Title
Do Search Algorithms Endanger Democracy? An Experimental Investigation of Algorithm Effects on Political Polarization

Journal
Journal of Broadcasting & Electronic Media, 64(2)

ISSN
0883-8151

Authors
Cho, Jaeho
Ahmed, Saifuddin
Hilbert, Martin
et al.

Publication Date
2020-05-01

DOI
10.1080/08838151.2020.1757365

Peer reviewed

Journal of Broadcasting & Electronic Media

ISSN: 0883-8151 (Print) 1550-6878 (Online) Journal homepage: https://www.tandfonline.com/loi/hbem20

Do Search Algorithms Endanger Democracy? An Experimental Investigation of Algorithm Effects on Political Polarization

Jaeho Cho, Saifuddin Ahmed, Martin Hilbert, Billy Liu & Jonathan Luu

To cite this article: Jaeho Cho, Saifuddin Ahmed, Martin Hilbert, Billy Liu & Jonathan Luu (2020): Do Search Algorithms Endanger Democracy? An Experimental Investigation of Algorithm Effects on Political Polarization, Journal of Broadcasting & Electronic Media

To link to this article: https://doi.org/10.1080/08838151.2020.1757365

Published online: 27 Apr 2020.

JOURNAL OF BROADCASTING & ELECTRONIC MEDIA
https://doi.org/10.1080/08838151.2020.1757365

Do Search Algorithms Endanger Democracy? An Experimental Investigation of Algorithm Effects on Political Polarization

Jaeho Cho[a], Saifuddin Ahmed[b], Martin Hilbert[a], Billy Liu[c], and Jonathan Luu[d]

[a]Department of Communication at University of California, Davis, CA, USA; [b]Wee Kim Wee School of Communication and Information at Nanyang Technological University, Singapore; [c]Software Engineer, Sacramento, CA, USA; [d]Software Engineer, San Diego, CA, USA

ABSTRACT

This study examines algorithm effects on user opinion, utilizing a real-world recommender algorithm of a highly popular video-sharing platform, YouTube. We experimentally manipulate user search/watch history by our custom programming. A controlled laboratory experiment is then conducted to examine whether exposure to algorithmically recommended content reinforces and polarizes political opinions. Results suggest that political self-reinforcement, as indicated by the political emotion-ideology alignment, and affective polarization are heightened by political videos – selected by the YouTube recommender algorithm – based on participants' own search preferences. Suggestions for how to reduce algorithm-induced political polarization and implications of algorithmic personalization for democracy are discussed.

The exponential increase in the amount of information on the Internet has made search engines crucial to the way people navigate online. Indeed, information recommender systems based on big data and algorithms are a key gateway through which most users navigate, select, and consume information online (Beam, 2014; Ricci et al., 2011). The search-recommend systems play an outsized role in facilitating online information consumption and energizing the Internet-based information and communication ecology (Hannak et al., 2013; Lazer, 2015). At the same time, given the degree to which search engines fundamentally shape the nature and type of

CONTACT Jaeho Cho jaecho@ucdavis.edu Department of Communication at University of California, University of California, Davis, 1 Shields Ave., Davis, CA 95616, USA

information encountered, concerns have arisen about how such algorithms impact users (Tufekci, 2018; Zuiderveen Borghesius et al., 2016).

Despite these increasing concerns and debates, little is known about the social and political implications of search algorithms (Hilbert et al., 2018; Thorson & Wells, 2015). The present study seeks to add to the understanding of algorithm effects by providing an empirically testing a theoretical framework. In so doing, instead of making assumptions about the way recommender algorithms operate, we follow an actual real-world recommender system of a highly popular video-sharing platform, YouTube, and experimentally manipulate the search and watch the history of the user by our custom Python script. An experiment is then conducted in a controlled laboratory setting to examine whether exposure to videos provided by YouTube recommender algorithm in response to the experimentally manipulated user search/watch history reinforces and polarizes political opinions. To the best of our knowledge, the current study is one of the first few to empirically assess the impact of search algorithms on user opinion, using observational real-world data (in the form of the behavior of the YouTube recommender system), combined with a closely controlled offline experiment.

Investigating whether and how search engines influence user opinion sheds light on the debate about the role the increasingly complex online information and communication environments play in a democracy. A great deal of research suggests that citizens enabled by a high degree of user control, tend to selectively seek information and maintain communication networks consistent with their political orientations, which eventually reinforces the partisan divide in society (Bennett & Iyengar, 2008; Gentzkow & Shapiro, 2011). On the other hand, there also exists ample evidence that, taking advantage of the increasing choice opportunities in the online sphere, citizens often expose themselves to a range of ideas that cut across partisan or ideological lines (Beam, 2014; Garrett & Stroud, 2014; Song, Cho, & Benefield, 2020). This research points out that information consumption and political discussion online are perhaps more complicated than suggested by the selective exposure literature (Bakshy et al., 2015; Flaxman et al., 2016). Within this context, the present study examines whether and how search algorithms influence user opinion. Previous research has primarily focused on individual motivations and social network characteristics as factors shaping the nature of the effects of online information consumption on the user. However, by considering search algorithms as a *system* factor responsible for such effects, our study adds another layer of complexity to our understanding of how online information environments influence the processes of public opinion formation and democracy.

Algorithmic Personalization

Recommender algorithms installed on online platforms help users efficiently access the information they are looking for while also reducing information overload (Hannak et al., 2013; Lazer, 2015). Based on some kind of input, recommender systems select a set of items from a large pool of available items and recommend a list of information links for users to choose (Resnick & Varian, 1997). The input can be in the form of a search term or some other preference statement, and the selection is traditionally based on the personal history of the user (content-based filtering), or the cross-section of other users (collaborative filtering), or a mix of both (Balabanović & Shoham, 1997).

This transaction between the user and algorithm has a two-fold meaning. While the search algorithms provide the public greater, more efficient access to more relevant information, the provided input by the public also allows online platforms to unobtrusively collect a trove of user data. Since search terms reveal users' underlying interests and preferences, algorithms can make informed guesses about each user and personalize searches and recommendations. This interactive process is similar to what Webster (2011) called *duality* of digital media (or *structuration* in Giddens's terms) – "a process through which agents and structures mutually reproduce the social world" (p. 45). As data about information search and consumption for a specific user accumulates, algorithms' ability to personalize is sharpened in subsequent searches. Thus, while search engine use generally facilitates information consumption, it may also increasingly shape the types of information users encounter. To the extent that user inputs (e.g., search terms, clicks, etc.) reflect their existing political preferences, search engine use could eventually lead to the so-called "filter bubble," a notion that refers to the like-minded information environment created by personalized algorithmic recommendations and the resulting segregation of users in their own ideological and political views (Pariser, 2011).

