**Who is Curating My Political Feed?**

**Characterizing Political Exposure of Registered U.S. Voters on Twitter**

Anonymous submission

**Abstract**

Social media platforms offer people a variety of options for engaging with politics, from following elected officials directly to discussing politics with social peers. Despite recent major advances in research into online political exposure through the lenses of selective exposure, filter bubbles, and ideological echo chambers, little is known about the fundamental questions of what types of political actors people are exposed to on social media and how these distinctive types vary across socio-demographic groups. We address this gap in the literature by analyzing unique panel data on more than 600,000 registered U.S. voters on Twitter during the 2020 U.S. Presidential campaign to identify distinct types of political consumers and how they vary in terms of socio-demographics. Our findings suggest that most of the population has a meaningful share of political content available from social peers, that the majority of this content originates from traditional sources of political information (media organizations, journalists, and politicians), and that media organizations are the dominant and direct source of political information on Twitter for nearly 20% of the sample population. These results advance our understanding of the ways in which citizens learn about politics in new media and pave the way for next-step research to identify the causal effects of exposure to distinct curators of political content on individuals’ political attitudes and political behavior.

Keywords: political exposure, social media, online media diets, Twitter, curated flows

# 

# Introduction

The tectonic shifts in the media environment and the rise of social media platforms over the past two decades have significantly changed the ways in which people are exposed to news and political information worldwide [(Fletcher and Nielsen, 2018)](https://www.zotero.org/google-docs/?HMtDVg). This trend has been particularly rapid in the United States, as Americans are now exposed to news more often through social media than through print; indeed, for younger generations social media has become the dominant channel for news [(Shearer, 2018)](https://www.zotero.org/google-docs/?CluboY).

In this increasingly networked media environment, the information populating one’s feeds is an amalgamation of curation decisions taken by others, including social peers, journalists, politicians, advertisers, and proprietary ranking algorithms [(Thorson and Wells, 2016)](https://www.zotero.org/google-docs/?2KCWwf). For example, Bakshy et al. [(2015)](https://www.zotero.org/google-docs/?Us6Fpv) showed how selective exposure on Facebook is partially determined by Facebook’s News Feed ranking algorithm, and determined more dominantly by the individual’s choice of whom to follow. Of course, the effects of social media and digital media use writ large extend beyond the online world, and a growing body of research shows mobilization effects, where digital media use is associated with more traditional forms of political participation offline [(Oser and Boulianne, 2020; Vaccari et al., 2015; Vaccari and Valeriani, 2021)](https://www.zotero.org/google-docs/?J6TdrH). Therefore, it is no surprise that issues of power and control [(Barzilai-Nahon, 2008)](https://www.zotero.org/google-docs/?xv53IS), limitations of free speech [(Morrow et al., 2021)](https://www.zotero.org/google-docs/?qk4n1h), and individual choice [(Bakshy et al., 2015)](https://www.zotero.org/google-docs/?9GSzQV) in political exposure on social media are among the most contested topics of our time.

Despite the clear importance of advancing scholarly and real-world knowledge about political exposure on social media, we know relatively little about two key parameters of political exposure, namely the prevalence of different types of actors in people’s social feeds, and how political exposure varies across socio-demographic groups. Currently, no social media platform provides precise individual-level or comprehensive aggregate-level information about exposure, which poses a key obstacle to advancing research on these crucial topics. The Social Science One initiative [(King and Persily, 2020)](https://www.zotero.org/google-docs/?GcD1GK) does provide aggregate information about viewership, but this information is limited in several important respects: it is currently limited to Facebook data; it includes only URLs and not all political content; it does not distinguish eligible from non-eligible voters; and it does not provide information about the user that posted the content. In lieu of more precise measurements, researchers have made recent contributions on the abovementioned topics by relying on self-reported measures of political consumption and general-purpose web tracking data, while acknowledging the serious selection bias challenges inherent in this approach [(Guess, 2021; Wojcieszak et al., 2022b)](https://www.zotero.org/google-docs/?8jAs9A).

In this study, we build on the extant literature on political exposure by using a research design that leverages a large panel of U.S. registered voters and their activity on Twitter. The combination of these two data sources creates the opportunity to ask and answer research questions that correspond to the two key parameters of political exposure noted above: What are the types of political exposure on social media in terms of quantity and composition from different types of actors (RQ1)? How do these types of political exposure vary across socio-demographic groups (RQ2)? To answer these research questions, we build on the curated flows theoretical framework [developed by Thorson and Wells (2016)](https://www.zotero.org/google-docs/?Ky9nUg) to identify the political content available to registered U.S. voters on Twitter and curated by different actors, including media organizations, journalists, politicians, opinion leaders, and social peers. For this, we use clustering methods that identify the prototypical modes of political exposure in terms of the breakdown of the actors responsible for this content exposure, and identify the socio-demographic covariates of each distinctive cluster.

Our contributions using this novel methodological approach are twofold. First, referring to a large and representative sample of registered U.S. voters on Twitter, we provide new empirical evidence about the prototypical modes of political exposure, in terms of both the quantity of political content and the composition of different actors who curate it. Second, we present findings on the varying levels and compositions of political exposure by different socio-demographic groups of registered U.S. voters on Twitter. Taken together, our contributions begin to address some of the most basic, yet unanswered, questions at the heart of the curated flow framework and social media communications: Who are the most significant curators in political communication, and for whom?

# The Importance of Political Exposure Online and on Social Media

Numerous studies show that online political exposure and information consumption on social media are related to political attitudes and behaviors, both online and offline. For example, Valeriani and Vaccari (2016) found that accidental exposure to information on social media is positively associated with online political participation in multiple national contexts. A recent meta-analysis concluded that incidental exposure, an unintended form of exposure that is common on social media, is positively associated with a variety of pro-democratic attitudes and behaviors, including news use, political knowledge, political participation, expressive engagement, and political discussion [(Nanz and Matthes, 2022)](https://www.zotero.org/google-docs/?Xw5WQL). Weeks et al. (2017) further found that counter-attitudinal incidental exposure on social media drove processes of selective exposure among stronger partisans, which subsequently led to greater political information-sharing. In contrast, overreliance on the news found on social networks is negatively associated with important socio-political indicators of political knowledge, political interest, and voting (Gil de Zúñiga et al., 2019).

