media repertoires and practices? Findings from a qualitative study. *International Journal of Communication, 13*, 853–873.

Stier, S., Breuer, J., Siegers, P., & Thorson, K. (2020). Integrating survey data and digital trace data: Key issues in developing an emerging field. S*ocial Science Computer Review, 38*(5), 503–516. doi:10.1177/0894439319843669

Strömback, J., Djerf-Pierre, M., & Shehata, A. (2013). The dynamics of political interest and news media consumption: A longitudinal perspective. *International Journal of Public Opinion Research, 25*(4), 414–435. doi:10.1093/ijpor/eds018

Tewksbury, D., & Maddex, B. (2001). Accidentally informed: Incidental news exposure on the World Wide Web. *Journalism & Mass Communication Quarterly, 78*(3), 533–554. https://doi.org/10.1177/107769900107800309

Thorson, K. (2020). Attracting the news: Algorithms, platforms, and reframing incidental exposure. *Journalism, 21*(8), 1067–1082. doi:10.1177/1464884920915352

Thorson, K., Cotter, K., Medeiros, M., & Pak, C. (2019). Algorithmic inference, political interest, and exposure to news and politics on Facebook. *Information, Communication & Society, 24*(2), 183–200. doi:10.1080/1369118X.2019.1642934

Thorson, K., & Wells, C. (2016). Curated flows: A framework for mapping media exposure in the digital age. *Communication Theory, 26*(3), 309–328. doi:10.1111/comt.12087

Tichenor, P., Donohue, G. A, & Olien, C. N. (1970). Mass media flow and differential growth in knowledge. *Public Opinion Quartely, 34*(2), 159–170. doi:10.1086/267786

Trilling, D., & Schönbach, K. (2015). Challenging selective exposure: Do online news users choose sites that match their interests and preferences? *Digital Journalism, 3*(2), 140–157. doi:10.1080/21670811.2014.899749

Vermeer, S., Trilling, D., Kruikemeier, S., & de Vreese, C. (2020). Online news user journeys: The role of social media, news websites, and topics. *Digital Journalism, 8*(9), 1114–1141. https://doi.org/10.1080/21670811.2020.1767509

Wei, L., & Hindman, D. B. (2011). Does the digital divide matter more? Comparing the effects of new media and old media use on the education-based knowledge gap. *Mass Communication and Society, 14*(2), 216–235. https://doi.org/10.1080/15205431003642707

Woo-Yoo, S., & Gil-de-Zúñiga, H. (2014). Connecting blog, Twitter, and Facebook use with gaps in knowledge and participation. *Communication & Society, 27*(4), 33–48. doi:10.15581/003.27.4.33-48

Zuckerberg, M. (2018, January 19). *Continuing our focus for 2018 to make sure the time we all spend on Facebook is time well spent* [Facebook status update]. Retrieved from https://www.facebook.com/zuck/posts/10104445245963251