Algorithmic personalization, which is invisible to users, has been widely adopted by many online platforms. Implemented through flexible and dynamically created websites, algorithms selectively estimate what information a user would like to see. Pariser (2011, p. 9), for example, reports a personal experience wherein two people ended up with different outcomes from an online search for "BP," with one getting investment news about British Petroleum and the other getting information about the Deepwater Horizon oil spill. Similarly, research on web filtering and personalization in e-commerce practices by Hannak et al. (2013) noted that an ambiguous query such as "router" on an e-commerce site might lead to suggestions about either networking devices or woodworking tools depending on user search or consumption history. When viewed in this light, the interactive search process between users and algorithms likely serves to reinforce users' prior preference structure, since recommending algorithms are responsive to users' past behaviors and thus search outcomes are reflective of user interest and preferences.

The underlying reason for the reinforcing tendency of algorithms is that algorithms are naturally designed to operate with, not against, human tendencies, one of which is that people feel more comfortable with consistency and confirmation (Kunda, 1990). Although people do not necessarily avoid attitude-discrepant messages (Ahn et al., 2014; Beam, 2014; Garrett et al., 2013), it is more likely that attitude-consistent messages that resonate with them and confirm their desired conclusions are preferred (Feldman et al., 2018; Gentzkow & Shapiro, 2011; Stroud, 2011). This means that tailoring search outcomes to reinforcing information is key to user engagement and satisfaction and thus to a platform's economic success.

A Spiral of Selectivity and User Opinion

It is proposed that there are two distinct but coexisting human motivations that account for citizens' choice and consumption of political information: directional vs. accuracy (Kunda, 1990, or defensive vs. accuracy in Hart et al., 2009, or cognitive consistency vs. understanding in Holbert et al., 2013). The directional (or defensive) motivation leads to selective exposure, encouraging individuals to seek information (or news sources) that they believe is congenial to their existing attitudes (Iyengar & Hahn, 2009; Stroud, 2011). In contrast, the accuracy motivation drives individuals to seek information that they expect to be of high informational value. With this motivation, individuals often cross ideological lines and glean (or at least do not avoid) information from the other side of the political spectrum (Garrett et al., 2013). Under the increasingly fragmented online information environments, the interplay of the two motivations might be a reason why citizens' information consumption pattern is often less clear cut between partisan selective exposure and cross-cutting exposure (Bakshy et al., 2015; Flaxman et al., 2016).

When coupled with the tendency of psychologically motivated selectivity, algorithmic tailoring and personalization have implications for user opinion. Selective exposure favoring information that confirms one's existing beliefs would be further amplified if individuals' information search and consumption are aided by recommender algorithms. Users' initial pattern of search term selection and information consumption – which is largely driven by prior attitudes and preferences – acquaints the algorithms with the users and thus allows for more accuracy in personalized recommendations. Such user-oriented recommendations, in turn, increase the likelihood of consumption (clicking) by the users, as the recommended messages fit their interests. Then, the consumption of pro-attitudinal information recommended by the algorithms gives reinforcing feedback to the algorithms, confirming the users' preferences and thus validating and enhancing algorithmic personalization. Through this positive feedback loop between the users and recommender algorithms, selective exposure – which starts off based on users' directional or defensive motivation – will be structurally facilitated by the invisible recommender systems.

The heightened selectivity as a result of the user–algorithm interaction induces a personally tailored information environment for each user which, in turn, shapes the user's opinion. A general conclusion of past research on selective exposure is that exposure to confirmatory information strengthens existing opinions (Stroud, 2011; Sunstein, 2001; Tewksbury & Riles, 2015). Pro-attitudinal information is easier to comprehend and accept as it does not create cognitive dissonance, nor provoke extra cognitive processing. That is, information that creates a psychological state of internal inconsistency is more difficult to parse and thus takes more time and effort to process. In contrast, congruent, new information easily fits and is readily integrated into one's existing mental model. Further, attitude-consistent information (or information from a like-minded source) is often viewed as more credible and of higher quality, and this perception not only encourages consumption of the information but also facilitates the processing of it (Miller et al., 1993). Exposure to such information then easily corroborates existing opinions and attitudes, making them more rigid and extreme.

When viewed from the broader literature in political psychology, this research on selective exposure suggests two ways how exposure to like-minded information via personalized algorithmic recommendations would shape users' political opinion: ideological reinforcement and affective polarization. First, schema theory in social cognition research has long suggested that human memory is a network of interconnected cognitive and affective constructs (Anderson & Bower, 1973; Lodge & Stroh, 1993). Such inter-attitudinal structures in a memory play a guiding role during the course of information processing from information selection to interpretation and integration of the information. On the other hand, the attitude structures are renewed and updated when the existing linkages between constructs are strengthened or discrete linkages are added by news constructs (Alba & Hasher, 1983; Nisbett & Ross, 1980). In light of the schema theory, scholars of public opinion have posited that individuals' political opinion is often organized in a coherent structure of attitudes in which domain-specific opinions are grounded in a broader set of political beliefs (or "super-ordinate values or postures" in Converse's terms) such as ideology

(Conover & Feldman, 1984; Converse, 1964; Hurwitz & Peffley, 1987; Kinder & Kam, 2009).

The hierarchical nature of attitude structure as applied to affect has long been recognized as a core component of attitude (Bagozzi & Burnkrant, 1979; Rosenberg & Hovland, 1960). That is, political affect – as "a summary judgment" and "a contemporary response to current circumstances" (Marcus, 2000, p. 228) – is, in large part, rooted in one's long-standing political predispositions (e.g., ideology). Individuals tend to have positive feelings toward an object on their ideological side while attaching negative affective "tags" to that on the opposite side (Fiske & Pavelchak, 1986). The structure of affect-ideology link would then likely be reinforced when individuals are exposed to pro-attitudinal information. Additional ingredients (reasons, factoids, or simple claims) from the information will reaffirm preexisting political ideology, which leads individuals to feel the target object through their ideological lens. Given that ideology serves as a foundation of political affect (Fiske & Pavelchak, 1986), the confirmation of existing ideology would likely result in stronger affect. The strengthened bonding of affect – as a summary judgment (Marcus, 2000) and a component of political attitude (Bagozzi & Burnkrant, 1979) – and ideology is what we refer to as ideological reinforcement.