Several recent studies informed by Thorson and Wells’ curated flows framework have shown that the impact of political messaging also depends on the type of actor delivering it, as the same political message received from different types of sources may have divers impacts on attitudes and behavior. For example, recent research has indicated that statements by celebrities and online influencers seem to affect the public’s real-world beliefs compared to similar statements by non-celebrities [(Alatas et al., 2019; Alrababa’h et al., 2021; Suuronen et al., 2021)](https://www.zotero.org/google-docs/?YXhlxR). Regarding media sources, research has shown that high levels of exposure to media outlets with high levels of political content shape political knowledge and behavior, including the propensity to vote [(de Vreese and Boomgaarden, 2006)](https://www.zotero.org/google-docs/?CHYeJh). Research by Graham et al. [(2015)](https://www.zotero.org/google-docs/?oMoBRy) into peer networks showed that over half of the political discussions in online forums in the United Kingdom led to at least one political action. The importance of the clear identification of actors is evident in Taylor et al.’s [(2022)](https://www.zotero.org/google-docs/?eOPQuL) large-scale longitudinal field experiment, which showed that content provided by anonymous sources has less impact on viewers’ opinions and behaviors compared to content shared by identified individuals with known reputations. Taken together, this emerging research indicates that the messenger’s identity may be as important as the message itself.

# Who is Heard and by Whom in Political Communication

A central element in democratic theory is allowing the expression of ideas to facilitate public information-sharing and deliberation [(Habermas, 1984)](https://www.zotero.org/google-docs/?xC5jjq). While research on the ways citizens construct their information diets precedes the digital era [(Katz and Blumler, 1974; Sears and Freedman, 1967)](https://www.zotero.org/google-docs/?GN5bql), the shift to online media, accompanied by the weakening of traditional gatekeepers and the context collapse that are common on social media [(Davis and Jurgenson, 2014)](https://www.zotero.org/google-docs/?2AenPu), warrants renewed attention to the fundamental question of who is being heard in modern political communication. Addressing this question is important for advancing our understanding of the extent to which social media, and information systems more broadly, fulfill their egalitarian potential [(Allen, 2015)](https://www.zotero.org/google-docs/?XcTn0A) or reinforce old political structures and power as the weapon of the strong [(Hindman, 2009)](https://www.zotero.org/google-docs/?6uuyS1).

As noted, the theoretical and empirical importance of examining who is being heard is highlighted by Thorson and Wells’ [(2016)](https://www.zotero.org/google-docs/?xOjqAM) discussion of the role of individual-level “curation” for understanding media exposure and its effects. While individuals choose whom to follow on social media, the notion of curation emphasizes the agency of external actors over the composition of an individual’s social media feed. In particular, the curated flows framework identifies a number of key actors, including social peers, journalists, politicians, and advertisers as well as proprietary ranking algorithms. Merten (2021) explored the decisions (e.g., follow, block, or hide) that users report taking in response to news curation by others. However, there is little empirical work showing the relative prevalence of different actors in the public’s political exposure. Two notable exceptions are the recent work by Wojcieszak et al. [(2022b)](https://www.zotero.org/google-docs/?uGXImo), which sheds new light on the channels (search engines, social media, aggregators, etc.) that lead people to news, and that by Jürgens and Birgit (2022), which measured the diversity of news accessed through different channels. Nevertheless, in order to advance our understanding of the media effects of social media and to gain better insights into the ways political learning takes place on such social platforms (Bode, 2016), we need to heed Prior’s (2009) call for better measurement of news exposure. Currently, little is known about the amount of political content to which people are exposed on social media, and the different kinds of actors involved in conveying this information. Therefore, our first research question is:

**(RQ1) What are the prototypical types of political exposure on social media both in terms of overall quantity and composition from different types of actors?**

A key component of political exposure is political ideology and the range of ideas being represented. Some recent work indicates that exposure to political content through online social networks may serve to increase political polarization [(Bail et al., 2018; Garrett et al., 2014; Shmargad and Klar, 2020)](https://www.zotero.org/google-docs/?RPQmr1). However, other studies indicate that social media exposure through weak ties and the visibility of social endorsements reduces polarization by offering diversity of exposure [(Barberá, 2015; Messing and Westwood, 2014)](https://www.zotero.org/google-docs/?hsuFLC). While people more frequently follow those media and politicians’ accounts that align with their ideology [(Eady et al., 2019; Wojcieszak et al., 2022a)](https://www.zotero.org/google-docs/?YgbcnN), there is still substantial overlap among people’s news diets [(Guess, 2021)](https://www.zotero.org/google-docs/?XFPHxq). Given that there is no consensus about the polarizing effects of media or social media [(Prior, 2013; Zhuravskaya et al., 2020)](https://www.zotero.org/google-docs/?m7rhi9), it is especially important to refine our understanding of political exposure on social media and to consider it jointly with political ideology.

Socio-demographic characteristics are also linked to political consumption. Consistent with Vaccari and Valeriani’s [(2021)](https://www.zotero.org/google-docs/?srZ9VS) call to move beyond the “one-effect-fits-all” fallacy, we draw on prior literature to assess how key socio-demographic characteristics relate to distinctive types of political exposure. First, there is a well-documented age gradient observed in individuals’ levels of interest in politics and their self-efficacy [(Verba et al., 1995)](https://www.zotero.org/google-docs/?KY9X8d). As younger generations increasingly obtain their news from social media [(Shearer, 2018)](https://www.zotero.org/google-docs/?MIQ5v6), it is important to study the types of political content to which they are exposed. In addition, research shows that those with traditionally advantaged socio-demographic backgrounds (e.g., male, older) are more active politically, including making efforts to seek out political content [(Schlozman et al., 2018)](https://www.zotero.org/google-docs/?mGwZf1). Yet, research suggests that social media and online participation may have differential mobilization effects that recruit younger groups and women more actively into politics [(Oser et al., 2013; Oser and Boulianne, 2020)](https://www.zotero.org/google-docs/?g61F5N). A possible explanation is that publics and counterpublics pay attention to different issues on social media [(Jackson and Foucault Welles, 2015; Shugars et al., 2021)](https://www.zotero.org/google-docs/?jDCaoj). Therefore, our second research question is:

**(RQ2) How do the prototypical types of political exposure on social media vary for distinctive socio-demographic groups?**

We now turn to the methodological challenges and opportunities involved in making robust inferences about the political exposure of citizens on social media.

# Measuring Political Exposure in the Digital Age: Challenges and Opportunities

Although survey data have long been a leading source of information about political consumption habits, researchers are actively seeking ways to improve data’s accuracy [(Berinsky, 2017; Guess, 2015)](https://www.zotero.org/google-docs/?qGau64). In the context of social media, prior work has shown that there could be large discrepancies between actual and reported frequency of posting about politics [(Guess et al., 2019; Henderson et al., 2021)](https://www.zotero.org/google-docs/?L7fgtT).

