Routledge
Taylor & Francis Group

Non-News Websites Expose People to More Political Content Than News Websites: Evidence from Browsing Data in Three Countries

Magdalena Wojcieszak ⓘ[a,b], Ericka Menchen-Trevino ⓘ[c], Bernhard Clemm von Hohenberg[d], Sjifra de Leeuw[e], João Gonçalves[f], Sam Davidson[g], and Alexandre Gonçalves[h]

[a]Department of Communication, University of California, Davis, California, USA; [b]Amsterdam School of Communication Research, Amsterdam, The Netherlands; [c]School of Communication, American University, Washington DC, Massachusetts, USA; [d]GESIS – Leibniz-Institute for the Social Sciences, Mannheim, Germany; [e]Dutch Ministry of Finance, Den Haag, The Netherlands; [f]Department of Media and Communication, Erasmus University Rotterdam, Rotterdam, The Netherlands; [g]Linguistics Department, University of California, Davis, California, USA; [h]Department of Journalism, University of Illinois, Urbana-Champaign, IL, USA

ABSTRACT

Most scholars focus on the prevalence and democratic effects of (partisan) news exposure. This focus misses large parts of online activities of a majority of politically disinterested citizens. Although political content also appears outside of news outlets and may profoundly shape public opinion, its prevalence and effects are understudied at scale. This project combines three-wave panel survey data from three countries (total N = 7,266) with online behavioral data from the same participants (over 106M visits). We create a multi-lingual classifier to identify political content both in news and outside (e.g. in shopping or entertainment sites). We find that news consumption is infrequent: just 3.4% of participants' online browsing comprised visits to news sites. Only between 14% (NL) and 36% (US) of these visits were to news about politics. The overwhelming majority of participants' visits were to non-news sites. Although only 1.6\% of those visits related to politics, in absolute terms, citizens encounter politics more frequently outside of news than within news. Out of every 10 visits to political content, 3.4 come from news and 6.6 from non-news sites. Furthermore, exposure to political content outside news domains had the same – and in some cases stronger - associations with key democratic attitudes and behaviors as news exposure. These findings offer a comprehensive analysis of the online political (not solely news) ecosystem and demonstrate the importance of assessing the prevalence and effects of political content in non-news sources.

Theorists emphasize the contributions of news media to effective democracy (Delli Carpini & Keeter, 1996) and researchers examine whether people consume news (Allen et al., 2020; Yang et al., 2020) and the effects of this consumption (Feezell, 2018; Strömbäck, 2005). And

yet, a sizable proportion of the American population sees news as complex or boring, is averse to partisan politics (Klar et al., 2018), and avoids news altogether (Feldman et al., 2013; Prior, 2007). This is especially the case in the online environment, which offers nearly unlimited content and unprecedented possibilities to customize individual media diets. In fact, news makes up only 1.4% of News Feed on Facebook (Meta, 2022), down from 4% in 2018 (Zuckerberg, 2018), only between 2% (Wojcieszak, de Leeuw, et al., 2021) and 7%–9% (Guess, 2021) of all URLs visited by large samples of Americans are news domains (Stier et al., 2022), and – across mobile and desktop – news comprises only 4.2% of total online consumption (Allen et al., 2020).

Clearly, the persistent focus on news and its effects is missing huge parts of online activities of a majority of citizens. This focus, moreover, may inaccurately portray the consumption and the effects of *political* content in *non-news* websites. In the "old" media environment, citizens saw politics in soft news shows (Andersen, 2019; Baum, 2003; Prior, 2003). The literature on incidental exposure similarly suggests that internet users *report* encountering news by chance online and on social media (Tewksbury et al., 2001; Valeriani & Vaccari, 2016), and these reported exposures may foster political knowledge and engagement (Baum, 2003; Feezell, 2018; Heiss & Matthes, 2019; Kwak et al., 2020).

Because this work relies on largely unreliable self-reports of (incidental) (news) exposure (Prior, 2009; Thorson, 2020), it cannot accurately portray online political ecosystem. In turn, studies that use online traces primarily focus on *partisan* news use and its effects in the U.S (Guess et al., 2021; Wojcieszak, de Leeuw, et al., 2021), not on exposure to political content outside of (partisan) news websites. Needless to say, partisan news attracts a small fraction of the population (Prior, 2013; Wojcieszak, de Leeuw, et al., 2021) and the U.S. is far from representative of other countries and media systems globally. Missing is evidence on whether users internationally actually encounter politics in ostensibly nonpolitical spaces (e.g., an article about COVID in *Women's Health* magazine) and how these encounters compare to news consumption in terms of both prevalence and democratic effects.

These open questions are important to our understanding of where citizens learn about politics and with what effects. To the extent that news consumption is beneficial, the low absolute levels of news use in the population may be problematic (Putnam, 2000). Inattentive and politically disengaged citizens may be susceptible to misinformation, fall for populist actors, and ultimately swing elections (Achen & Bartels, 2006; Bartels, 1996; Brennan, 2017). The 2016 U.S. presidential election or Brexit are prime examples (Fording & Schram, 2017; Somin, 2016). Yet, if people encounter political content outside news sites and *if* these encounters enhance important democratic outcomes, the worries about declining news audience – although nontrivial – may be overstated.

We rely on a large project that combines panel surveys and online behavioral data from three countries (total $N = 7,266$).[1] Participants submitted their online browsing data (desktop only, no mobile) via our open-source tool, Web Historian (Menchen-Trevino, 2016). From over 106 million visits to over 65 million URLs in over 655,000 domains, we match the visited domains to identifiable news domains as determined by our open-source lists of news sites per country, for a total of 2,366 news domains visited across the countries. We also identified YouTube channels, Facebook pages, and Twitter handles of all the news media organizations on our lists and assigned these exposures accordingly. Furthermore, we develop a multilingual BERT-based neural binary classifier to categorize the content visited

in news *and also* non-news domains as related to politics or not (e.g., the aforementioned article about COVID in *Women's Health* magazine would be categorized as political content outside news).

We use these behavioral data to assess the prevalence of exposure to political content *outside* news domains in people's online browsing. We compare it with the prevalence of exposure to news domains and also – to offer additional nuance and put news exposure in perspective – to political content within news. Furthermore, we take advantage of panel surveys and leverage over-time within-respondent variation to examine the effects of these behaviorally tracked exposures on both beneficial (e.g., participation, political trust, support for compromise) as well as negative (e.g., misperceptions, attitude and affective polarization) outcomes. This undertaking represents the most comprehensive comparative evidence on online political ecosystem to date.

We offer a key finding; most exposures to politics online take place *outside* news sites, far more than within news. Although only 1.6%[2] of visits to non-news sites were to political content, in absolute terms, citizens in the three countries read about politics *more* frequently outside of news sites. For 10 pieces of political content read online, 6.6 come from non-news websites and only 3.4 from news. In fact, only 3.4% of visits were to news sites and only between 14% (NL) and 36% (US) of these visits were to political content within news sites. We also find that the effects of exposure to politics outside news are astonishingly similar to those of news exposure. Before presenting our data, we outline the theorizing and evidence on low news consumption levels and on incidental exposure to news and politics.

News, News Avoidance, Incidental Exposure

Since Alexis de Tocqueville's seminal book on democracy in America (Tocqueville, 1863), journalism has been seen as central to an informed public and a well-functioning democratic society (Delli Carpini & Keeter, 1996). By covering the issues of the day, news media increase political knowledge and inform prospective voters (Chaffee et al., 1994) and by providing information about opportunities for political involvement, news media stimulate participation (Lemert, 1984; Norris, 2000; Strömbäck, 2005). This positive link between news media use and various beneficial outcomes has been found for both offline and digital media and across countries (Beckers et al., 2021; Kenski & Stroud, 2006; Ohme, 2020).

Yet, despite the theoretical centrality of news, documented benefits of news use, and the unprecedented availability of news in the current media environment, large groups of citizens avoid news, either actively or unintentionally (Newman et al. 2019; Skovsgaard & Andersen, 2020) and for a variety of reasons. Most generally, people use media that satisfy their needs and desires (Katz et al., 1973) and – for many – current affairs are not as gratifying or appealing as entertainment and other nonpolitical fare. In addition, negativity and conflict in news coverage generate stress or anxiety (Villi et al., 2022), the sheer amount of news may lead to news overload and fatigue (Song et al., 2017), and also some citizens may be uninterested in politics, perceive it as irrelevant, or lack the skills to navigate the news system (Park, 2019; Schmitt et al., 2018). News avoidance can entail active or intentional resistance or rejection of news or be an unintentional byproduct of increased choice in the current media environment (Skovsgaard & Andersen, 2020), both of which can have negative consequences for the democratic system.

The online environment can further exacerbate these tendencies as it makes it easier than ever to tune out of news and public affairs. Citizens have unprecedented amount of content at their disposal and can easily immerse themselves in movies, sports, or celebrity life. In fact, although engagement with news and entertainment is not mutually exclusive (Huang & Yang, 2022), news consumption does decrease the more choice is available to people, a pattern detected in the U.S (Feldman et al., 2013; Prior, 2007) and in Europe (Aalberg et al., 2013). Ultimately, large groups of society may further withdraw from the political process, become less knowledgeable, more susceptible to populism or misinformation, and less able to vote in accordance with their personal or group interests (Prior, 2007). To the extent that effectively functioning democracies require citizens who are up-to-date on news and politically active (Delli Carpini & Keeter, 1996), these low overall patterns are problematic.