Second, selective exposure to like-minded information would also likely result in affective polarization – a process of opinion formation in which individuals develop diverging evaluations of attitude objects (e.g., in-group vs. out-group candidates) in a partisan context (Abramowitz & Webster, 2016; Iyengar et al., 2012; Levendusky, 2018). Recent work on public opinion suggests that there is an increasing tendency that people, especially partisans, defend the in-group and dislike the out-group (Iyengar et al., 2012; Levendusky, 2018). This tendency toward affective polarization would likely be magnified when people are exposed to like-minded information, as it endorses their existing partisan attachment and evaluation. If a person is exposed to counter-attitudinal information, her attitudinal ambivalence would increase, thereby reducing affective polarization (Huckfeldt et al., 2004; Mutz, 2006). Yet, algorithmic personalization based on user preference for attitude-consonant information would structurally limit the opportunity to encounter attitude-dissonant information that potentially challenges existing affective attachment to the in-group and/or antipathy toward the out-group.

As discussed earlier, confirmatory algorithmic recommendations and attitude reinforcement are initiated by user inputs, which are often search terms. A key assumption is that the terms that users enter for their online searches are reflective of what they care about and what they believe. Non-self-report, behavioral data about user interests and beliefs allow algorithms to hone further recommendations. Viewed in this light, it is possible that the reinforcing spiral between users and algorithms could be balanced if a diverse set of search terms, at least some of which do not mirror the users', are used. Yet, given the human tendency toward psychological consistency (Kunda, 1990), cognitive salience of counter-attitudinal ideas in the user mind would not be strong enough to drive routine searches online. Indeed, for ideas to be spelled for searches, they should somehow exist and be salient in users' mind chronically or temporarily. Thus, it is less likely that individuals focus on counter-attitudinal ideas for their searches unless motivated otherwise.

Nonetheless, a broadening of search terms could be possible when individuals pick up ideas discussed and shared within their social networks to use in searches. Although social network construction is self-selective, ideas available in social networks, especially those based on online platforms, could go beyond the boundary of self and be more diverse than self-generated ideas (Kwak et al., 2010; Song et al., 2020). Further, in the context of social media, Messing and Westwood (2014) found that individuals often consume news from ideologically misaligned sources when it is recommended by their friends. As social endorsement reduces partisan selectivity toward like-minded information, social media use is an opportunity to "increase users' exposure to a variety of news and politically diverse information" (Messing & Westwood, 2014, p. 1056).

Exposure to and learning about alternative ideas through social networks do not guarantee using the ideas in one's own online searches. However, if individuals are *motivated* to explore ideas that they encounter in their social networks, their search behaviors would likely be less bounded by their own interests and preferences and thus train algorithms on its own. That is, if searches are crafted based on one's perception (or memory) of ideas shared in social networks, the resulting algorithmic recommendations would likely contain a wider range of information or perspectives. Exposure to information that is more diverse or less agreeable in relation to one's own thoughts and beliefs could then lessen self-reinforcement and opinion polarization (Mutz, 2006).

To summarize, algorithmic recommendations are largely an outcome of iterative interactions between users and algorithms, and the algorithmic recommendations are different depending on what search terms users enter. If, as suggested by the notion of selective exposure, information search is driven by a user's own perspectives, algorithms recommend information the user would agree with. In contrast, if a user's search term selection somehow goes beyond her routine repertoire, algorithms would provide more diverse recommendations. It is thus expected that the patterns of ideological self-reinforcement and affective polarization would differ depending on what search terms are used in using a search platform. Drawing on the reasoning provided above, we propose two hypotheses as follows:

H1: Participants who are exposed to content recommended by algorithms provided with self-generated search terms will exhibit a stronger association between candidate evaluation and political ideology – than participants exposed to content recommended by algorithms provided with socially-generated search terms – and those exposed to content recommended by algorithms provided no prior search terms (the control group) (Ideological reinforcement hypothesis).

H2: Participants who are exposed to content recommended by algorithms provided with self-generated search terms will exhibit stronger affective polarization – than participants exposed to content recommended by algorithms provided with socially-generated search terms – and those exposed to content recommended by algorithms provided no prior search terms (the control group) (Affective polarization hypothesis).

Method

Experimental Procedure

An experiment that manipulated YouTube's video recommendation was conducted in a controlled laboratory. A total of 108 undergraduate students in a university in the western United States participated in the experiment. All participants were volunteers and were awarded course extra credit for their participation (Sex: 68% female; Race: 23% White, 3% Black, 53% Asian, 12% Other; Age: M = 22.06, SD = 2.65). The experiment was carried out on November 14–18, about 1 week after the U.S. presidential election in 2016 (related methodological details can also be found in Hilbert et al., 2018).

To experimentally manipulate YouTube's video algorithmic recommendation, we first measured participants' initial topic preferences, which serve as basic input data (i.e., search terms) for recommender algorithms. After signing up for the study and filling out consent forms, participants took a short preliminary survey, which presented a list of 13 terms or statements for each of the two candidates in the 2016 US presidential election, Hillary Clinton and Donald Trump. Of the 26 terms/statements, 16 (8 per candidate) were copied from the two candidates' official campaign websites (https://www.donaldjtrump.com/contract/for Trump campaign and https://www.hillaryclinton.com/issues/for Clinton campaign), and the rest (5 for each candidate) were selected from Google Trend. Some examples are as follows: "Making college debt free and reducing student debt" and "Build a wall against illegal immigration at the Mexican border" from the candidates' campaign websites; "Trump lies" and "Lock her up" from Google Trend (see Appendix 1 for the full list of 26 terms/statements). These 26 items were randomly mixed prior to being presented to participants.