Digital trace data provide new and complementary ways to measure individuals’ behavior directly and are often collected through dedicated software installed by participants [(Flaxman et al., 2016)](https://www.zotero.org/google-docs/?rfgEwk). For example, Guess (2021) uses web browsing data combined with survey responses to characterize Americans’ media consumption habits and examine whether internet use facilitates selective exposure to like-minded views. While this approach provides the most comprehensive picture of both objective and subjective measures of political engagement, the research is often limited to a few thousand participants who are willing to volunteer their data. In addition to raising selection issues, the sample quickly becomes statistically underpowered for obtaining accurate descriptions of subgroups and heterogeneity of activity [(Hughes et al., 2021)](https://www.zotero.org/google-docs/?yoMKGY). The challenge of directly measuring political exposure for the field as a whole is clearly identified by Amsalem and Zoizner’s (2023) comprehensive meta-analysis of learning about politics on social media. They observed that most relevant studies do not include any direct measure of political exposure and lack sufficient sample size to estimate heterogeneous effects.

A recently-developed alternative approach for directly gathering data on individuals’ behavior is to use publicly available social media data. Despite the meaningful changes to Twitter’s ownership and policies since 2022, it has been a uniquely important social media platform for investigating exposure to political content of a large sample of users due to the active engagement of media outlets and political figures on the platform, including during the observation period of the current study [(Bail et al., 2018; Barberá, 2015; Eady et al., 2019; Guess, 2021)](https://www.zotero.org/google-docs/?YJ1Xln). In 2021, around one in five (23%) of Americans reported using Twitter [(Odabaş, 2022)](https://www.zotero.org/google-docs/?5Zyg2N), and of these, almost seven in ten said they receive their news regularly through the platform [(Mitchell et al., 2021)](https://www.zotero.org/google-docs/?T8mVMP). While Twitter users in the United States were found to be younger and more likely to be Democrats in comparison to the general public [(Wojcik and Hughes, 2019)](https://www.zotero.org/google-docs/?us2UEu), prior work has shown that differences between Twitter users and non-users are due mostly to the demographic composition of social media users, which can be addressed by controlling for demographic variables [(Mellon and Prosser, 2017)](https://www.zotero.org/google-docs/?Tu2FFc). Importantly, rigorous empirical work on the representativeness of Twitter users has shown some modest demographic differences between Twitter users and the general population that can be accounted for analytically [(Hughes et al., 2021)](https://www.zotero.org/google-docs/?slG7Vx).

The recently-developed methodological approach of analyzing publicly available social media data is an important contribution to the existing literature, as no social media platform currently offers public access to data about individuals’ exposure to distinctive types of political content, and as a result, scarce research has directly measured it. An increasingly prominent approach for approximating exposure involves the collection of content posted on the accounts followed by the focal user on social media [(Eady et al., 2019; Grinberg et al., 2019)](https://www.zotero.org/google-docs/?j2KhVU). As described in Grinberg et al. (2019), this approach does not guarantee exposure, i.e., that an individual actually saw a particular post, but it does directly speak to the content available to people in their social feeds from their ego network.

The following section details how the current study applies this recently-developed novel methodological approach to measuring potential political exposure in terms of the content available from social peers.

# Data and Methods

## *Twitter Panel and Political Exposure*

This research is based on a sampling frame of over 1.5 million Twitter users who were successfully matched to public U.S. voter registration records. Following the approach described in prior work (e.g., see Grinberg et al., 2019 and Shugars et al., 2021), the matching process used the Twitter Decahose, a 10% random sample of all tweets, to identify 290 million profiles that posted content between January 2014 and March 2017. The profiles were then matched against voter records provided by TargetSmart in October 2017 for all 50 U.S. states and the District of Columbia. A Twitter account was matched to a voter record if its user’s full name exactly matched and it represented the only person with that name in either the city- or state-level geographic area specified in both datasets. While the reliance on full names and disclosed locations eliminated many fake, automated (bot), and organizational accounts, it did raise concerns about potential selection bias. However, rigorous comparison of this panel with a gold-standard survey conducted by the Pew Research Center showed that only small demographic and ideological differences exist between the two samples of registered U.S. voters [(Hughes et al., 2021)](https://www.zotero.org/google-docs/?li1xsm). Importantly, this matched dataset provided comprehensive data on individuals’ social media behavior on Twitter as well as basic socio-demographic information. Age and gender were taken directly from public voter registration records, while race/ethnicity and party affiliation are based on TargetSmart inferences (see validation in Appendix B of Shugars et al. 2021).

The primary dataset used in this work was a set of 606,112 panel members for whom we had at least one indication of activity on Twitter during the 2020 Presidential election (August through November 2020). This set included all individuals who posted or liked at least one tweet during the four months of the study period. By focusing on a period of a presidential election, we examined potential political exposure at its peak [(Grinberg et al., 2019; Peterson et al., 2021)](https://www.zotero.org/google-docs/?xZpWIk). See Appendix A for socio-demographic information about these active panel members. As our target population is restricted to registered U.S. voters on Twitter, we make no claims about the important, yet omitted, populations of eligible non-registered voters or inactive Twitter users.

To model potential political exposure, we followed the approach used in prior work that approximates individuals’ exposure using the content available from the accounts they follow [(Eady et al., 2019; Grinberg et al., 2019)](https://www.zotero.org/google-docs/?BpRW7Y). This modeling approach acknowledges that appearing in the feed is a necessary condition for exposure, without claiming that it is a sufficient condition. Our panel followed a total of 51 million unique Twitter users, which made it impossible to collect all of their tweets during the observation period through the Twitter API. We addressed this limitation by analyzing the content in the Twitter Decahose posted by accounts followed by the panel, similar to the approach use by Grinberg et al. (2019). An important limitation of this technique is that it provides only a sample of the content from an individual’s network. Specifically, this technique does not include ads and does not consider algorithmic ranking. Yet, in lieu of more precise information from social media platforms about exposure, this approach reflects the most accurate and reproducible estimate currently available to the public about the composition of people’s social feeds.

To identify political tweets, we trained a Machine Learning classifier and validated its accuracy against human coders, similar to the approach used in prior work [(Bakshy et al., 2015; Eady et al., 2019; Grinberg et al., 2019)](https://www.zotero.org/google-docs/?zjVOdV). The classifier resulted in a precision of 88.8% and a recall of 80.0% for tweets about U.S. politics, and a recall of 96.4% when further restricted to the subcategory of election-related tweets. More details about the classifier and its validation are in Appendix B.

## *Identifying Different Actors in Political Exposure*

Following the curated flows framework, we examined different types of actors that curate political content for individuals in our panel. We focused on four types of actors directly mentioned in Thorson and Wells’ (2016) framework – media organizations, journalists, politicians, and social peers – and include a fifth category identified as important in recent research: opinion leaders who have been identified as important in recent research. Specifically, opinion leaders have large followerships on social media, nonpolitical even more than political opinion leaders [(Mukerjee et al., 2022)](https://www.zotero.org/google-docs/?PTj0Rs), and a demonstrated ability to influence public opinion [(Alatas et al., 2019)](https://www.zotero.org/google-docs/?np0tlN). To date, however, the share of political content originating from opinion leaders’ accounts in day-to-day political exposure has not been directly quantified.