At the same time, citizens can encounter political information without seeking it and without much effort. For instance, movie goers saw newsreels before the movie (Downs, 1957) and television viewers learned about politics and public affairs when watching *Good Morning America*, *The Oprah Winfrey Show*, and other soft news programs that mixed entertainment and politics (Andersen, 2019; Baum, 2003; Baum & Jamison, 2006; Prior, 2003). Because these programs discuss cooking or celebrities, they attract viewers whose primary motivation is not current affairs, but who nevertheless learn about these issues.

The potential for such inadvertent exposure is yet greater online. The burgeoning literature on incidental exposure suggests that web portals, social media, or search engines are key avenues to news online, directing traffic to news websites (Stier et al., 2022; Wojcieszak, Menchen-Trevino, et al., 2021). In fact, many users report encountering politics by chance, such as when going online for trivia or romance (Wojcieszak & Mutz, 2009) or connecting with friends and family on social media (Valeriani & Vaccari, 2016). These incidental exposures to politics, whether in televised infotainment (Andersen, 2019; Baek & Wojcieszak, 2009; Baum, 2003; Baum & Jamison, 2006; Moy et al., 2005) or online (Feezell, 2018; Heiss & Matthes, 2019; Kwak et al., 2020; Weeks et al., 2022), are cross-sectionally and – to a lesser extent – causally related to news use, political knowledge, participation, and political discussion (Nanz & Matthes, 2022).

Although these findings are promising, they primarily rely on self-reports of exposure to (soft) news programs, which are unreliable and subject to various biases, such as recall or social desirability (Prior, 2009). Given that it is yet more challenging to recall *incidental* exposure as one that – by definition – is only tangential to users' experience, these reports may not be accurate. In fact, the high percentages of users *reporting* inadvertent exposure to news and politics are surprising, given that what users see online is partly shaped by recommendation algorithms and social networks. As such, those who do not like or follow public affairs and whose network is politically disinterested are unlikely to encounter politics online (Wells & Thorson, 2017). For these reasons, research finds that public affairs comprise 1.8% of the average Facebook feed of students, with the median student liking zero pages from journalists or news organizations (Wells & Thorson, 2017). In a related vein, those who are better informed are more likely to encounter and notice more exposures as political and thus it is still the engaged individuals who are *found by the news* (Thorson, 2020). As importantly, by focusing exclusively on news versus incidental exposure to politics, extant scholarship misses other kinds of exposures online.

To date, the focus on news use and the methodological limitations of prior work have generated important gaps in our knowledge about the online political ecosystem, especially internationally. Missing is evidence on whether internet users actually encounter political information outside news websites, not merely report such encounters. It is also important to assess how the extent of these encounters compares to the extent of news consumption in general and that of the consumption of political content within news sites. News sites do not only feature hard news, as the majority of content therein concerns nonpolitical topics. As such, visiting a news site is not synonymous with exposure to politics, as many people go to *CNN* or *Fox News* for sports, recipes, or entertainment (Flaxman et al., 2016; Guess, 2021; Merten, 2021).

Our project fills these gaps. We rely on a combination of over-time surveys and online behavioral data from the same participants in three countries to address four progressively specific questions: (1) What is the prevalence of exposure to political content *outside* news domains in individual browsing? (2) What is the prevalence of exposure to news domains (total) and also to political content in news domains. Moreover, once we acknowledge that news is only one of many places providing politically relevant information, (3) how do the effects of exposure to political content outside news domains compare to those of exposure to news websites? Lastly, (4) how do these patterns differ across party and media systems?

With regard to the effects of the tracked exposures, we aim to offer a comprehensive evidence by testing both positive and negative outcomes. As aforementioned, exposure to news media and political content stimulates citizen engagement in a range of civic and political activities, from signing a petition to protesting (**e.g.**; Stromback & Shehata, 2010). Also, by showing how the democratic process works and how political decisions are made, news media – in the aggregate – should enhance political trust and support for political compromise, an essential part of democratic decision-making. And yet, despite these benefits, there are reasons to believe that news exposure may have a range of *adverse* effects, which have received little attention in past work and which – nevertheless—can theoretically be linked to news exposure. For one, negativity is one of the core journalistic values (Galtung & Ruge, 1965) and so news media tend to focus on clashes between political groups, feature uncivil debates (Levendusky & Malhotra, 2016; Martin & Yurukoglu, 2017), and cover politics as a game or a horse-race. This may lead people to see the system at large as failing (Cappella & Jamieson, 1996), the elites as evil (Mutz, 2007), and society as sharply divided (Levendusky & Malhotra, 2016; Mutz, 2006). Such negativity may also make individuals anxious or angry. Second, theories of public opinion formation suggest that exposure to elite cues and party communication can distort citizens' policy preferences (Mullinix, 2016; Nicholson, 2012) and make people's partisan identities more salient. Because citizens encounter these cues and (increasingly polarized) elite communication via media, news exposure can polarize attitudes (Druckman et al., 2013; Levendusky, 2013) and intensify out-group hostility (Levendusky & Malhotra, 2016; see also Wojcieszak et al., 2022, for a review).

Context

The patterns of news consumption, news avoidance, and incidental exposure are shaped by political context, national culture, and societal norms (Toff & Kalogeropoulos, 2020). In fact, news exposure and avoidance differ between countries (Newman et al. 2019; Villi et al.,

2022). To at least partly capture these contextual influences, we test our questions in three countries with different media and political systems: the U.S., the Netherlands, and Poland. Recent categorizations see the U.S. as a polarized liberal media system, with fragmented news markets, politicization of content and funding, and uneven professionalization (Brüggemann et al., 2014; Nechushtai, 2018). The U.S. is also a highly polarized two-party system. In turn, the NL is a stable democracy with a multi-party system, democratic corporatist media, a tradition of external pluralism stimulated by the government, but with
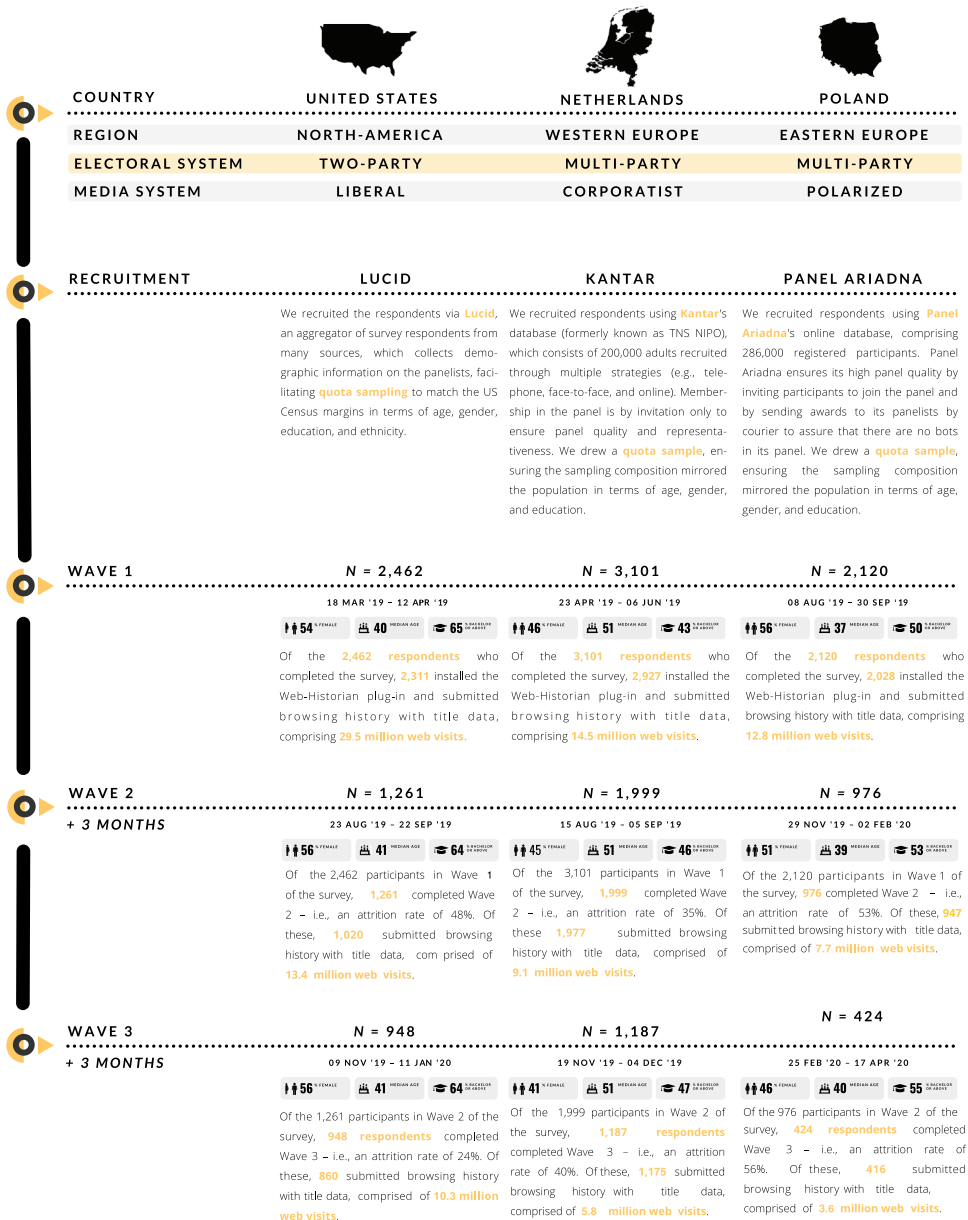

COUNTRY	UNITED STATES	NETHERLANDS	POLAND
REGION	NORTH-AMERICA	WESTERN EUROPE	EASTERN EUROPE
ELECTORAL SYSTEM	TWO-PARTY	MULTI-PARTY	MULTI-PARTY
MEDIA SYSTEM	LIBERAL	CORPORATIST	POLARIZED

RECRUITMENT	LUCID	KANTAR	PANEL ARIADNA
	We recruited the respondents via Lucid, an aggregator of survey respondents from many sources, which collects demographic information on the panelists, facilitating quota sampling to match the US Census margins in terms of age, gender, education, and ethnicity.	We recruited respondents using Kantar's database (formerly known as TNS NIPO), which consists of 200,000 adults recruited through multiple strategies (e.g., telephone, face-to-face, and online). Membership in the panel is by invitation only to ensure panel quality and representativeness. We drew a quota sample, ensuring the sampling composition mirrored the population in terms of age, gender, and education.	We recruited respondents using Panel Ariadna's online database, comprising 286,000 registered participants. Panel Ariadna ensures its high panel quality by inviting participants to join the panel and by sending awards to its panelists by courier to assure that there are no bots in its panel. We drew a quota sample, ensuring the sampling composition mirrored the population in terms of age, gender, and education.