Participants were then asked to rank the statements twice following two separate instructions. The first prompt asked, "If you would search online for specific topics of the 2016 presidential election that interest you personally, what would these topics be? Please rank 10 (out of 26) of the following topics according to your personal interest." Because the goal of this task is to identify topics that interest each participant the most, it was unnecessary for them to rank all 26 items. In addition, ranking 26 items would be challenging and thus prone to errors. Rather, participants were asked to be selective, concentrating on their top 10 topics, and they were also instructed to take their time and rearrange their ranking until they become confident of the ranking. Next, we asked separately, with the same list of 26 statements, "When online (social media, email, etc.), which of the following topics are you likely to see posted or recommended by your friends and online circles? Please rank the top 10 (out of 26) statements that could come from one of your online contacts." The first identifies "self" search preference while the second taps into "social" preference at least in participants' perception.[1]

The rankings were then used to create two lists (self vs. social) of search terms individually tailored based on participants' search preferences. To do this, two steps were taken. First, statements, if selected, were transformed into words that can capture the gist of their meaning. This was done because search terms are typically in the form of a small group of words, which makes the search more targeted and efficient. For example, the statement, "Build a wall against illegal immigration at the Mexican border," was reduced to "Build wall Mexican"; "Wall Street reform to make Wall Street must work for Main Street" to "Wall Street reform" (see Appendix 1). Terms adopted from Google Trend were not modified, as they were already in the form of words such as "Crooked Hillary" or "Trump's sexual assault."

We then created a list of search terms for each participant based on their responses configured in the form of simple words. Although we asked participants to select and rank their top 10 topics, some participants only ranked 7. To standardize, we decided to use the top seven topics chosen by participants. With the seven topics (or keyword phrases) per participant, we created a list of 70 search terms which consisted of top seven topics in the order ranked by each participant (7), all possible pair-wise combinations of top seven topics (21), randomly constructed 35 combinations involving three of the top seven topics (35), and the seven topics in reverse order (7). This combinatorial logic was developed in an effort to achieve two goals: 1) making input data (i.e., search terms) reflective of participants' *strongest* opinions and 2) generating a list of search terms long enough to affect the YouTube recommender engine while avoiding using the same terms in the same order. The decision of how many search terms are necessary for training a search/recommender system would ideally be informed by the architecture of the algorithm, and it would vary system by system. With no knowledge of how the YouTube algorithm is

[1]Although aided and also limited by the pre-selected lists, the identification of two kinds of topic preference, self and social, was made in a controlled manner.

designed, we made the decision of using 70 terms based on our own pilot observations of its adaptive sensitivity with some different numbers of inputs. Overall, the same practice was done twice with each ranking, self-generated and socially generated, producing two lists of topic preference (70 terms each), each of which was used to train the YouTube recommender system based on either "self" preference or "social" preference.

Next, we created a brand new YouTube account for each participant for the sole purpose of this study. We then used a custom Python script in PyCharm compatible with a Chrome browser extension to bias these YouTube accounts with each corresponding users' search preferences, a randomly selected one-third with self-preference and another one-third with social-preference.[2] Once functional, our custom script automatically logged into an unbiased, brand-new YouTube account created for the purpose of this experiment and searched YouTube by taking the first item of the list. Then, the first video recommended by YouTube was played for 7 s by the script, which let it registered into the YouTube account's watch history.[3] Our script repeated the same process for the next 69 search items in the list to bias the YouTube algorithm for each account created for each participant by running the individually pre-determined search terms and playing recommended videos. Once the process of biasing the search/watch history of each YouTube account is done, our custom-programmed script scraped all videos recommended by YouTube for each account, which are available at youtube.com/feed/recommended, and made a list of algorithm-recommended videos. Then, the first five videos of the list were employed as experimental treatments. Through our manual check, nonpolitical videos were excluded once identified, with the assumption that users when in search for information about a specific topic (the presidential election of a week ago in this case) would be less likely to watch videos not relevant to their search. Such nonpolitical videos were not frequent and they were mostly about trendy topics at the time such as popular songs, movies, TV shows, sporting events, etc.[4] In a few cases, we found up to two nonpolitical

[2]The Python script is available at https://github.com/martinhilbert/YouTube-recommender-bias. Note that the script might be outdated because YouTube updates the platform regularly.

[3]The first recommended video was selected based on previous evidence suggesting that users often click on top-ranked search results (Bakshy et al., 2015).

[4]We have no clue why nonpolitical videos, although only in a few cases, were recommended in response to "political" search terms. Yet, given that most of these videos carry the content of what was trending online at the time, we speculate that the overall popularity of content on the platform might be somehow considered in the recommender algorithm, which gives some of the trending videos a higher chance to be recommended. It would be certainly interesting to examine how nonpolitical videos are recommended to users in search for political content and how the users would process the videos if they watched them.

videos from the first five and replaced them with the next videos in the recommended list.

In sum, through all these processes, we prepared a set of unique five videos for each participant, based either on the participant's own preferences or what they would expect to see from their social environment, and conducted a laboratory experiment with three conditions – self, social, and control. For the control condition, our script searched YouTube with the term, "the 2016 Presidential Election," using an unbiased account on the day of the experiment and selected the first five videos from search results. These videos were used as a baseline control – since they were the videos that a user who was in search for videos about the same topic (i.e., the presidential election of past week) – but did not have a specific search/watch history would encounter**,** per YouTube's recommendation at the time. It is a dynamic baseline, as the current YouTube environment simultaneously affects both the personalized recommendations and the trending videos.

After checking in, participants were given an individual space equipped with a computer, completed a pre-exposure questionnaire assessing their emotional reaction toward each of the two major-party candidates, watched the experimentally arranged videos, and finally finished a post-exposure questionnaire about their emotional reactions toward the two candidates.

Measures

With the pre-exposure questionnaire, we measured emotions each participant initially had about two candidates – two positive emotions, joy, and hope; three negative emotions, anger, fear, and sadness; and an overall feeling thermometer. The five discrete emotions have long been identified as primary human emotional states (Ekman et al., 1969; Frijda, 1988) and have been considered as evaluative judgments in public opinion research (Brader, 2006; Marcus et al., 2000). In addition, the valence-based dimensional approach (i.e., feeling thermometer) has also been widely used in assessing individuals' overall feeling toward an attitude object (Iyengar et al., 2012; Redlawsk, 2006). Both approaches are adopted in the present study.