To identify accounts of different actors, we relied on manually curated lists of accounts by recent academic works, developed methods to identify additional accounts, and validated the accuracy of our inferences and the robustness of results. We identified media organizations by using the list of media organizations in McCabe et al. [(2022)](https://www.zotero.org/google-docs/?PZDv03), which started with a seed list of known media organizations and used snowball sampling to expand it iteratively. We supplemented this list with the media organizations listed in Wojcieszak et al. [(2022a)](https://www.zotero.org/google-docs/?BEqCQw). We also relied on Wojcieszak et al.’s [(2022a)](https://www.zotero.org/google-docs/?O5vQaX) extensive list to identify 1,951 journalists’ accounts.

Politician accounts were identified through an original list compiled by linking an official list of the 116th Members of Congress (MoC) names to a list of MoC accounts on Twitter [(Wrubel and Kerchner, 2020)](https://www.zotero.org/google-docs/?cRNucn). Our list of 927 accounts includes both the accounts of MoC and their election campaigning accounts, which is important for capturing all messages originating from politicians during an election cycle. We supplemented this list with 51 additional politicians’ accounts found in Wojcieszak et al. [(2022a)](https://www.zotero.org/google-docs/?uLWxGQ). Appendix C details our identification strategy for politicians’ accounts.

To identify opinion leaders, we relied on the manually labeled list of accounts of nonpolitical opinion leaders (e.g. public figures, popular brands, celebrities) by Mukerjee et al. (2022), and extend it using Bail et al.’s [(2018)](https://www.zotero.org/google-docs/?yByALo) approach of considering as an opinion leader any account followed by 15 or more active MoC. Since the accounts followed by multiple MoC may themselves belong to media organizations, journalists, or politicians, we used a combination of automatic and manual annotation of accounts using profile information to distinguish opinion leaders from other actor types. Validating this approach using a hold-out random sample of accounts resulted in 80.0% accuracy, which is considerably higher than when using random assignment with four categories. Appendix D details our identification strategy for opinion leader accounts.

We considered as a social peer any account that did not appear on any of the abovementioned lists of media organizations, journalists, politicians, or opinion leaders. Importantly, however, the same content can be attributed to multiple users due to the complex nature of retweets, quotes, replies, or mentions. To support different attributions of content and interpretations of the results, we distinguished between *direct* and *indirect* exposure. Direct exposure results from directly following the accounts of media organizations, journalists, politicians, or opinion leaders. Indirect exposure is mediated through social peers who retweet, quote, mention, or reply to a tweet by these actors. To complete our documentation of the analytical work we conducted to identify distinctive actors in political exposure, Appendix E validates account inferences and robustness, and Appendix F provides summary information about all curating actors analyzed in this study.

## 

## *Measuring Political Alignment of Curating Actors*

Modeling the ideological leaning of curating actors is fundamental to assessing people’s online media diets, and different methods have been proposed for this purpose. Our analysis focused on three aspects of citizens’ political exposure: (i) exposure to left- and right-leaning MoC; (ii) exposure to left- and right-leaning opinion leaders; and (iii) exposure to political news sites based on the ideological alignment of people who share links to this site.

For MoC, we considered their party affiliation to be representative of their political leaning, excluding four Independents and one Libertarian. For news sites, we followed Bakshy et al.’s [(2015)](https://www.zotero.org/google-docs/?CTTnD4), approach by estimating political leaning based on the political alignment of people sharing links to the website. Appendix G provides further details on our estimation of the political alignment of news sites. For opinion leaders, we inferred their political leanings based on the composition of MoC who followed them.

.

## *Clustering Methodology to Infer Prototypical Types of Political Exposure*

We used state-of-the-art clustering methods to identify prototypical types of political exposure on Twitter. To this end, we took into account several categories of information about panel members’ political exposure: (i) the overall magnitude of political exposure and its share of the user’s available feed; (ii) the curating sources (partitioned by direct and indirect exposure); and (iii) the ideological leaning of the political actors and news sites in the feed. Appendix H provides the full list and description of the 15 features that we analyzed to provide measures of these three key categories of political exposure.

For clustering, we used the common approach with this type of large and complex dataset to reduce the dimensionality of the data first [(Allaoui et al., 2020; Grootendorst, 2022](https://www.zotero.org/google-docs/?BCJz7O)[)](https://www.zotero.org/google-docs/?3FIPoL), and only then applied the clustering algorithm. Specifically, we used Uniform Manifold Approximation and Projection (UMAP) to reduce dimensionality [(McInnes et al., 2020)](https://www.zotero.org/google-docs/?3FIPoL), and then applied the clustering algorithm of HDBSCAN, which determined the optimal number of clusters, subject to minimum cluster size, and is robust to outliers [(McInnes et al., 2017)](https://www.zotero.org/google-docs/?OGCwz0).

# Results

In this section, we report results regarding our two research questions: what are the prototypical types of political exposure on Twitter?; and how does the distribution of these exposure types vary across distinctive socio-demographic groups?

To identify robust patterns of political exposure, users who did not meet a minimum threshold of political exposure were assigned to a separate cluster of “nonpolitical” users. Consistent with prior research, we set this threshold at one observed political tweet a day on average in the Decahose, i.e., a total of 122 observations throughout the entire election period [(Grinberg et al., 2019)](https://www.zotero.org/google-docs/?P8V5Yh). Based on this criterion, 8.9% of the population was directly assigned to the nonpolitical cluster.

Clustering the political exposure of users with a minimal amount of exposure to politics resulted in seven clusters that covered 99.1% of the population, 0.4% of accounts that the algorithm identified as outliers, and three small clusters with several hundred people that together amounted to 0.5% of the population. We then omitted these outlier and small-cluster accounts from further analysis and focused on the core exposure patterns identified in 99.1% of the population.

Figure 1 presents the prototypical types identified by clustering the political exposure of panel members. Each point in Figure 1A represents an individual and their political consumption at the reduced two-dimensional space computed by the UMAP algorithm with its color designating

|  |
| --- |
|  |
| ***Figure 1: Prototypical types of individual political exposure.***  *Each point in panel (A) represents the political exposure of a single panel member, reduced to two dimensions using the UMAP algorithm, and colored by the cluster assignment obtained from HDBSCAN. Panel (B) shows the median number of political tweets available to individuals per day (left, darker-colored bars), and their percentage out of all tweets available to them on Twitter (right, lighter-colored bars). Cluster labels and their share in the population are specified on the x-axis. Colors are consistent between the two figure panels. Ninety-five percent bootstrapped CIs are omitted from the figure due to their small magnitude, which is upper bounded by 27 exposures to tweets and 0.28% in the share of politics.* |

its cluster assignment. Points that are closer together represent individuals with similar properties of political exposure. Figure 1B shows the median amount of political exposure available in people’s ego networks[[1]](#footnote-1) and its share out of all content available to people on Twitter, for each cluster separately. For example, the cluster referred to as Media Superconsumers consists of 4.7% of the population, and this cluster’s median user had nearly 6,000 political tweets available to them each day, comprising 52% of their total daily tweet exposure.