WAVE 1	N = 2,462	N = 3,101	N = 2,120
	18 MAR '19 – 12 APR '19	23 APR '19 – 06 JUN '19	08 AUG '19 – 30 SEP '19
	54 % FEMALE 40 MEDIAN AGE 65 % BACHELOR OR ABOVE	46 % FEMALE 51 MEDIAN AGE 43 % BACHELOR OR ABOVE	56 % FEMALE 37 MEDIAN AGE 50 % BACHELOR OR ABOVE
	Of the 2,462 respondents who completed the survey, 2,311 installed the Web-Historian plug-in and submitted browsing history with title data, comprising 29.5 million web visits.	Of the 3,101 respondents who completed the survey, 2,927 installed the Web-Historian plug-in and submitted browsing history with title data, comprising 14.5 million web visits.	Of the 2,120 respondents who completed the survey, 2,028 installed the Web-Historian plug-in and submitted browsing history with title data, comprising 12.8 million web visits.

WAVE 2 + 3 MONTHS	N = 1,261	N = 1,999	N = 976
	23 AUG '19 – 22 SEP '19	15 AUG '19 – 05 SEP '19	29 NOV '19 – 02 FEB '20
	56 % FEMALE 41 MEDIAN AGE 64 % BACHELOR OR ABOVE	45 % FEMALE 51 MEDIAN AGE 46 % BACHELOR OR ABOVE	51 % FEMALE 39 MEDIAN AGE 53 % BACHELOR OR ABOVE
	Of the 2,462 participants in Wave 1 of the survey, 1,261 completed Wave 2 – i.e., an attrition rate of 48%. Of these, 1,020 submitted browsing history with title data, comprised of 13.4 million web visits.	Of the 3,101 participants in Wave 1 of the survey, 1,999 completed Wave 2 – i.e., an attrition rate of 35%. Of these 1,977 submitted browsing history with title data, comprised of 9.1 million web visits.	Of the 2,120 participants in Wave 1 of the survey, 976 completed Wave 2 – i.e., an attrition rate of 53%. Of these, 947 submitted browsing history with title data, comprised of 7.7 million web visits.

WAVE 3 + 3 MONTHS	N = 948	N = 1,187	N = 424
	09 NOV '19 – 11 JAN '20	19 NOV '19 – 04 DEC '19	25 FEB '20 – 17 APR '20
	56 % FEMALE 41 MEDIAN AGE 64 % BACHELOR OR ABOVE	41 % FEMALE 51 MEDIAN AGE 47 % BACHELOR OR ABOVE	46 % FEMALE 40 MEDIAN AGE 55 % BACHELOR OR ABOVE
	Of the 1,261 participants in Wave 2 of the survey, 948 respondents completed Wave 3 – i.e., an attrition rate of 24%. Of these, 860 submitted browsing history with title data, comprised of 10.3 million web visits.	Of the 1,999 participants in Wave 2 of the survey, 1,187 respondents completed Wave 3 – i.e., an attrition rate of 40%. Of these, 1,175 submitted browsing history with title data, comprised of 5.8 million web visits.	Of the 976 participants in Wave 2 of the survey, 424 respondents completed Wave 3 – i.e., an attrition rate of 56%. Of these, 416 submitted browsing history with title data, comprised of 3.6 million web visits.

Figure 1. Overview project flow.

guarantees of editorial freedom (Hallin & Mancini, 2004). Lastly, Poland is a newer democracy, a highly polarized post-communist country, with a liberal polarized pluralist media system, public service broadcasting controlled by the government, and one in which media-politics relations are tense since the 2015 elections, with legislation limiting press freedom. These three cases, although by no means extensive or representative, offer a chance to assess cross-country variations in behaviorally tracked exposures to news and to political content outside news.

Data and Methods

We rely on a 3-wave panel study, in which, every three months, the same participants in the three countries completed 20-minute surveys and submitted – after extensive informed consent – their browsing data. Sampling and recruitment in the US was done by Lucid, an aggregator of respondents from many sources that collects demographic information on the panelists, facilitating quota sampling to match the US Census margins. In the Netherlands, the random sample was drawn from Kantar's database, which consists of 200,000 adults recruited through multiple strategies (e.g., telephone, face-to-face, and online). Membership in the panel is by invitation only to ensure quality and representativeness. In Poland, we used Panel Ariadna that invites citizens to join its panel of 286,000 registered participants and sends awards to its panelists by courier to assure that there are no bots in its panel and that an individual does not register multiple times. Across the countries, quotas on age, gender, and education were enforced (and on ethnicity in the U.S.). Figure 1 shows the project's timeline, survey and trace data. Supplementary Materials (SM) C summarizes the demographics of the samples of all the waves as compared to the general populations and shows there was no significant attrition between waves across socio-demographics. The project was approved by the Ethical Board of the European Research Council (ERC) and University of Amsterdam (see SM B).

Digital Trace Data

At each wave, prior to taking the survey, participants submitted their browsing data via our open-source tool that allows for transparent data sharing, Web Historian. All the details on the tool, informed consent, data visualizations, and the steps taken by the participants are shown in SM A Web Historian is a browser extension that accesses respondents' browser history stored on their computers (we do not have mobile data, an issue we address in the discussion), displays it to them using interactive visualizations (e.g. network graph of websites visited, word cloud of used search terms, searchable table of browser history), and allows them to submit it to researchers following an extensive informed consent process (see SM A.4). Web Historian collects up to 90 days of one's browsing history, and so we use data that span three 90-day time periods before each of the three waves for a total of up to nine months of continuous individual-level trace data per country. In total, we have 27 months of browsing, matched with individual-level characteristics of the participants from three distinct democracies.

Behavioral Measures

Online News Exposure. Web Historian records data at the visit level, i.e., each visit is a record in the data and includes a timestamp, the URL of the site visited, and the title of the page. Having data at the visit level allows us to calculate how often participants visited news websites over the 270-day period per country by matching the visited domain (e.g., nytimes. com) to identifiable news domains on our lists of news organizations per country. We composed these lists from several sources: manually identified news domains from Alexa's Top 1000 web domains list; the 1000 most browsed domains in our trace data; and the 1000 most shared domains by politicians on Twitter. Given the size of the country and the numerous media markets in the U.S., we augmented the U.S. list with any missing news organizations from usnpl.com, a website that includes a large number of local news outlets.

Our U.S. list contains a total of 5,400 news organizations (of which 2041 were visited by the participants); the Dutch list contains 294 outlets (256 visited); the Polish list 298 organizations (291 visited). SM E.1 shows how the list was created, SM Table E.10 presents the list of news domains comprising 90–95% of visits, and https://github.com/ercexpo/ Github offers the publicly available lists. Furthermore, we account for the fact that citizens consume news on social media platforms. While we cannot access one's Facebook News Feed or Twitter timeline, so we do not know whether people merely saw news or political information on social media, we can see if participants visited a news website's Facebook page, a tweet from the Twitter handle of a news organization, or a video from the YouTube channel of a website on our list. We identified Facebook pages, Twitter handles, and YouTube channels of all the news organizations on our lists and assigned these exposures accordingly. Out of all the visits to these platforms, only 0.5% were to news organizations and only 34% of these on-platform news exposures were to political content.

Political exposure. To determine whether participants consumed political news within the news sites they visited and also saw political content outside news domains, we develop a multilingual BERT-based neural binary classifier. We trained the model using titles (i.e. the sentence-like string authored by the website and displayed on the browser 'tab'). SM E.2 presents all the details. Graduate annotators, two per country, manually labeled a total of 14,232 article titles for whether they were political or not. We conceptualize *politics* rather broadly, including references to both political figures, policies, elections, news events (e.g., impeachment inquiry, the primaries) *as well as* issues such as climate change, immigration, healthcare, gun control, sexual assault, racial, gender, sexual, ethnic, and religious minorities, the regulation of large tech companies, and crimes involving guns. Ten percent of the titles were jointly annotated (Krippendorff's Alpha .934).