Contemporary research has suggested that discrete emotions, even those that share the same valence (e.g., anger, fear, sadness), have differential behavioral implications. According to the functional theory of emotion, anger, and fear facilitate action readiness while sadness often leads to behavioral withdrawal (Frijda, 1988). Further, the affective intelligence theory (Marcus et al., 2000) distinguishes between anger and fear. The former induces habit/conviction-based actions through disposition system whereas the latter motivates thought/information-based actions via the surveillance system. Taken together, given the differential nature of discrete emotions, the *degree* of algorithm effects on user opinion at the affective level might vary across different categories of emotions. Despite the possibility, however, it is less likely that the overall *direction* (and *nature*) of effects in reinforcement and polarization changes, depending on emotional states. The main goal of this study is to test the hypotheses with a wide array of emotions above and beyond the overall feeling thermometer. Yet, it will be certainly fruitful for future research to examine more nuanced effects of the algorithm across different discrete emotions.

Specifically, five discrete emotions were measured by asking participants how frequently they felt each emotion toward each candidate on a 5-point scale**, ranging** from "Never" (1), "Some of the time," "About half the time," "Most of the time," to "Always" (5): $\text{joy}_{\text{ClintonPre}}$: M = 2.13, SD = 1.18; $\text{joy}_{\text{TrumpPre}}$: M = 1.34, SD = 0.74; $\text{hope}_{\text{ClintonPre}}$: M = 2.63, SD = 1.14; $\text{hope}_{\text{TrumpPre}}$: M = 1.57, SD = 0.87; $\text{anger}_{\text{ClintonPre}}$: M = 2.23, SD = 0.94; $\text{anger}_{\text{TrumpPre}}$: M = 3.35, SD = 1.32; $\text{fear}_{\text{ClintonPre}}$: M = 1.94, SD = 0.97; $\text{fear}_{\text{TrumpPre}}$: M = 3.29, SD = 1.25; $\text{sadness}_{\text{ClintonPre}}$: M = 1.89, SD = 0.99; $\text{sadness}_{\text{TrumpPre}}$: M = 3.12, SD = 1.39. Prior to the video viewing, we also asked participants to rate their emotional valence on a feeling thermometer running from negative (0) to positive (100): feeling $\text{thermometer}_{\text{ClintonPre}}$: M = 46.13, SD = 24.31; feeling $\text{thermometer}_{\text{TrumpPre}}$: M = 21.94, SD = 24.50.

The same set of questions was used to measure participants' feelings toward the presidential candidates after watching five videos: $\text{joy}_{\text{ClintonPost}}$: M = 2.16, SD = 1.20; $\text{joy}_{\text{TrumpPost}}$: M = 1.33, SD = 0.76; $\text{hope}_{\text{ClintonPost}}$: M = 2.71, SD = 1.20; $\text{hope}_{\text{TrumpPost}}$: M = 1.47, SD = 0.73; $\text{anger}_{\text{ClintonPost}}$: M = 2.03, SD = 0.98; $\text{anger}_{\text{TrumpPost}}$: M = 3.32, SD = 1.32; $\text{fear}_{\text{ClintonPost}}$: M = 1.63, SD = 0.84; $\text{fear}_{\text{TrumpPost}}$: M = 2.98, SD = 1.36; $\text{sadness}_{\text{ClintonPost}}$: M = 1.78, SD = 0.85; $\text{sadness}_{\text{TrumpPost}}$: M = 2.82, SD = 1.39; feeling $\text{thermometer}_{\text{ClintonPost}}$: M = 51.57, SD = 24.57; feeling $\text{thermometer}_{\text{TrumpPost}}$: M = 25.44, SD = 25.30.

Participants' political ideology was measured by a question commonly used in the literature in political science (and the American National Election Studies) – "When it comes to politics, do you usually think of yourself as extremely conservative (1), conservative, slightly conservative, moderate, or middle of the road, slightly liberal, liberal, or extremely liberal (7)?" (M = 4.76, SD = 1.21). Political interest was measured by three questions asking if participants are interested in following (a) what is going on in politics, (b) the results of the 2016 presidential election, and (c) the government led by Trump administration (on a 5-point scale with "Not interested at all" being 1 and "Absolutely interested" being 5; Reliability alpha = .79; M = 3.84, SD = .91). Both political ideology and interest were measured only once in the pre-exposure questionnaire.

Using the measures of affective candidate evaluations and political ideology, the two key concepts of the present study (i.e., ideological reinforcement and affective polarization) were assessed empirically. First, ideological reinforcement was examined by the correlation between affective candidate evaluations and political ideology which quantifies the extent to which the two are aligned as a function of the experimental treatment (Cho et al., 2018). Second, affective polarization was measured as the difference between affective evaluations of the two candidates which captures the degree to which attitudes about in-group and out-group candidates diverge.

Results

Randomization Check

A series of chi-square and ANOVA tests were conducted to check between-condition randomization. The results indicate that every relevant pre-exposure variable (i.e., discrete emotions, feeling thermometer measures, ideology, political interest, sex, race, college grade level) is randomly distributed across conditions with the exception of age. The case of age is due to several outliers. When we included age in our analyses as a covariate, the results were substantially the same as those we reported below.

Ideological Reinforcement

To test whether exposure to videos recommended by the YouTube search algorithm reinforces individuals' existing political attitude structure as indicated by the bonding of opinion and ideology, we assessed the correlation between political ideology (i.e., on a scale of 1 to 7, with 7 being "extremely liberal") and the relative affective assessment of the two candidates, calculated by subtracting the post-exposure emotional rating for one candidate from that for the other. Taking advantage of the pre- and posttest measurements, partial correlation analyses were conducted. That is, the estimated correlation coefficients indicate how closely aligned with one's predispositional (pre-exposure) political ideology one's emotional evaluation of the candidates is after watching videos, with the corresponding pre-exposure affective assessment being held constant. This correlation is then statistically compared by z-statistics across the three conditions (i.e., self, social, and control) to see whether ideological reinforcement occurs differently depending on which video treatment participants received.

As shown in Table 1, the results suggest that the correlation between anger and political ideology in the "self" condition (r = −.521, p < .01) is statistically greater than those in the other two conditions, "social" (r = −.004, n.s.) and control (r = .071, n.s.). Given how the two variables were measured (i.e., the relative anger = anger toward Clinton – anger toward Trump; ideology with "Liberal" high), the negative correlation across all conditions indicates, not

Table 1. Testing the ideological reinforcement hypothesis: Partial correlations between post-exposure affective assessments of candidates and political ideology across conditions with pre-exposure affective assessments controlled.