Cluster labels reflect our assessment of the distinctive features of each cluster in terms of its size in the population, the amount of political exposure it represents, and the composition of political curators and sources in the cluster. In particular, we labeled the clusters as follows: one cluster as Nonpolitical due to low level of exposure to politics (not meeting our minimal threshold); one cluster as Opinion Leaders (OL)-Oriented due to an elevated level of exposure to opinion leaders; one cluster as Average Consumers based on its large share in the population (50.5%); two clusters as Partisan (Partisan Left and Partisan Right) due to the political alignment of content; and three clusters as Media-Oriented, Media-Oriented++, and Superconsumers, based on their increasingly elevated levels of media consumption.

Figure 1 provides two key observations. First, most of the population has a meaningful share of politics in their Twitter feeds. This finding is consistent with prior work showing that Twitter users are above average in their political engagement [(McClain et al., 2021)](https://www.zotero.org/google-docs/?H6izt0). With the exception of the two clusters with the lowest share of political involvement (Nonpolitical and OL-Oriented), all other clusters, which account for nearly 90% of the population, have 8% or more of politics in their feeds. Even if this finding is applicable only to registered U.S. voters on Twitter and, to a lesser extent, to voters on other social media platforms, it does present a picture of an engaged public during an election cycle. Second, we observe that the Partisan Left and Partisan Right clusters exhibit very similar levels of political consumption, and that the media-oriented clusters, which, together, amount to 15–24% of the population, have a larger share of politics in the content available from social peers than do the Partisan clusters. Next-step research can use these findings to investigate the causal relationship between the overall level of direct and indirect media exposure and subsequent attitudes and behaviors such as ideological polarization.

In addition to the overall level of exposure to politics, the clusters we identified vary in the composition of political exposure from distinctive curating actors. Figure 2A shows the breakdown of political exposure by different actor types, including media organizations, journalists, politicians, opinion leaders, and social peers. Lighter-colored bars indicate indirect exposure, where the focal user received content from a peer that referred to a media organization, journalist, politician, or opinion leader (i.e., through a retweet).

|  |
| --- |
|  |
| ***Figure 2A: The composition of political exposure across clusters.***  *The share of politics curated by different actor types (y-axis) across clusters (x-axis). Darker-colored bars represent direct exposure to media organizations, journalists, politicians, opinion leaders, and social peers. Lighter-colored bars represent indirect exposure to media organizations, journalists, politicians, or opinion leaders through social peers.* |

by either retweeting, quoting, replying, or mentioning them. For example, the group of Average Consumers receives over four times more indirect exposure (22.8%) than direct exposure (5.2%) to politicians. In stark contrast, the Media Superconsumers group receives nearly 90% of their political exposure directly from media organizations with very little indirect exposure. A similar pattern appears for the OL-Oriented cluster, which receives more than 70% of its political exposure directly from opinion leaders.

Focusing again on the six clusters representing the bulk of the population (Average, Partisan Left, Partisan Right, and Media Consumers; nearly 90%), Figure 2 shows that more than half of political exposure for these clusters comes from traditional sources of political information: media organizations, journalists, and politicians. The overall share of these traditional sources in political exposure increases with the increased share of politics in the feed (from left to right, as shown in Fig. 1), with the notable exception of the small OL-Oriented cluster. The clusters also vary considerably in terms of direct and indirect exposure. Nonpolitical Consumers are only indirectly exposed to traditional sources, while Average Consumers receive some direct exposure to media organizations and politicians and considerably more indirect exposure to traditional sources through their peer networks. Partisans have the largest share of political exposure directly from politicians and journalists and relatively little direct exposure to media organizations compared to the Media-Oriented clusters. Leaving aside the more extreme Superconsumers, we see that the Media-Oriented and Media-Oriented++ clusters, which total nearly 20% of the sample population, receive about half of their political exposure directly from media organizations. Taken together, these findings highlight the importance of considering both direct and indirect exposure to traditional sources as well as to opinion leaders, particularly for people who receive a smaller share of political content.

We now turn to our second research question, which focuses on how different socio-demographic groups engage with different types of political consumption. Figure 3 shows how age, gender, ethnicity, and party affiliation are distributed across the different exposure types. Specifically, the figure shows the average age estimate for each cluster and the percentage of women, Caucasians, and registered Democrats in each cluster (See Appendix A for further details on socio-demographic characteristics). The dashed horizontal line in each panel designates the sample average as a baseline for comparison.

|  |
| --- |
|  |
| ***Figure 3: Socio-demographic characteristics among different political exposure types.***  *Note: Sample averages are marked in a gray dashed line. Ninety-five percent bootstrapped CIs are shown (mostly occluded due to their small size).* |

Examining the figure’s presentation of the clusters’ socio-demographic characteristics using the same x-axis ordering of the clusters as in prior figures, i.e., by increasing share of political exposure from left to right, reveals several key observations. First, there is a clear positive association between age and share of political exposure, as the literature would predict (e.g., Verba, Schlozman and Brady, 1995). In theory, the ego network of young adults, who generally participate less in politics, could have compensated for their lack of direct exposure. In practice, we observe that the youngest age cohort in our data (18-29) is overrepresented in the nonpolitical group by 19.2% relative to their proportion in the sample. This suggests that the little exposure to politics of young adults may, in part, be a network effect. Second, the figure shows meaningful gender and ethnicity differences between Partisan Right and Partisan Left. The fact that the two Partisan clusters have different demographic characteristics, yet a similar breakdown of actors in their political feeds, suggests that there may be some commonalities in the polarization processes across political ideology. Third, the OL-Oriented cluster is distinctively young, male, and non-Caucasian. Together with its small size in the sample (1.7%) and overrepresentation of opinion leaders, this seems like a niche cluster that gets exposed to politics incidentally through nonpolitical opinion leaders.

Finally, we find that the Media Oriented and Media Oriented++ clusters, which together comprise nearly 20% of the sample population, have significantly higher percentages of women, registered Democrats, and older adults. Prior work has documented a partisan gender gap in American politics [(Doherty et al., 2018)](https://www.zotero.org/google-docs/?7aK6Kn), with women more likely to identify as Democrats. However, to the best of our knowledge, no prior work has shown such large gender differences in political consumption directly from media organizations.