These coded articles were then split into training, validation, and testing datasets. Our training set consisted of 10,629 article titles from news and non-news websites (3,956 from US sites, 3,534 from Dutch sites, and 3,139 from Polish sites). Our validation and test sets consisted of 1,181 and 2,422 titles, respectively, selected at random. The model was fine-tuned for four epochs with 10% of the remaining coded data set aside during training for validation. To account for the imbalance between our positive and negative classes, we trained with a 1:4 weighting. Additionally, we used a 0.1 dropout, weight decay of 0.05, a learning rate of 2e-5, and 10% warm-up. Finally, two annotators per country validated the output from the classifier, manually checking 1,156 titles. The annotators disagreed with

roughly 5% of the output, finding that the classifier may be more likely to label nonpolitical content as political than vice versa. This means that our study – if at all – over-reports exposure to political content.

In short, this extensively validated model allows us to categorize the content visited in news *and also* non-news domains as related to politics or not with high accuracy (93%, Precision .92, Recall .91, F1 .915). We use this classification to construct measures of exposure to political content inside *and* outside of news sites. To illustrate, if a person reads about a political event or issue (e.g., 'Azerbaijan army's music video reveals a weapon bought in secret') on a non-news domain, this visit is classified as political content outside of news domains. Had the person read this story on *CNN* or *FOX*, this particular visit would have been classified as political exposure within a news domain. In the main analyses, news homepages are not categorized as political news visits and SM F.3 presents the results considering homepages as such because news homepages expose users, at least cursorily, to political news.[3]

We calculate two exposure measures to show that our results hold when relying on alternative measurements. First, we assign to each participant a score consisting of the total number of unique URLs per day visited within the identified news sites, so as to not count page reloads and other browser or website idiosyncrasies as additional visits. Second, we calculate the time spent on these websites. We consider a visit to a new page as beginning the time spent on that page until the next visit occurs, with a five minute inactivity timeout, a standard way of adjusting for the periodic nature of web browsing events when constructing a time-based behavioral measurement (Council & Bureau, 2013). Across all countries and waves, the median time span between visits was 5 seconds, and 3.3% of visits had a time span of 5 minutes or more.

Outcome Measures

We estimate the effects of exposure to political content outside news and to news domains on a range of democratically relevant attitudes, cognitions, and behaviors. In SM (and also to (c) political content within news. SM D.1 details question wording and summary statistics, including distributions of both raw measures and constructed indices.

To measure *participation*, we asked the participants which of nine political activities they are likely to undertake in the near future (e.g., "sign a petition") and we summed the individual activities. To assess *support for compromise*, we averaged answers to four statements (e.g., "I want politicians who work together"). We also averaged the responses to questions about participants' *trust* in several institutions (e.g., the government, police, courts). *Attitude polarization* was measured by averaging responses to items gauging whether participants agreed more with a liberal (minimum) or a conservative (maximum) stance on salient issues or policies (e.g., climate change, immigration, economy). We folded each item so that moderate positions indicate a low level, and the extreme positions a high level of attitude polarization. We used two measures of *affective polarization*: a feeling thermometer toward supporters of the opposite party (or the most distant party in PL and NL), and also toward "people who oppose your views" on a set of issues per country. We reverse-coded these thermometers so that a higher value signifies a more negative feeling.

To gauge *misinformation endorsement*, we included a set of true and false statements in each country (e.g., "President Trump's grandfather was a pimp and a tax evader"). Our

measure is an index of incorrect answers so that a higher value represents being more misinformed. We also assessed two perceptions of the political system: *attribution of malevolence* that averaged responses to five questions asking how much respondents think that the opposing party wants to hurt the country and *perceived polarization* that averages responses to four questions asking participants how much they perceive the political climate as polarized.

Analytical Approach

To answer the descriptive questions, we use W1 trace data. The over-time analyses include those participants whose online data contained recorded visits on at least 7 days within each wave. In these analyses, we rely on the 9 months of trace data per country. We use trace data from the 90 days before W1 to construct W1 measures, data from the approximately 90 days before W2 to construct the W2 measures, and data from the approximately 90 days between W2 and W3 to construct W3 exposure measures. We assign to each participant a score based on the total number of unique URLs per person per day to the identified news sites.

In the predictive analyses, we rely on multilevel random intercepts models to estimate the effects while accounting for the nested structure of the data (subjects nested in waves), thereby "controlling" for the individual's typical level on the outcome variable. In terms of temporal order, each observation in our data has a value for the predictor (i.e., exposure) and a value for the outcome, with the former measured before the outcome (i.e., the browsing data capture the three months before the survey). As such, any association between exposure and the tested outcome of the same wave cannot be due to reverse causality, although we cannot exclude the possibility that the outcome of a former wave influences the predictor of a later wave.

For each outcome, we estimate two models: one testing the effects of exposure to political content outside news domains and another testing effects from exposure to news domains. This is because we are interested in the effects of each behavior and in comparing them to each other, not in whether political non-news exposure has effects above and beyond news exposure, or vice versa. In each model, we control for the overall number of deduplicated visits for each person/wave to isolate the specific exposure type tested from a person's overall browsing. We also include gender, age, and education as covariates. We adjust the significance tests for multiple comparisons controlling for the false discovery rate (Benjamini et al., 2006). Coefficients are presented as regression tables in SM G.1. We scale all outcome variables (except participation) to a range between 0 and 100, so that coefficients denote the one percentage-point increase in the outcome for a one-unit increase in the predictor. Since participation is measured with count variable (between 0 and 9), we use a negative binomial multilevel model for this outcome. Here, for a one-unit increase in the predictor, the logs of expected counts of the outcome is expected to change by the coefficient. We transformed all exposure variables by adding one and taking the log.

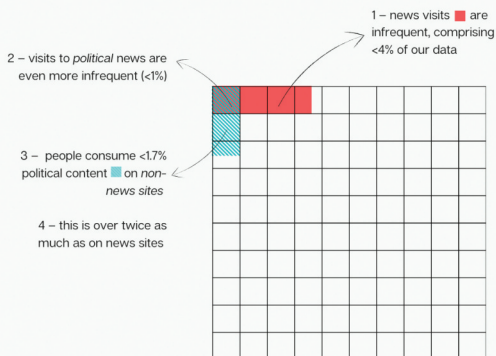| Summary Table	United States	Netherlands	Poland
No. Respondents Wave 1	2,311	2,927	2,028
All visits	29,482,222	14,548,136	12,829,578
News visits (%)	694,129 (2.4%)	610,545 (4.2%)	476,246 (3.7%)
Political news visits (%)	250,140 (0.8%)	88,235 (0.6%)	128,111 (1.0%)
Non-news visits with political content (%)	597,526 (2.0%)	177,706 (1.2%)	203,461 (1.6%)

Figure 2. Relative prevalence of the tested exposures.

Results

Figure 2 summarizes browsing data from 7,266 participants from the three countries. These data contained over 56.9 M unique visits in Wave 1. The median participant provided browsing data spanning 88 days, actively browsed on 57 days, and visited 247 different domains. Below, we present results based on the number of visits to unique URLs per day per respondent and SM Figure G.9 shows robustness checks using time spent. As this key descriptive evidence shows, on average only 3.4% of visits comprised visits to news sites. The significant differences between the countries show greater online news consumption in the two European systems than in the U.S. (NL 4.2%, PL 3.7% vs US 2.4%; $p < .001$). Even there, however, online news exposure is *not* frequent.

As Figure 2 also details, an even smaller fraction of trace data comprised visits to political news. Our extensively validated classifier categorized the visited 56.8 M titles within all domains as related or not related to political issues or events (see SM E.2 for details). On average, only 25.8% of the visits were to political content within news sites, less than 1% of all visits. Again, there were significant between-country variations. The Dutch, who visited news the most, read political news the least (only 14% of news browsing was political news), followed by the Poles (27%) and the Americans, who read the least news but most political news (36%). On average, the majority of visits were to nonpolitical articles within news sites, such as sports or weather. In short, 96.6% of what people saw online was not news and most news visits were not to political news.[4] These bleak patterns offer an incomplete picture, however. Do people read about politics outside news sites (e.g., an article about racial discrimination in *Sports Illustrated*)? Across the countries, 1.6% of non-news visits were to content related to politics. The U.S. had the greatest percentage of political browsing outside of news (2.0% of all visits), followed by PL (1.6%), and the NL (1.2%). Although this seems like a small percentage, it is *higher* than that of exposure to political news in the aggregate browsing. Given the immense popularity of non-news websites, in absolute terms, exposure to politics outside news is *greater* than exposure to politics within news pages, as seen in Figure 3. Whereas an average participant read only one political news article for every 122 web pages, they read one political non-news piece for every 58 pages visited. Especially in the U.S., but also in the two European countries, exposure to politics outside news is more

Summary Figure

1 – news visits ▮ are infrequent, comprising <4% of our data

2 – visits to *political* news are even more infrequent (<1%)

3 – people consume <1.7% political content ▮ on *non-news* sites

4 – this is over twice as much as on news sites

1 – Less than 4% of our data are visits to *news sites*

> other
> news sites

Citizens rarely visit news sites. In the Netherlands, 610,545 out of 14,548,136 visits were visits to news websites (4.2%); In Poland, 476,246 out of 12,829,578* visits were visits to news websites (3.7%); and in the United States, 694,129 out of 29,482,222 were visits to news websites (2.4%).

UNITED STATES	2.4%
NETHERLANDS	4.2%
POLAND	3.7%
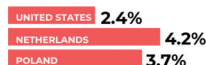
*The total number of visits reported here, only comprise visits for which we were able to obtain the title. This is approximately 89% of all visits we recorded.