	Conditions		
Affective assessments	Self (n = 35)	Social (n = 38)	Control (n = 35)
$Anger_{Clinton-Trump}$	−.52***[ab]	−.004[a]	.07[b]
$Fear_{Clinton-Trump}$	−.16	−.09	−.45***
$Sadness_{Clinton-Trump}$	−.32*[a]	.11[ab]	−.48***[b]
$Joy_{Clinton-Trump}$	.27	−.05	.32*
$Hope_{Clinton-Trump}$	.17	.11	.27
Feeling $thermometer_{Clinton-Trump}$	.17	.09	.39**

Notes. Each affective assessment was constructed as a relative measure by subtracting ratings for Trump from ratings for Clinton, with higher scores indicating a stronger preference for Clinton over Trump. Entries are partial correlation coefficients between each post-exposure affective assessment and political ideology (Liberal = high) with the corresponding pre-exposure affective assessment controlled (* p <.10; ** p <.05; *** p <.01). Coefficients with the same superscript in each row are statistically different as tested by z-statistics (p <.05, one-tailed).

surprisingly, that participants who identified themselves as liberal tend to feel less anger toward Clinton than toward Trump. That is, candidate assessment on the emotional level is more closely aligned with political ideology for participants who watched YouTube-recommended videos based on a manipulated "self" watch history than for participants who watched videos recommended based on a "social" watch history or no watch history at all (i.e., the control group). Likewise, the correlation between political ideology and sadness is significantly stronger in the "self" condition (r = −.322, p < .10) than that in the "social" condition (r = .111, n.s.).

A similar, albeit not statistically significant, the pattern was observed for all other measures of emotion. The ideology-emotion alignment in participants' minds is stronger for those in the "self" preference condition, compared to those in the "social" preference condition: self r = −.155 (n.s.) vs. social r = −.087 (n.s.) for fear, self r = .270 (n.s.) vs. social r = −.052 (n.s.) for joy; self r = .173 (n.s.) vs. social r = .111 (n.s.) for hope; self r = .165 (n.s.) vs. social r = .087 (n.s.) for overall valence as indicated by the feeling thermometer measure. Further, the ideology-emotion correlation in the control condition is, in general, greater than that in the "social" condition: social r = −.004 (n.s.) vs. control r = .071 (n.s.) for anger, social r = −.087 (n.s.) vs. control r = −.445 (p < .01) for fear, social r = .111 (n.s.) vs. control r = −.482 (p < .01) for sadness, social r = −.052 (n.s.) vs. control r = .320 (p < .10) for joy; social r = .111 (n.s.) vs. control r = .272 (n.s.) for hope; social r = .087 (n.s.) vs. control r = .393 (p < .05) for the feeling thermometer measure. Of these, the difference between the social condition and the control condition for sadness turned out to be statistically significant, based on z-statistics.

To summarize, an overall pattern observed in the results is that the degree of the association between emotional assessments of candidates and political ideology is stronger in the "self" condition than in the "social" condition. This suggests that ideological reinforcement, as indicated by the political emotion-ideology alignment, is heightened by political videos selected by the YouTube recommender algorithm based on participants' own search preferences and is muted to some degree if *perceived* preferences of participants' online social circle are used as inputs. This pattern is particularly pronounced for negative emotions such as anger and sadness.

Affective Polarization

To test whether exposure to videos recommended by the YouTube search algorithm leads to polarized affective assessments of candidates, we ran a series of ANCOVA comparing scores of affective polarization calculated by taking the absolute value of the difference between affective assessments of the two candidates (Fiorina & Abrams, 2008; Iyengar et al., 2012), measured after video exposure. Pre-exposure affective polarization was considered covariate for each ANCOVA.

The results suggest that there were no significant differences in affective polarization across the three conditions (i.e., self, social, and control), especially when it came to specific discrete emotions (i.e., anger, fear, sadness, joy, and hope) (see Table 2). Yet, the overall pattern, although not statistically significant, is that affective polarization is lowest in the "social" condition, with the exception of joy, which is lowest in the "self" condition. This general pattern is also observed, this time at a statistically significant level, for the feeling thermometer. As shown in Table 2, participants in the "social" condition reported a lower level of polarized affective assessments (M = 31.89, SD = 2.64) in terms of overall emotional valence (i.e., feeling thermometer) than those in the "self" condition (M = 39.87, SD = 2.74) and in the control condition (M = 42.88, SD = 2.75) ($F(2, 104) = 4.47$, $p < .05$). Post hoc comparison confirms that a significant difference was between the "social" condition and each of the other two conditions, respectively, with the difference between the "self" and the control conditions not being significant.

In sum, the results indicate that affective polarization is not heightened by videos YouTube recommends – based on either users' self-preferences or preferences of users' social networks – as affective polarization in both conditions is no greater than that in the control condition. However, there is a tendency for videos recommended based on "social" preferences to reduce affective polarization, a tendency which is statistically meaningful, especially when affective assessments of candidates are measured in the form of overall valence on the feeling thermometer.

Additional Analysis

Although not hypothesized, we also expect that algorithm effects on opinion reinforcement and polarization would be stronger for those who are

Table 2. Testing the affective polarization hypothesis: Mean differences in affective polarization measures across conditions.

	Conditions			
Affective polarization	Self (n = 35)	Social (n = 38)	Control (n = 35)	*F(df)*
Anger	1.85 (.13)	1.61 (.13)	1.78 (.13)	.86(2, 104)
Fear	1.62 (.17)	1.30 (.16)[a]	1.80 (.17)[a]	2.37(2, 102)#
Sadness	1.57 (.16)	1.15 (.15)	1.53 (.16)	2.23(2, 104)
Joy	1.06 (.16)	1.27 (.15)	1.42 (.16)	1.32(2, 104)
Hope	1.56 (.16)	1.37 (.16)	1.80 (.16)	1.81(2, 104)
Feeling thermometer	39.87 (2.74)[a]	31.89 (2.64)[ab]	42.88 (2.75)[b]	4.47(2, 104)*

Notes. Affective polarization was measured as the absolute value of the difference between affective assessments of the two candidates, Clinton, and Trump. Pre-exposure affective polarization was considered as a covariate for each ANCOVA. Entries are estimated marginal means with standard errors in parentheses. Means with the same superscript in each row are statistically different as tested by posthoc pairwise comparison (p <.05). For F-test, # p <.10; * p <.05.

engaged in politics. This possibility was tested by including participants' political interest as a moderator. To extend H1, political interest was specified as a factor moderating the relationship between candidate assessments and ideology across conditions. As described earlier, the measure of ideology runs from "Extremely Conservative (1)" to "Extremely Liberal (7)," and candidate assessments were measured, with higher scores indicating a stronger preference for Clinton over Trump. Political interest is a composite measure consisting of three measures of psychological engagement in the 2016 election and politics.