# 

# Discussion

Much of the discussion about societal factors that may be contributing to democratic backsliding in advanced democracies – including rising populism, decreasing trust in media and the political establishment, increased polarization, and misinformation – has been linked to the increased prevalence of digital media, in particular social media. Social platforms are, indeed, widely adopted as a source of political information and are a primary source for many young adults. These trends in political content exposure call for a better theoretical understanding of political exposure on these platforms, including next-step causal examination of the impact of different types of political exposure on subsequent political attitudes and political behaviors. Robust analysis of these phenomena requires new computational methods for making valid inferences based on digital trace data that complement traditional methods.

Grounded in the curated flows theoretical framework, this work contributes to the conceptualization and measurement of actors responsible for this curation. The empirical findings describe the types of actors responsible for political content distribution to registered U.S. voters on Twitter, and the demographic characteristics of distinctive types of political consumers. We found that most of the population on the platform was exposed to non-negligible amounts of political content during the 2020 U.S. Presidential election, ranging on an average day from 87 political tweets (8% of the overall feed) to a few thousand political tweets (52% of the overall feed). Notably, more than half of political tweets originated from traditional sources of political information: media organizations, journalists, and politicians. The observational findings of the current study pave the way to investigate the causal impact of political content curation by distinctive actors on individuals’ subsequent attitudes and behaviors, such as left-right ideological polarization and affective polarization.

Another key finding is that media organizations were an important source of political information for a large proportion of the sample, with much of this exposure taking place directly and without any mediation by peers. These findings contribute to the debate about the erosion of traditional gatekeepers, as most media organizations on our lists have fundamentally the same editorial processes that Kurt Lewin [(1943)](https://www.zotero.org/google-docs/?DO3jGE) wrote about when he first introduced Gatekeeping Theory. Our results show that a substantial proportion of consumers of political content on Twitter *choose* to replicate traditional gatekeeping in new media. Future research could investigate the curation roles and impacts of these media-oriented individuals in their local networks and examine the role of media organizations in influencing subsequent political attitudes and behaviors of specific socio-demographic groups.

Along with these contributions, this research has several important limitations mentioned earlier. First, while the findings are likely to capture the political exposure of American adults on Twitter in 2020, which represented about a fifth of American adults [(Odabaş, 2022)](https://www.zotero.org/google-docs/?gINu4T), it much less clear how they will generalize to other populations and social media platforms without direct measurement. Previous research has found some similar media effects to Twitter and the more widely-used Facebook (e.g., Valenzuela et al. 2018). However, numerous studies have emphasized the importance of considering specific contextual features in the relationship between social media use and political behavior (e.g., Vaccari and Valeriani, 2022). Additional comparative research is needed to fully contextualize these findings. A second key limitation is the empirical focus of our analysis on potential political exposure, meaning content that is available to people but not necessarily seen by them. Although this is a limitation that affects all scholarship on these topics, it is important to note that the differences between these two populations may be systematically affected by factors such as the time of day when individuals visit their feeds, the duration of their visits, and the algorithmic content ranking conducted by social platforms. A third key limitation is that, having relied on manually curated lists and verification for identifying distinctive curation actors (e.g., media organizations, opinion leaders), we cannot guarantee the comprehensiveness of the lists. For example, the list of politicians does not include state and local politicians, who may have different levels of exposure and audiences.

There are also several avenues for future work to expand this research. In terms of theory, the curated flows framework puts much of its emphasis on the actor doing the curation. Our study shows that there is room to expand the theory to consider the producer of the content in addition to the curator while the content is propagating through the network. Content attribution is also a major challenge that warrants methodological contributions. Furthermore, future research can examine how the different types of political content exposure are related to pro-democratic attitudinal measures known to be crucial for robust democratic functioning, such as political knowledge and political efficacy. In addition, the current study paves the way for next-step experimental research that makes clear causal identification of how different types of political consumers engage in, and mobilize for, political action both online and offline.

# Acknowledgments

The research was approved by [anonymized] Institutional Review Board.

**Funding:** [anonymized for submission]

**Data and materials availability:** All data and code necessary to evaluate the final conclusions in this paper are available in a public GitHub repository [anonymized for submission] and in Harvard Dataverse [anonymized for submission]. Although the Twitter panel was originally constructed by linking two public resources (public voter records and public Twitter activity), we do not release any personally identifying information, as doing so could expose individuals to additional unanticipated risks, violate their privacy expectations [(Williams et al., 2017)](https://www.zotero.org/google-docs/?OaiQRq), and breach Twitter’s Terms of Service.

# 

# References

[Aggarwal, C., Hinneburg, A., Keim, D., 2002. On the Surprising Behavior of Distance Metric in High-Dimensional Space. First Publ Database Theory ICDT 200 8th Int. Conf. Lond. UK January 4–6 2001 Jan Van Den Bussche Eds Berl. Springer 2001, pp 420-434 Lect. Notes Comput. Sci. 1973.](https://www.zotero.org/google-docs/?FHjHwm)

[Alatas, V., Chandrasekhar, A.G., Mobius, M., Olken, B.A., Paladines, C., 2019. When celebrities speak: A nationwide twitter experiment promoting vaccination in Indonesia. National Bureau of Economic Research.](https://www.zotero.org/google-docs/?FHjHwm)

[Allaoui, M., Kherfi, M.L., Cheriet, A., 2020. Considerably Improving Clustering Algorithms Using UMAP Dimensionality Reduction Technique: A Comparative Study. pp. 317–325.](https://www.zotero.org/google-docs/?FHjHwm)

[Allen, D., 2015. Reconceiving public spheres, in: From Voice to Influence: Understanding Citizenship in a Digital Age. University of Chicago Press, p. 179.](https://www.zotero.org/google-docs/?FHjHwm)

[Alrababa’h, A., Marble, W., Mousa, S., Siegel, A.A., 2021. Can Exposure to Celebrities Reduce Prejudice? The Effect of Mohamed Salah on Islamophobic Behaviors and Attitudes. Am. Polit. Sci. Rev. 115, 1111–1128. https://doi.org/10.1017/S0003055421000423](https://www.zotero.org/google-docs/?FHjHwm)

[Bail, C.A., Argyle, L.P., Brown, T.W., Bumpus, J.P., Chen, H., Hunzaker, M.B.F., Lee, J., Mann, M., Merhout, F., Volfovsky, A., 2018. Exposure to opposing views on social media can increase political polarization. Proc. Natl. Acad. Sci. 115, 9216–9221. https://doi.org/10.1073/pnas.1804840115](https://www.zotero.org/google-docs/?FHjHwm)