2 – Less than 1% of our data are *political news*

> other
> political news

Consumption of political content on these news sites is even more infrequent. In the Netherlands, 87,919 visits (14.4% of news visits, 0.6% of all visits) were visits to political content on news websites; In Poland, 118,585 visits were visits to news websites (24.9% of news visits, 0.9% of all visits); and in the United States, 248,498 visits were visits to news websites (35.8% of news visits, 0.8% of all visits).

UNITED STATES	0.8%
NETHERLANDS	0.6%
POLAND	0.9%
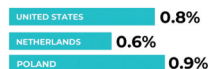
3 – More than 1.6% of our data are visits to *political content on non-news sites*
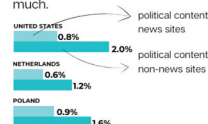
> political content non-news sites
> other
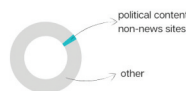
Citizens also consume political content on non-news sites. In the Netherlands, 174,578 visits were political visits on non-news sites (1.2%); In Poland, 205,273 visits were political visits on non-news sites (1.6%); and in the United States, 589,644 were political visits on non-news sites (2.0%).

UNITED STATES	2.0%
NETHERLANDS	1.2%
POLAND	1.6%

4 – More than twice as much *political content* on non-news sites than on news sites

> political content non-news sites
> political content news sites
> other

Finally, our data shows that people consume around twice as much political content on non-news sites than on news sites. In the Netherlands this is 2x as much; in Poland, this is 1.8x as much; and in the United States, this is 2.5x as much.

> political content news sites
> political content non-news sites

UNITED STATES	0.8%	2.0%
NETHERLANDS	0.6%	1.2%
POLAND	0.9%	1.6%

Figure 3. Prevalence of exposure to news, political content within news domains, and political content outside news domains.

prominent than within news.[5] To offer some insight into the tested phenomena, we provide a list of the main non-news websites in each country that expose individuals to political information in SM E.13, E.14, and E.15.

Additional Analyses

As robustness checks, we calculate the length of time spent at each page to test if the findings hold when changing from a visit-based to a time-based measurement approach. As SM Figure G.9 details, all the results are nearly identical when using duration rather than the number of visits. To offer additional insight, we also show that the tested exposures were relatively stable during the nine months of data collection. SM Figure F.6 shows that exposures to news outlets, political news, and politics outside news all increased during the elections in the NL and PL, but were otherwise largely unaffected by other country-specific and international events. Furthermore, we examine ideological asymmetries in the tested patterns (see SM F.4), finding that the political left

and the political right consumed roughly the same amounts of news and politics outside news (with the aggregate and within-country differences not surpassing .3%; see SM Table F.23).

Also, we assess whether there are differences in these exposures among different groups of citizens, focusing on variables traditionally associated with lower news use (i.e., political interest, education, age, gender, and race in the U.S.). The work on incidental exposure suggests that people with lower political interest and those who use less news are more likely to consume infotainment and inadvertently encounter politics online. Indeed, this was the case in our data. Citizens with low political interest, as determined based on W1 self-reports (SM Table F.20) and behavioral traces indicating the extent of exposure to online news websites (SM Table F.21), consistently encounter more politics outside news sites than within. The patterns for the other demographics are different, with the more highly educated (apart from the NL) and the older individuals encountering more politics outside news; the differences for gender are less pronounced (apart from PL; SM F.16, F.17, F.18, F.19 details the results).

Last but not least, we present the descriptive results after removing survey-taking visits from our data. Due to the nature of online samples and their incentive structure, people participating in an online panel of one survey organization may be participants in numerous other panels. This is particularly true of respondents recruited through a panel aggregator, such as Lucid, the company we used in the U.S. The presence of survey-taking visits may influence our conclusions about the prevalence of visits to news and political content outside news domains. We identified these survey visits and re-ran our descriptive and predictive analyses with these visits excluded. SM E.3 details our approach to identifying these sites and SM F.6 shows the percent of news and political non-news visits after removing survey visits. Removing these made the percent of news visits (4.4%) and of political visits outside news domains (US 1.4%, NL 1.2%, PL 1.5%) relatively equal across
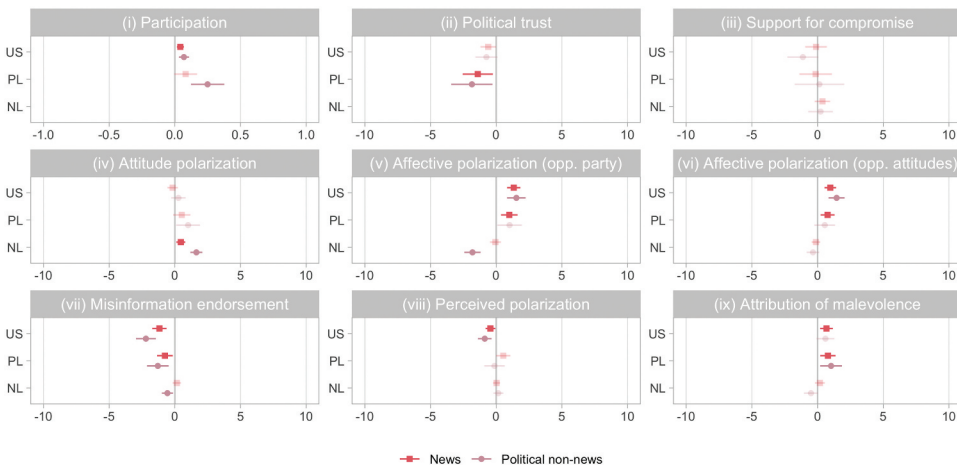
Figure 4. Effects of exposure to news and to political content outside news. Coefficients are from two different random-intercept models, each regressing the outcome on a type of exposure (news or political non-news), as well as a set of controls. These two models correspond to "model 1" and "model 2" in the full regression tables in SM G.1.

the countries, without changing the overall conclusion that political visits outside news expose people to politics more frequently than news.

Democratic Effects

We now examine how these exposures influence key outcomes, each measured using multiple indicators across the waves. Full regression tables are shown in SM G.1. Figure 4 illustrates the results. Each panel shows an outcome measure, and plots the coefficient estimate of interest for each model, grouped by country. Estimates are shown as transparent when they are not statistically significant at the 95% level *after* FDR adjustment. The figure plots the effects of exposure to political content outside news, our core interest, as compared to news exposure (modeled separately). A consistent picture emerges: over-time exposure to political content outside news predicts as strongly – and in some cases even more strongly – all the tested outcomes as news exposure.

Addressing the beneficial outcomes, exposure to political content outside news predicts increased participatory intentions (e.g. protesting or signing a petition) in the U.S. and PL (Facet *i*).[6] This over-time effect is as strong as that of exposure to news. As this count outcome is modeled as a negative binomial, a coefficient of 0.4 for political non-news in PL means that someone with one unit more of logged political non-news exposure would have 20% higher participation, all else constant ($e^{0.4} = 1.2$). We find evidence for a negative relationship between exposure to political content outside news and trust in PL (Facet *ii*). For this outcome (and all the remaining ones), the meaning of a coefficient of, for example, −1.84 (political non-news in PL) is that a one-unit decrease in exposure to political content outside news is associated with a reduction of 1.84 points on a scale from 0 to 100, all else constant. No significant effects on support for compromise emerge in any country (Facet *iii*). In short, tracked online exposure to politics outside news and to news have parallel effects on political engagement and limited over-time effects on trust and support for compromise.

We turn to negative outcomes, testing if the tracked exposures predict attitude polarization over-time (i.e., changes in attitude extremity on five issues per country). In the U.S. and PL, we find null relationships, while in the NL exposure to political content outside news predicts *increases* in attitude extremity (Facet *iv*). This relationship in the NL is comparable and even stronger than that with news exposure. We also consider affective polarization, i.e., hostility toward out-partisans and citizens with opposite policy beliefs, each measured in three ways, see SM D.1. In addition to accounting for different facets of affective polarization, using multiple measures also ensures that the detected patterns are not due to any specific measurement alone and that the results are robust to contexts and out-groups. In the U.S. and PL, exposure to political content outside news increases hostility toward out-party supporters, and – in the U.S. – also toward those of with opposing views. These effects are similar to those of news exposure. In the NL, in contrast, exposure to politics outside news predicts decreases in affective polarization toward out-party supporters (Facet *v* and *vi*).

Further, in all three countries, news exposure and even more so political exposure outside news decreases misinformation endorsement (Facet *vii*). Findings regarding perceived polarization show that in the U.S., exposure to politics outside news decreases

perceived polarization, but exposure to news does not. Lastly, in PL, exposure to news and to politics outside news predicts increased attribution of malevolence. In the U.S., only news exposure has this negative effect.

To contextualize these findings, we estimated the marginal means for significant effects.[7] These numbers illustrate how exposure to political content outside news often exerts stronger effects than news exposure, for both positive and negative outcomes. For instance, for participation in the U.S., the predicted difference between exposure at one SD below average and one SD above the mean is 0.23 for political content outside news and 0.18 for news. The analogous differences in terms of misinformation endorsement in NL is 3.85 for political content outside news and only 2.79 for news (on the scale from 0–100). The same holds true for the negative outcomes. For instance, in the NL, the difference in attitude polarization between someone with one SD below political exposure outside news (51.9) and a person with exposure one SD above the mean (57.0) is more than 5 points. The analogous shift from low to high levels of news exposure is only from 53.6 to 55.6. These numbers illustrate that consuming politics outside news domains can have a tangible impact on key outcomes, whether positive or negative.