A three-way interaction (ideology x experimental condition [each of two dummy variables with control being the reference category] x interest) was assessed in a regression equation predicting post-exposure candidate evaluation, with pre-exposure candidate evaluation included as a control. The three-way interaction was not significant; likewise, two-way interactions involving interest were not significant either. To extend H2, in the ANCOVA model, political interest was added as a random factor and specified to interact with the fixed factor (i.e., conditions), and the results for the interaction were not significant. Although there were no significant findings, we still think it is an important direction for future research.

Discussion

The results of our experiment showed an overall pattern of opinion reinforcement and polarization after exposure to algorithm-recommended content, especially when algorithms were fed with self-generated search terms. In contrast, such a pattern of reinforcement and polarization was less salient when algorithms were fed with socially generated search terms. Thus, algorithmic recommendations personalized by user behavior data, if unchecked, has the potential to solidify personal political convictions and encourage polarized opinions. If our findings from YouTube can be generalized to search engines and algorithms running on other online platforms, algorithms and big data are responsible, at least in part, for the increasingly divisive political culture in many contemporary democracies. Due to the programmed responsiveness to past user interests and preferences, algorithms serve as a confirmatory communication partner or an echo chamber that reassures and reinforces users' prior beliefs and fosters extremism, as observed in the case of politically homogeneous communication networks. Viewed in this light, algorithms, although helping users navigate a massive amount of information, add to political selectivity that separates people from contradictory views.

On the other hand, algorithm effects on user opinions are not technologically determined. Rather, the role-played by users in such effects is both relevant and important, since algorithmic recommendations are a product of users' search terms and subsequent consumption of information. As shown in the results, ideas *perceived* to be discussed in one's social circles, if adopted as search terms by the user, weaken reinforcing effects of algorithms. Thus, there is the possibility that political selectivity can be lessened by algorithms if a range of search terms that goes beyond users' political beliefs and identity is used. Given the reinforcing dynamics between user inputs and algorithmic personalization, key to the nature of algorithm effects is how people search and consume information. Existing evidence suggests that the pattern of individuals' information seeking and consumption is more nuanced and multifaceted than assumed in the notion of selective exposure and homophily. People, in general, do not necessarily avoid counter-attitudinal information (Beam, 2014; Garrett et al., 2013), and further seek such information when they believe it helps them understand and be accurate on the situation at hand (Hart et al., 2009; Holbert et al., 2013; Kunda, 1990). In addition to the internal motivation toward accuracy and understanding, social endorsement trumps partisan selectivity in shaping individuals' choice of what information to consume (Messing & Westwood, 2014). Coupled with this research, our results suggest that algorithm effects on opinion reinforcement and polarization are conditional and would be better understood when considering how and to what extent cross-cutting exposure, as opposed to selective exposure, occurs.

This work is only a first step toward theorizing and empirically testing algorithm effects on the user. As such, there are several points that need to be acknowledged as limitations of the present study and considered in future research on algorithm effects. First, this study utilized an actual real-world, not proxy, algorithm on a highly popular online platform, a fact that greatly enhances the external validity and relevance of the findings. However, online platforms operating in different contexts program and run their algorithms differently. This is why online search and recommender algorithms are essentially black boxes. Although personalization by users' past behavior is a key mechanism, the complex and dynamic algorithmic systems handling big data make it difficult to predict outcomes (Lazer, 2015). As such, it would be useful to see more evidence of algorithm effects across different platforms.

Second, user opinion, in this study, was operationalized as affective assessments of two major candidates of the presidential election that was just finished. It is thus highly likely that the opinion targets (i.e., presidential candidates) are much more salient and polarizing in participants' minds than any other issues. Further, the overall body of available information online about the election and candidates was also likely highly polarized at the time when this experiment was conducted. This unique situation might amplify the reinforcing effects by algorithmic personalization as perhaps participants already have strong opinions and information recommended by algorithms might contain partisan or polarizing cues. Future research should consider additional topics to assess whether algorithms reinforce and polarize user opinion.

Third, the way search preferences were measured is another area of limitation. In the present study, participants were asked to choose topics of their interest from a pre-selected set. Although all of the listed topics were adopted from popular real-word sources (i.e., candidates' official campaign websites and Google), this approach only allows participants to indicate their preferences constrained within the list. It would be fruitful to measure search preferences in a more natural, free-form setting, which would enhance the external validity of the measurement. Further, we also suggest a cautious reading of the results concerning the 'social' condition in which search terms were crafted based on participants' *perception* of topics/ideas being shared within their online social circles. We acknowledge that this is not the way that individuals routinely do searches. The results should be interpreted as showing algorithm effects when users are somehow *motivated* to explore ideas that they think their friends care about. Although less realistic, this perception-based "social" condition allowed us to test algorithm effects on users when search inputs go beyond their own preferences and interests. Future research should look at ways to encourage users to broaden their search preferences.

Finally, we cannot draw a firm conclusion from a single experiment with college students. As compared to the average population, college students are in general more active on the Internet and tend to be friends with a more diverse group of people on their online social networks (Pfeil et al., 2009). The relatively open social networks likely lead to a wider diversity of search terms gleaned from their online friends, which could potentially explain the results at least in part. For a more generalizable conclusion, future research should test algorithm effects with a larger sample beyond college students.

Conclusion

In all, it is well worth exploring the impact of algorithms on user opinion. Algorithms certainly facilitate information search and consumption by helping users find relevant information more easily. This upside comes with a risk, however. When it comes to controversial social and political topics, algorithmic personalization could also serve as an echo chamber strengthening users' prior convictions and thus aggravating the existing political chasm in society. Fortunately, the results of our study suggest that the reinforcing effects of algorithms diminish as users employ search terms cued by social networks, which provides clues into how to build socially responsible algorithms. Once again, this study is only a preliminary investigation into the political implications of search algorithms. Our results call for more research on questions such as how concerns and views larger than the boundary of self can be incorporated into one's search terms either by user intention or algorithm, how users select and process algorithm-recommended information that goes beyond their comfort zone, and what impact that a broader range of recommendations would have on user satisfaction and economic viability of platforms.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Jaeho Cho (Ph.D. University of Wisconsin-Madison) is Professor of Communication at the University of California, Davis. His research interest concerns how and under what conditions information and communication contribute to an informed citizenry and a vibrant democracy.