[Bakshy, E., Messing, S., Adamic, L.A., 2015. Exposure to ideologically diverse news and opinion on Facebook. Science 348, 1130–1132. https://doi.org/10.1126/science.aaa1160](https://www.zotero.org/google-docs/?FHjHwm)

[Barberá, P., 2015. How Social Media Reduces Mass Political Polarization. Evidence from Germany, Spain, and the U.S. 46.](https://www.zotero.org/google-docs/?FHjHwm)

[Barzilai-Nahon, K., 2008. Toward a theory of network gatekeeping: A framework for exploring information control. J. Am. Soc. Inf. Sci. Technol. 59, 1493–1512. https://doi.org/10.1002/asi.20857](https://www.zotero.org/google-docs/?FHjHwm)

[Berinsky, A.J., 2017. Measuring Public Opinion with Surveys. Annu. Rev. Polit. Sci. 20, 309–329. https://doi.org/10.1146/annurev-polisci-101513-113724](https://www.zotero.org/google-docs/?FHjHwm)

[Davis, J.L., Jurgenson, N., 2014. Context collapse: Theorizing context collusions and collisions. Inf. Commun. Soc. 17, 476–485. https://doi.org/10.1080/1369118X.2014.888458](https://www.zotero.org/google-docs/?FHjHwm)

<de Vreese, C.H., Boomgaarden, H., 2006. News, political knowledge and participation: The differential effects of news media exposure on political knowledge and participation. Acta Polit. 41, 317–341. https://doi.org/10.1057/palgrave.ap.5500164>

[Doherty, C., Kiley, J., O’Hea, O., 2018. Wide Gender Gap, Growing Educational Divide in Voters’ Party Identification. Pew Research Center.](https://www.zotero.org/google-docs/?FHjHwm)

[Eady, G., Nagler, J., Guess, A., Zilinsky, J., Tucker, J.A., 2019. How Many People Live in Political Bubbles on Social Media? Evidence From Linked Survey and Twitter Data. SAGE Open 9, 215824401983270. https://doi.org/10.1177/2158244019832705](https://www.zotero.org/google-docs/?FHjHwm)

[Flaxman, S., Goel, S., Rao, J.M., 2016. Filter bubbles, echo chambers, and online news consumption. Public Opin. Q. 80, 298–320.](https://www.zotero.org/google-docs/?FHjHwm)

[Fletcher, R., Nielsen, R.K., 2018. Are people incidentally exposed to news on social media? A comparative analysis. New Media Soc. 20, 2450–2468.](https://www.zotero.org/google-docs/?FHjHwm)

<Garrett, R.K., Gvirsman, S.D., Johnson, B.K., Tsfati, Y., Neo, R., Dal, A., 2014. Implications of pro- and counterattitudinal information exposure for affective polarization: Partisan media exposure and affective polarization. Hum. Commun. Res. 40, 309–332. https://doi.org/10.1111/hcre.12028>

[Graham, T., Jackson, D., Wright, S., 2015. From everyday conversation to political action: Talking austerity in online “third spaces.” Eur. J. Commun. 30, 648–665. https://doi.org/10.1177/0267323115595529](Graham, T., Jackson, D., Wright, S., 2015. From everyday conversation to political action: Talking austerity in online \“third spaces.\” Eur. J. Commun. 30, 648–665. https://doi.org/10.1177/0267323115595529)

[Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., Lazer, D., 2019. Fake news on Twitter during the 2016 U.S. presidential election. Science 363, 374. https://doi.org/10.1126/science.aau2706](https://www.zotero.org/google-docs/?FHjHwm)

[Grootendorst, M., 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure.](https://www.zotero.org/google-docs/?FHjHwm)

[Guess, A., Munger, K., Nagler, J., Tucker, J., 2019. How accurate are survey responses on social media and politics? Polit. Commun. 36, 241–258.](https://www.zotero.org/google-docs/?FHjHwm)

[Guess, A.M., 2021. (Almost) Everything in moderation: New evidence on Americans’ online media diets. Am. J. Polit. Sci. ajps.12589. https://doi.org/10.1111/ajps.12589](https://www.zotero.org/google-docs/?FHjHwm)

[Guess, A.M., 2015. Measure for measure: An experimental test of online political media exposure. Polit. Anal. 23, 59–75. https://doi.org/10.1093/pan/mpu010](https://www.zotero.org/google-docs/?FHjHwm)

[Habermas, J., 1984. The theory of communicative action. Beacon Press.](https://www.zotero.org/google-docs/?FHjHwm)

[Henderson, M., Jiang, K., Johnson, M., Porter, L., 2021. Measuring Twitter use: Validating survey-based measures. Soc. Sci. Comput. Rev. 39, 1121–1141.](https://www.zotero.org/google-docs/?FHjHwm)

[Hindman, M., 2009. The Myth of Digital Democracy. Princeton University Press.](https://www.zotero.org/google-docs/?FHjHwm)

[Hughes, A.G., McCabe, S.D., Hobbs, W.R., Remy, E., Shah, S., Lazer, D.M.J., 2021. Using administrative records and survey data to construct samples of tweeters and tweets. Public Opin. Q. 85, 323–346. https://doi.org/10.1093/poq/nfab020](https://www.zotero.org/google-docs/?FHjHwm)

[Jackson, S.J., Foucault Welles, B., 2015. Hijacking# myNYPD: Social media dissent and networked counterpublics. J. Commun. 65, 932–952.](https://www.zotero.org/google-docs/?FHjHwm)

[Katz, E., Blumler, J.G., 1974. The uses of mass communications: Current perspectives on gratifications research.](https://www.zotero.org/google-docs/?FHjHwm)

[King, G., Persily, N., 2020. A new model for industry–academic partnerships. PS Polit. Sci. Polit. 53, 703–709.](https://www.zotero.org/google-docs/?FHjHwm)

[Lewin, K., 1943. Forces behind food habits and methods of change. Bull. Natl. Res. Counc. 108, 35–65.](https://www.zotero.org/google-docs/?FHjHwm)

[McCabe, S., Green, J., Wan, A., Lazer, D., 2022. New Tweetscores.](https://www.zotero.org/google-docs/?FHjHwm)

[McClain, C., Widjaya, R., Rivero, G., Smith, A., 2021. The behaviors and attitudes of US adults on Twitter. Pew Research Center.](https://www.zotero.org/google-docs/?FHjHwm)

[McInnes, L., Healy, J., Astels, S., 2017.HBSCAN: Hierarchical density based clustering. J. Open Source Softw. 2. https://doi.org/10.21105/joss.00205](https://www.zotero.org/google-docs/?FHjHwm)

[McInnes, L., Healy, J., Melville, J., 2020. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.](https://www.zotero.org/google-docs/?FHjHwm)

[Mellon, J., Prosser, C., 2017. Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. Res. Polit. 4. https://doi.org/10.1177/2053168017720008](https://www.zotero.org/google-docs/?FHjHwm)