Additional Analyses

To again check the robustness of the effects, SM Figure G.9 shows that the coefficients are nearly identical in direction, magnitude, and significance when we use the time spent on each page rather than the number of visits as the predictor. We also re-estimated models among both the left- and right-leaning participants (see SM G.1) by adding cross-level interactions. We find largely insignificant interactions, with a few exceptions shown in SM G.1. In addition, we also estimate models with both political exposure outside news and news exposure as predictors in the same model (SM Figure G.9), which examines the effect of one exposure while controlling for the other (although it is vulnerable to the issues of multicollinearity). Even when controlling for news exposure outside news, we still find effects of the latter, even though these effects are slightly less pronounced. Furthermore, in the tables in SM G.1, we also compare the effects of political exposure outside news to those of exposure to political *and* nonpolitical content within news. Lastly, we also present the models re-estimated after excluding survey-taking visits in SM Figure G.1. This alternative approach does not change the direction or size of coefficients substantively, although some become more, and others less, statistically significant.

Discussion

For centuries, democratic theorists have viewed news media as crucial to democracy, for decades, social scientists have tested the benefits of news consumption, and for years public observers have worried about the polarizing tone of (partisan) news coverage. Yet-—as we argue-—this focus is limited given that many citizens in the U.S. and elsewhere are *not* interested in news and do not go online for politics. Relying on large-scale data combining over-time survey self-reports with a total of 27 months of online browsing from three countries ($N = 7,266$; 106 M visits), we offer a comparative evaluation of citizens' encounters with politics beyond news websites.

We find that citizens prefer online entertainment, shopping sites, and celebrity gossip over news and public affairs. That may be nothing new (Prior, 2005). Yet, our data also show that citizens' visits to non-news sites are the dominant source of political information. Even though politics in these kinds of sites comprised only 1.6% of all visits, the aggregate popularity of webmail, entertainment, shopping sites, or celebrity gossip means that an average citizen encounters most political content *outside* news. People, especially Americans and especially those with low political interest (as shown in SM Table F.20 and F.21), encounter politics more frequently outside news outlets than within (Wojcieszak & Mutz, 2009).

In contrast, the results regarding news exposure are bleak. Again, this finding is not new, but it is insufficiently emphasized in public debates and academic research. Across the countries, whether two-party (U.S.) or multi-party (NL, PL) systems, highly polarized (U.S, PL) or not (NL), with strong (NL) or weak (U.S.) public service broadcasting, news is a drop in the overall ocean of what citizens do online (3.4% of the overall browsing). We furthermore note that political content within news sites was even a smaller drop in this ocean of content: visits to news about political issues and events comprised less than 1% of all visits and 25% of all news visits in our data (and roughly 50% when also including homepage visits as political news visits). Many news visits were to check sports or find an oatmeal cookie recipe. Again, most political exposures in our data occur outside news.

In fact, we suspect that *political* exposures outside news are greater than what we present. We cannot capture political information people merely see on social media (e.g., embedded videos, headlines, posts) and instead examine a more meaningful engagement with content (i.e., clicking on the URL and landing on a story). Because nonpolitical content on platforms communicates political information (Settle, 2018) and political memes are popular, we may be under-estimating exposure to political content outside news. Although the same could be said about *news* exposure, the lack of social media data may not underestimate its volume in any dramatic way. Not only do we identify and include exposure to news organizations on Twitter, Facebook, and YouTube, but also our low estimates are consistent with other trace data. As mentioned, only 1.4% of Facebook's News Feed is news (Meta, 2022), public affairs comprise 1.8% of the average News Feed of students (Wells & Thorson, 2017), and only about 1 in 300 outbound clicks from Facebook correspond to substantive news (Flaxman et al., 2016; see also Karnowski et al., 2017; Vermeer et al., 2020). In our data, although social media browsing from Twitter, Facebook, and YouTube made up 8.7% of all visits, only 0.5% of those visits were to the identified news organizations, and these made up only 1.2% of all news visits overall. In short, the lack of social media data may lead us to underestimate the volume of exposure to *political* content on platforms without significantly underestimating *news* exposure.

That said, we acknowledge that – as similar studies – we do not account for the over-arching information ecology of our participants, such as their news exposure offline or on mobile devices. For most people, television is the dominant news source (Allen et al., 2020) and we do not have access to these exposures offline, as occurring in one's household. In addition, many people self-report using mobile for news (Newman et al., 2022) and we do not have access to these exposures. Including mobile, we speculate, would not change the patterns presented, as news exposures are similarly low on desktop and mobile (Allen et al., 2020), and would not allow us to address our core questions about *political* exposures (i.e., inasmuch as most people use apps, mobile data are domain-level only). In sum, researchers

cannot access all the information about all the outlets people see, and likely never will. Studies that have more complete data, e.g, TV, mobile, and online browsing (Allen et al., 2020), can only make aggregate claims, and individual-level studies using online traces, such as ours, have to accept that these traces do not represent the totality of what users read or hear online and offline. Because an average American spends over 7 hours a day online (Moody, 2021), with this figure being comparable in Poland (Papilot.pl, 2021), and because opportunities for easy access to news and for incidental exposure to politics are greater online than in the "traditional" media environment, our data – although incomplete – still offer important insights.

Lastly, we show that exposures to political content outside news and to news have similar over-time effects on various outcomes. In short, both enhance participatory intentions and lower misinformation endorsement. Yet, at the same time, both kinds of exposures decrease trust in PL, likely due to numerous controversies related to the government, which aims to curtail the freedom of the courts, press, and citizens. Exposure to these efforts and their criticism by the EU may undermine citizens' trust in state institutions. In addition, the tested exposures have a range of aforementioned negative effects, such as exacerbating attitude polarization and affective polarization. Similarly, we find suggestive evidence that exposure to political content outside news and to news enhances attribution of malevolence and – in PL – also perceived polarization.

We suspect that the nearly identical effects of these two types of exposures are due to the fact that non-news content about politics also conveys elite cues, which activate people's partisan identities and influence policy preferences, out-group attitudes, and perceptions of the system (Druckman et al., 2013; Mullinix, 2016; Nicholson, 2012). Our findings suggest that exposure to these elite cues can have similar consequences regardless of where these cues are encountered, whether in news or non-news sites. In turn, we speculate that the between-country differences in the effects are due to country-specific events and the idiosyncrasies of national politics. Future research should collect online traces in more countries to offer larger scale comparative evidence and complement the rich self-reported data on news exposure/avoidance offered by the Reuters Institute Digital News Report (Newman et al. 2019).

Our findings have key implications. Much of what scholars and observers debate is foreign to the daily lives of a majority of citizens. The consumption of news, vastly over-reported in surveys and neatly isolated in most experiments, accounts for a small fraction of people's online activities. It is the other 96.6% of these activities that carry the potential to introduce people to politics. Given the popularity of entertainment, these non-news sites – as newsreels in the cinema and soft news on TV – reach the largely disengaged citizens who can learn about COVID vaccine as a by-product of following Australian Open on a sports website. Inasmuch as such encounters with politics have some potential to pull disengaged citizens into the democratic process (as suggested by its dominance among the disinterested citizens and its positive links with participation), these encounters could have cumulative benefits. This is especially as many citizens are turning away from news, partly because it is depressing, negative (Galtung & Ruge, 1965), and disproportionately focuses on political conflict (Levendusky & Malhotra, 2016; Martin & Yurukoglu, 2017; Mutz, 2006).

Estimating these over-time effects is an important direction for future work. In addition, questions such as "how to encourage greater political exposure?" or "how to make news more personally relevant?" should come to the forefront of scholarly endeavors. Public

observers worry about partisan news coverage and the growing polarization between the political left and the right in the U.S. and beyond. Yet these worries may be missing the forest for the trees as the primary divide (Krupnikov & Ryan, 2022) is between the news junkies and those who withdraw from news and politics altogether. Pulling the latter back into the democratic process, therefore, has the potential of minimizing the (illusion) of polarization (Druckman et al., 2021) and also making the political system more equitable and representative.

Notes

1. The dataset is available at https://osf.io/nd2wj/?view_only=870a157ed80f42aaba08b14b58a1 b2daOSF. It includes data from 7,428 people, 7,266 of whom submitted sufficient data for analysis in wave 1 of the study.
2. Since we have three samples, we use the average percent across countries.
3. Home pages visits are defined as visits to our news URLs with an empty path component of the URL. URL components are the following: [protocol:][//host[:port]][path][?query][#fragment] (IBM Corporation, 2021), thus the URL https://www.nytimes.com/live/2021/12/02/world/ biden-omicron-variant-covid?source=facebook would have the protocol "https," the host "www.nytimes.com" the path "/live/2021/12/02/world/biden-omicron-variant-covid" and the query "source=facebook," with exceptions for portal sites with separate news home pages, and sites that use query strings to fetch articles (the replication code will be made available on https://github.com/Github).
4. SM Figure F.7 categorizes news homepage visits as 'political news exposure.' Even when news homepages are categorized as political news visits, only 44% of visits to news websites expose users to political news, roughly 1.5% of all visits.
5. Amongst the participants who submitted at least 7 days of data in Wave 1 (US = 2,220, NL = 2,761, PL = 1,899), many had zero or only one visit to news (US = 11%, NL = 14.6%, PL = 6.1%). In contrast, many fewer had no or only one visit to political content outside of news (US = 2.3%, NL = 6% PL = 1.3%).
6. This outcome was not measured in the NL.
7. Regarding the control variables, we predict values for a woman with high education level, at median age and with average total browsing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Notes on contributors

Magdalena Wojcieszak (Ph.D. University of Pennsylvania) is a Professor of Communication, Affiliate Faculty, Designated Emphasis in Computational Social Science and Member, Graduate Group in Computer Science at the University of California Davis, and an Associate Researcher (ERC EXPO, Principal Investigator) University of Amsterdam. Her research interests include political communication and (social) media use and effects.

Ericka Menchen-Trevino (Ph.D. Northwestern University) is an Assistant Professor at the American University. Her research interest include research methodology, digital media and political communication.

Bernhard Clemm von Hohenberg is a Research Associate at the GESIS Leibniz Institute for the Social Sciences. His research in computational social science focuses on misinformation and media effects, using experimental methods, surveys and trace data.

Sjifra de Leeuw (Ph.D. University of Amsterdam) is a Senior Policy Advisor at the Dutch Ministry of Finance in Den Haag, the Netherlands.

João Gonçalves (Ph.D. Universidade do Minho) is an Assistant Professor at Erasmus University Rotterdam. His research focuses on AI, computational social science, and algorithms.

Sam Davidson is a Ph.D. student in the Linguistics Department at UC Davis focusing on computational language modeling and natural language processing.

Alexandre Goncalves (Ph.D. Columbia University) is an Assistant Professor of Journalism at the College of Media, University of Illinois Urbana-Champaign. His research at the intersection of computer science and communication focuses on the role of social media in the rise of new social and political movements.

ORCID

Magdalena Wojcieszak (iD) http://orcid.org/0000-0001-5456-4483
Ericka Menchen-Trevino (iD) http://orcid.org/0000-0002-5029-8269

Data availability statement

Data and code to reproduce all results are available at https://osf.io/nd2wj/?view_only=870a157ed80f42aaba08b14b58a1b2da.

Open scholarship

This article has earned the Center for Open Science badges for Open Data and Open Materials through Open Practices Disclosure. The data and materials are openly accessible at https://osf.io/nd2wj/?view_only=870a157ed80f42aaba08b14b58a1b2da.

References

Aalberg, T., Blekesaune, A., & Elvestad, E. (2013). Media choice and informed democracy: Toward increasing news consumption gaps in Europe? *The International Journal of Press/politics*, *18*(3), 281–303. https://doi.org/10.1177/1940161213485990

Achen, C. H., & Bartels, L. M. (2006). It feels like we're thinking: The rationalizing voter and electoral democracy. *Annual Meeting of the American Political Science Association, Philadelphia*, PA, USA, 30.

Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, *6*(14), 1–6. https://doi.org/10.1126/sciadv.aay3539

Andersen, K. (2019). An entrance for the uninterested: Who watches soft news and how does it affect their political participation? *Mass Communication and Society*, *22*(4), 487–507. https://doi.org/10.1080/15205436.2019.1585544

Baek, Y. M., & Wojcieszak, M. E. (2009). Don't expect too much! learning from late-night comedy and knowledge item difficulty. *Communication Research*, *36*(6), 783–809. https://doi.org/10.1177/0093650209346805

Bartels, L. M. (1996). Uninformed votes: Information effects in presidential elections. *American Journal of Political Science*, *40*(1), 194–230. https://doi.org/10.2307/2111700

Baum, M. A. (2003). Soft news and political knowledge: Evidence of absence or absence of evidence? *Political Communication*, *20*(2), 173–190. https://doi.org/10.1080/10584600390211181

Baum, M. A., & Jamison, A. S. (2006). The Oprah effect: How soft news helps inattentive citizens vote consistently. *The Journal of Politics*, *68*(4), 946–959. https://doi.org/10.1111/j.1468-2508.2006.00482.x

Beckers, K., Van Aelst, P., Verhoest, P., & d'Haenens, L. (2021). What do people learn from following the news? a diary study on the influence of media use on knowledge of current news stories. *European Journal of Communication*, *36*(3), 254–269. https://doi.org/10.1177/0267323120978724

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, *93*(3), 491–507. https://doi.org/10.1093/biomet/93.3.491

Brennan, J. (2017). Against democracy. In *Against democracy*. Princeton University Press.

Brüggemann, M., Engesser, S., Büchel, F., Humprecht, E., & Castro, L. (2014). Hallin and Mancini revisited: Four empirical types of western media systems. *Journal of Communication*, *64*(6), 1037–1065. https://doi.org/10.1111/jcom.12127

Cappella, J. N., & Jamieson, K. H. (1996). News frames, political cynicism, and media cynicism. *The Annals of the American Academy of Political and Social Science*, *546*(1), 71–84. https://doi.org/10.1177/0002716296546001007

Chaffee, S. H., Zhao, X., & Leshner, G. (1994). Political knowledge and the campaign media of 1992. *Communication Research*, *21*(3), 305–324. https://doi.org/10.1177/009365094021003004

Council, M. R., & Bureau, I. A. (2013). *Audience Reach Measurement Guidelines Version 2.0*. https://www.iab.com/wp-content/uploads/2015/06/AudienceReachMeasurementGuidelines.pdf

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

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, *65*(2), 135–150. https://doi.org/10.1086/257897

Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2021). (Mis-)estimating affective polarization. *The Journal of Politics*, *715603*. https://doi.org/10.1086/715603

Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How elite partisan polarization affects public opinion formation. *American Political Science Review*, *107*(1), 57–79. https://doi.org/10.1017/S0003055412000500

Feezell, J. T. (2018). Agenda setting through social media: The importance of incidental news exposure and social filtering in the digital era. *Political Research Quarterly*, *71*(2), 482–494. https://doi.org/10.1177/1065912917744895

Feldman, L., Stroud, N. J., Bimber, B., & Wojcieszak, M. E. (2013). Assessing selective exposure in experiments: The implications of different methodological choices. *Communication Methods and Measures*, *7*(3–4), 172–194. https://doi.org/10.1080/19312458.2013.813923

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption [publisher: Oxford academic]. *Public Opinion Quarterly*, *80*(S1), 298–320. https://doi.org/10.1093/poq/nfw006

Fording, R. C., & Schram, S. F. (2017). The cognitive and emotional sources of trump support: The case of low-information voters. *New Political Science*, *39*(4), 670–686. https://doi.org/10.1080/07393148.2017.1378295

Galtung, J., & Ruge, M. (1965). The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of Peace Research*, *2*(1), 64–90. https://doi.org/10.1177/002234336500200104

Guess, A. M. (2021). (Almost) everything in moderation: New evidence on Americans' online media diets. *American Journal of Political Science*, *65*(4), 1007–1022. https://doi.org/10.1111/ajps.12589

Guess, A. M., Barberá, P., Munzert, S., & Yang, J. (2021). The consequences of online partisan media. *Proceedings of the National Academy of Sciences*, *118*(14). https://doi.org/10.1073/pnas.2013464118

Hallin, D. C., & Mancini, P. (2004). *Comparing media systems: Three models of media and politics [Google-books-ID: 954NJChZAGoC]*. Cambridge University Press.

Heiss, R., & Matthes, J. (2019). Does incidental exposure on social media equalize or reinforce participatory gaps? evidence from a panel study. *New Media & Society*, *21*(11–12), 2463–2482. https://doi.org/10.1177/1461444819850755

Huang, S., & Yang, T. (2022). No trade-offs between news and entertainment: Evidence from online engagement data. *New Media & Society*, 14614448211063900. https://doi.org/10.1177/14614448211063899

IBM Corporation. (2021). The components of a url. https://prod.ibmdocs-production-dal-6099123ce774e592a519d7c33db8265e-0000.us-south.containers.appdomain.cloud/docs/en/cics-ts/5.2?topic=concepts-components-url

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. https://doi.org/10.1016/j.chb.2017.06.041

Katz, E., Blumler, J. G., & Gurevitch, M. (1973). Uses and gratifications research. *Public Opinion Quarterly*, *37*(4), 509. https://doi.org/10.1086/268109

Kenski, K., & Stroud, N. J. (2006). Connections between internet use and political efficacy, knowledge, and participation. *Journal of Broadcasting & Electronic Media*, *50*(2), 173–192. https://doi.org/10.1207/s15506878jobem5002_1

Klar, S., Krupnikov, Y., & Ryan, J. B. (2018). Affective polarization or partisan disdain? untangling a dislike for the opposing party from a dislike of partisanship. *Public Opinion Quarterly*, *82*(2), 379–390. https://doi.org/10.1093/poq/nfy014

Krupnikov, Y., & Ryan, J. B. (2022). *The other divide [Google-books-ID: qnRUEAAAQBAJ]*. Cambridge University Press.