[Messing, S., Westwood, S.J., 2014. Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. Commun. Res. 41, 1042–1063.](https://www.zotero.org/google-docs/?FHjHwm)

[Mitchell, A., Shearer, E., Stocking, G., 2021. News on Twitter: Consumed by Most Users and Trusted by Many. Pew Research Center.](https://www.zotero.org/google-docs/?FHjHwm)

[Morrow, G., Swire-Thompson, B., Polny, J.M., Kopec, M., Wihbey, J.P., 2021. The emerging science of content labeling: Contextualizing social media content moderation. J. Assoc. Inf. Sci. Technol.](https://www.zotero.org/google-docs/?FHjHwm)

l<Mukerjee, S., Jaidka, K., Lelkes, Y., 2022. The political Landscape of the U.S. Twitterverse. Polit. Commun. 39, 565–588. https://doi.org/10.1080/10584609.2022.2075061>

[Nanz, A., Matthes, J., 2022. Democratic consequences of incidental exposure to political information: A meta-analysis. J. Commun. 72, 345–373.](https://www.zotero.org/google-docs/?FHjHwm)

[Odabaş, M., 2022. 10 facts about Americans and Twitter. Pew Res. Cent.](https://www.zotero.org/google-docs/?FHjHwm)

[Oser, J., Boulianne, S., 2020. Reinforcement effects between digital media use and political participation: A meta-analysis of repeated-wave panel data. Public Opin. Q.](https://www.zotero.org/google-docs/?FHjHwm)

<Oser, J., Hooghe, M., Marien, S., 2013. Is online participation distinct from offline participation? A latent class analysis of participation types and their stratification. Polit. Res. Q. 66, 91–101. https://doi.org/10.1177/1065912912436695>

[Peterson, E., Goel, S., Iyengar, S., 2021. Partisan selective exposure in online news consumption: evidence from the 2016 presidential campaign. Polit. Sci. Res. Methods 9, 242–258. https://doi.org/10.1017/psrm.2019.55](https://www.zotero.org/google-docs/?FHjHwm)

[Prior, M., 2013. Media and Political Polarization. Annu. Rev. Polit. Sci. 16, 101–127. https://doi.org/10.1146/annurev-polisci-100711-135242](https://www.zotero.org/google-docs/?FHjHwm)

[Schlozman, K.L., Brady, H.E., Verba, S., 2018. Unequal and unrepresented, in: Unequal and Unrepresented. Princeton University Press.](https://www.zotero.org/google-docs/?FHjHwm)

[Sears, D.O., Freedman, J.L., 1967. Selective exposure to information: A critical review. Public Opin. Q. 31, 194–213.](https://www.zotero.org/google-docs/?FHjHwm)

[Shearer, E., 2018. Social media outpaces print newspapers in the US as a news source. Pew Res. Cent. 10.](https://www.zotero.org/google-docs/?FHjHwm)

<Shmargad, Y., Klar, S., 2020. Sorting the news: How ranking by popularity polarizes our politics. Polit. Commun. 37, 423–446. https://doi.org/10.1080/10584609.2020.1713267>

<Shugars, S., Gitomer, A., McCabe, S., Gallagher, R.J., Joseph, K., Grinberg, N., Doroshenko, L., Foucault Welles, B., Lazer, D., 2021. Pandemics, protests, and publics: demographic activity and engagement on Twitter in 2020. J. Quant. Descr. Digit. Media 1. https://doi.org/10.51685/jqd.2021.002>

<Suuronen, A., Reinikainen, H., Borchers, N.S., Strandberg, K., 2021. When social media influencers go political: An exploratory analysis on the emergence of political topics among Finnish influencers. Javn. - Public 1–17. https://doi.org/10.1080/13183222.2021.1983367>

[Taylor, S.J., Muchnik, L., Kumar, M., Aral, S., 2022. Identity effects in social media. Nat. Hum. Behav. https://doi.org/10.1038/s41562-022-01459-8](https://www.zotero.org/google-docs/?FHjHwm)

[Thorson, K., Wells, C., 2016. Curated flows: A framework for mapping media exposure in the digital age. Commun. Theory 26, 309–328.](https://www.zotero.org/google-docs/?FHjHwm)

[Vaccari, C., Valeriani, A., 2021. Outside the bubble: Social Media and Political Participation in Western Democracies. Oxford University Press.](https://www.zotero.org/google-docs/?FHjHwm)

[Vaccari, C., Valeriani, A., Barberá, P., Bonneau, R., Jost, J.T., Nagler, J., Tucker, J.A., 2015. Political expression and action on social media: Exploring the relationship between lower-and higher-threshold political activities among Twitter users in Italy. J. Comput.-Mediat. Commun. 20, 221–239.](https://www.zotero.org/google-docs/?FHjHwm)

[Verba, S., Schlozman, K.L., Brady, H.E., 1995. Voice and Equality: Civic Voluntarism in American Politics. Harvard University Press.](https://www.zotero.org/google-docs/?FHjHwm)

[Williams, M.L., Burnap, P., Sloan, L., Jessop, C., Lepps, H., 2017. Users’ views of ethics in social media research: Informed consent, anonymity, and harm, in: The Ethics of Online Research. Emerald Publishing Limited.](https://www.zotero.org/google-docs/?FHjHwm)

[Wojcieszak, M., Casas, A., Yu, X., Nagler, J., Tucker, J.A., 2022a. Most users do not follow political elites on Twitter; those who do show overwhelming preferences for ideological congruity. Sci. Adv. 8, eabn9418. https://doi.org/10.1126/sciadv.abn9418](https://www.zotero.org/google-docs/?FHjHwm)

[Wojcieszak, M., Menchen-Trevino, E., Goncalves, J.F., Weeks, B., 2022b. Avenues to news and diverse news exposure online: Comparing direct navigation, social media, news aggregators, search queries, and article hyperlinks. Int. J. Press. 27, 860–886.](https://www.zotero.org/google-docs/?FHjHwm)

[Wojcik, S., Hughes, A., 2019. Sizing Up Twitter Users. Pew Res. Cent.](https://www.zotero.org/google-docs/?FHjHwm)

[Wrubel, L., Kerchner, D., 2020. 116th U.S. Congress Tweet Ids. https://doi.org/10.7910/DVN/MBOJNS](https://www.zotero.org/google-docs/?FHjHwm)

<Zhuravskaya, E., Petrova, M., Enikolopov, R., 2020. Political effects of the internet and social media. Annu. Rev. Econ. 12, 415–438. https://doi.org/10.1146/annurev-economics-081919-050239>

1. Absolute number estimates are based on the multiplication of observed amounts in the 10% random sample Decahose by 10. [↑](#footnote-ref-1)