Kwak, N., Lane, D. S., Weeks, B. E., Kim, D. H., & Lee, S. S. (2020). Now we're talking? understanding the interplay between online selective and incidental exposure and their influence on online cross-cutting political discussion. *Social Science Computer Review*, *40*(3), 579–597. https://doi.org/10.1177/0894439320920790

Lemert, J. (1984). News context and the elimination of mobilizing information: An experiment. *Journalism Quarterly*, *61*(2), 243–259. https://doi.org/10.1177/107769908406100201

Levendusky, M. S. (2013). Why do partisan media polarize viewers? *American Journal of Political Science*, *57*(3), 611–623. https://doi.org/10.1111/ajps.12008

Levendusky, M. S., & Malhotra, N. (2016). Does media coverage of partisan polarization affect political attitudes? *Political Communication*, *33*(2), 283–301. https://doi.org/10.1080/10584609.2015.1038455

Martin, G., & Yurukoglu, A. (2017). Bias in cable news: Persuasion and polarization. *American Economic Review*, *107*(9), 2565–2599. https://doi.org/10.1257/aer.20160812

Menchen-Trevino, E. (2016). Web historian: Enabling multi-method and independent research with real-world web browsing history data. https://doi.org/10.9776/16611

Merten, L. (2021). Block, hide or follow—personal news curation practices on social media. *Digital Journalism*, *9*(8), 1018–1039. https://doi.org/10.1080/21670811.2020.1829978

Meta. (2022). Widely viewed content report: What people see on facebook. https://transparency.fb.com/data/widely-viewed-content-report/

Moody, R. (2021). Screen time statistics: Average screen time in us vs. the rest of the world. https://www.comparitech.com/tv-streaming/screen-time-statistics/

Moy, P., Xenos, M. A., & Hess, V. K. (2005). Communication and citizenship: Mapping the political effects of infotainment. *Mass Communication and Society*, *8*(2), 111–131. https://doi.org/10.1207/s15327825mcs0802_3

Mullinix, K. (2016). Partisanship and preference formation: Competing motivations, elite polarization, and issue importance. *Political Behavior*, *38*(2), 383–411. https://doi.org/10.1007/s11109-015-9318-4

Mutz, D. C. (2006). *How the mass media divide us*. Brookings Institution Press.

Mutz, D. C. (2007). Effects of "in-your-face" television discourse on perceptions of a legitimate opposition. *American Political Science Review*, *101*(4), 621–635. https://doi.org/10.1017/S000305540707044X

Nanz, A., & Matthes, J. (2022). Democratic consequences of incidental exposure to political information: A meta-analysis. *Journal of Communication*, *72*(3), 345–373. https://doi.org/10.1093/joc/jqac008

Nechushtai, E. (2018). From liberal to polarized liberal? contemporary us news in Hallin and mancini's typology of news systems. *The International Journal of Press/politics*, *23*(2), 183–201. https://doi.org/10.1177/1940161218771902

Newman, N., Fletcher, R., Kalogeropoulos, A., & Kleis Nielsen, R. (2019). Reuters institute digital news report 2019. *Reuters Institute for the Study of Journalism*. https://www.digitalnewsreport.org/survey/2019/

Newman, N., Fletcher, R., Robertson, C. T., Eddy, K., & Kleis Nielsen, R. (2022). *Reuters Institute Digital News Report 2022*. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022

Nicholson, S. (2012). Polarizing cues. *American Journal of Political Science*, *56*(1), 52–66. https://doi.org/10.1111/j.1540-5907.2011.00541.x

Norris, P. (2000). *A virtuous circle: Political communications in postindustrial societies*. Cambridge University Press.

Ohme, J. (2020). Mobile but not mobilized? differential gains from mobile news consumption for citizens' political knowledge and campaign participation. *Digital Journalism*, *8*(1), 103–125. https://doi.org/10.1080/21670811.2019.1697625

Papilot.pl. (2021). Polacy spędzają w sieci 23 lata, 5 miesięcy i 16 dni swojego życia. https://www.papilot.pl/lifestyle/ciekawostki/57226/polacy-spedzaja-w-sieci-23-lata-5-miesiecy-i-16-dni-swojego-zycia

Park, C. S. (2019). Does too much news on social media discourage news seeking? mediating role of news efficacy between perceived news overload and news avoidance on social media. *Social Media + Society*, *5*(3), 2056305119872956. https://doi.org/10.1177/2056305119872956

Prior, M. (2003). Any good news in soft news? the impact of soft news preference on political knowledge. *Political Communication*, *20*(2), 149–171. https://doi.org/10.1080/10584600390211172

Prior, M. (2005). News vs. entertainment: How increasing media choice widens gaps in political knowledge and turnout. *American Journal of Political Science*, *49*(3), 577–592. https://doi.org/10.1111/j.1540-5907.2005.00143.x

Prior, M. (2007). *Post-broadcast democracy: How media choice increases inequality in political involvement and polarizes elections*. Cambridge University Press.

Prior, M. (2009). The immensely inflated news audience: Assessing bias in self-reported news exposure. *Public Opinion Quarterly*, *73*(1), 130–143. https://doi.org/10.1093/poq/nfp002

Prior, M. (2013). Media and political polarization. *Annual Review of Political Science*, *16*(1), 101–127. https://doi.org/10.1146/annurev-polisci-100711-135242

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon; schuster.

Schmitt, J. B., Debbelt, C. A., & Schneider, F. M. (2018). Too much information? predictors of information overload in the context of online news exposure. *Information, Communication & Society*, *21*(8), 1151–1167. https://doi.org/10.1080/1369118X.2017.1305427

Settle, J. E. (2018). *Frenemies: How social media polarizes america*. Cambridge University Press.

Skovsgaard, M., & Andersen, K. (2020). Conceptualizing news avoidance: Towards a shared understanding of different causes and potential solutions. *Journalism Studies*, *21*(4), 459–476. https://doi.org/10.1080/1461670X.2019.1686410

Somin, I. (2016). *Brexit and Political Ignorance*. https://www.washingtonpost.com/news/volokh-conspiracy/wp/2016/06/14/brexit-and-political-ignorance/

Song, H., Jung, J., & Kim, Y. (2017). Perceived news overload and its cognitive and attitudinal consequences for news usage in south korea. *Journalism & Mass Communication Quarterly*, *94*(4), 1172–1190. https://doi.org/10.1177/1077699016679975

Stier, S., Mangold, F., Scharkow, M., & Breuer, J. (2022). Post post-broadcast democracy? news exposure in the age of online intermediaries. *American Political Science Review*, *116*(2), 768–774. https://doi.org/10.1017/S0003055421001222

Strömbäck, J. (2005). In search of a standard: Four models of democracy and their normative implications for journalism. *Journalism Studies*, *6*(3), 331–345. https://doi.org/10.1080/14616700500131950

Stromback, J., & Shehata, A. (2010). Media malaise or a virtuous circle? exploring the causal relationships between news media exposure, political news attention and political interest. *European Journal of Political Research*, *49*(5), 575–597. https://doi.org/10.1111/j.1475-6765.2009.01913.x

Tewksbury, D., Weaver, A. J., & Maddex, B. D. (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. https://doi.org/10.1177/1464884920915352

Tocqueville, A. D. (1863). *Democracy in America [Google-books-ID: xZfibezcptec]*. Sever; Francis.

Toff, B., & Kalogeropoulos, A. (2020). All the news that's fit to ignore. *Public Opinion Quarterly*, *84* (S1), 366–390. https://doi.org/10.1093/poq/nfaa016

Valeriani, A., & Vaccari, C. (2016). Accidental exposure to politics on social media as online participation equalizer in Germany, Italy, and the United Kingdom. *New Media & Society*, *18*(9), 1857–1874. https://doi.org/10.1177/1461444815616223

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

Villi, M., Aharoni, T., Tenenboim-Weinblatt, K., Boczkowski, P. J., Hayashi, K., Mitchelstein, E., Tanaka, A., & Kligler-Vilenchik, N. (2022). Taking a break from news: A five-nation study of news avoidance in the digital era. *Digital Journalism*, *10*(1), 148–164. https://doi.org/10.1080/21670811.2021.1904266

Weeks, B. E., Lane, D. S., & Hahn, L. B. (2022). Online incidental exposure to news can minimize interest-based political knowledge gaps: Evidence from two u.S. elections. *The International Journal of Press/politics*, *27*(1), 243–262. https://doi.org/10.1177/1940161221991550

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*(1), 33–52. https://doi.org/10.1177/0894439315609528

Wojcieszak, M. E., de Leeuw, S., Menchen-Trevino, E., Lee, S., Huang-Isherwood, K. M., & Weeks, B. (2021). No polarization from partisan news: Over-time evidence from trace data. *The International Journal of Press/politics*, 19401612211047190. https://doi.org/10.1177/19401612211047194

Wojcieszak, M. E., Menchen-Trevino, E., Goncalves, J. F. F., & Weeks, B. E. (2021). Avenues to news and diverse news exposure online: Comparing direct navigation, social media, news aggregators, search queries, and article hyperlinks. *The International Journal of Press/politics*, *27*(4), 860–886. https://doi.org/10.1177/19401612211009160

Wojcieszak, M. E., & Mutz, D. C. (2009). Online groups and political discourse: Do online discussion spaces facilitate exposure to political disagreement? *Journal of Communication*, *59*(1), 40–56. https://doi.org/10.1111/j.1460-2466.2008.01403.x

Wojcieszak, M. E., von Hohenberg, B. C., Casas, A., Menchen-Trevino, E., de Leeuw, S., Gonçalves, A., & Boon, M. (2022). Null effects of news exposure: A test of the (un) desirable effects of a 'news vacation'and 'news binging'. *Humanities and Social Sciences Communications*, *9* (1), 1–10. https://doi.org/10.1057/s41599-022-01423-x

Yang, T., Majó-Vázquez, S., Nielsen, R. K., & González-Bailón, S. (2020). Exposure to news grows less fragmented with an increase in mobile access. *Proceedings of the National Academy of Sciences*, *117* (46), 28678–28683. https://doi.org/10.1073/pnas.2006089117

Zuckerberg. (2018). Zuckerberg 2018 post. Retrieved January 19, 2018, from https://www.facebook.com/zuck/posts/10104445245963251