Saifuddin Ahmed (Ph.D. University of California, Davis) is an assistant professor at the Wee Kim Wee School of Communication and Information at Nanyang Technological University. His research interests lie in political communication, election studies, and public opinion with a focus on the transformative role of new media in political inequality.

Martin Hilbert (Ph.D. Socio-Economic Science, F.A.U. Germany; Ph.D. Communication USC USA) is Professor at the University of California, Davis, associated with Communication, Computer Science, and the DataLab. He chairs the Computational Social Science program. His research interests include social algorithms, complex adaptive systems, and consciousness.

Billy Liu (B.S. University of California, Davis) is a software engineer.

Jonathan Luu (B.S. University of California, Davis) is a software engineer.

References

Abramowitz, A. I., & Webster, S. (2016). The rise of negative partisanship and the nationalization of US elections in the 21st century. *Electoral Studies*, *41*(March), 12–22. https://doi.org/10.1016/j.electstud.2015.11.001

Ahn, T. K., Huckfeldt, R., & Ryan, J. B. (2014). *Experts, activists, and interdependent citizens: Are electorates self-educating?* Cambridge University Press.

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, *93*(2), 203–231. https://doi.org/10.1037/0033-2909.93.2.203

Anderson, J. R., & Bower, G. (1973). *Human associative memory*. Winston & Sons.

Bagozzi, R. P., & Burnkrant, R. E. (1979). Attitude organization and the attitude–behavior relationship. *Journal of Personality and Social Psychology*, *37*(6), 913–929. https://doi.org/10.1037/0022-3514.37.6.913

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

Balabanović, M., & Shoham, Y. (1997). Fab: Content-based, collaborative recommendation. *Communications of the ACM*, *40*(3), 66–72. https://doi.org/10.1145/245108.245124

Beam, M. A. (2014). Automating the news: How personalized news recommender system design choices impact news reception. *Communication Research*, *41*(8), 1019–1041. https://doi.org/10.1177/0093650213497979

Bennett, W. L., & Iyengar, S. (2008). A new era of minimal effects? The changing foundations of political communication. *Journal of Communication*, *58*(4), 707–731. https://doi.org/10.1111/j.1460-2466.2008.00410.x

Brader, T. (2006). *Campaigning for hearts and minds: How emotional appeals in political ads work*. University of Chicago Press.

Cho, J., Ahmed, S., Keum, H., Choi, Y., & Lee, J. (2018). Influencing myself: Self-reinforcement through online political expression. *Communication Research*, *45*(1), 83–111. https://doi.org/10.1177/0093650216644020

Conover, P., & Feldman, S. (1984). How people organize the political world: A schematic model. *American Journal of Political Science*, *28*(1), 95–126. https://doi.org/10.2307/2110789

Converse, P. E. (1964). The nature of belief systems in mass publics. In D. E. Apter (Ed.), *Ideology and discontent* (pp. 206–261). Free Press.

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science*, *164*(3875), 86–88. https://doi.org/10.1126/science.164.3875.86

Feldman, L., Wojcieszak, M., Stroud, N. J., & Bimber, B. (2018). Explaining media choice: The role of issue-specific engagement in predicting interest-based and partisan selectivity. *Journal of Broadcasting & Electronic Media*, *62*(1), 109–130. https://doi.org/10.1080/08838151.2017.1375502

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annual Review of Political Science*, *11*(1), 563–588. https://doi.org/10.1146/annurev.polisci.11.053106.153836

Fiske, S. T., & Pavelchak, M. A. (1986). Category-based versus piecemeal-based affective responses: Developments in schema-triggered affect. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 167–203). Guilford Press.

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

Frijda, N. (1988). The law of emotion. *American Psychologist*, *43*(5), 349–358. https://doi.org/10.1037/0003-066X.43.5.349

Garrett, R. K., Carnahan, D., & Lynch, E. K. (2013). A turn toward avoidance? Selective exposure to online political information, 2004–2008. *Political Behavior*, *35*(1), 113–134. https://doi.org/10.1007/s11109-011-9185-6

Garrett, R. K., & Stroud, N. J. (2014). Partisan paths to exposure diversity: Differences in pro- and counterattitudinal news consumption. *Journal of Communication*, *64*(4), 680–701. https://doi.org/10.1111/jcom.12105

Gentzkow, M., & Shapiro, J. M. (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics*, *126*(4), 1799–1839. https://doi.org/10.1093/qje/qjr044

Hannak, A., Sapiezynski, P., Molavi Kakhki, A., Krishnamurthy, B., Lazer, D., Mislove, A., & Wilson, C. (2013). Measuring personalization of web search. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 527–538). Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee. Retrieved from http://dl.acm.org/citation.cfm?id=2488388.2488435

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*(4), 555–588. https://doi.org/10.1037/a0015701

Hilbert, M., Ahmed, S., Cho, J., Liu, B., & Luu, J. (2018). Communicating with algorithms: A transfer entropy analysis of emotions-based escapes from online echo chambers. *Communication Methods and Measures*, *12*(4), 260–275. https://doi.org/10.1080/19312458.2018.1479843

Holbert, R. L., Weeks, B. E., & Esralew, S. (2013). Approaching the 2012 U.S. Presidential election from a diversity of explanatory principles: Understanding, Consistency, and Hedonism. *American Behavioral Scientist*, *57*(12), 1663–1687. https://doi.org/10.1177/0002764213490693

Huckfeldt, R., Mendez, J. M., & Osborn, T. (2004). Disagreement, ambivalence, and engagement: The political consequences of heterogeneous networks. *Political Psychology*, *25*(1), 65–95. https://doi.org/10.1111/j.1467-9221.2004.00357.x

Hurwitz, J., & Peffley, M. (1987). How are foreign policy attitudes structured? A hierarchical model. *American Political Science Review*, *81*(4), 1099–1120. https://doi.org/10.2307/1962580

Iyengar, S., & Hahn, K. S. (2009). Red media, blue media: Evidence of ideological selectivity in media use. *Journal of Communication*, *59*(1), 19–39. https://doi.org/10.1111/j.1460-2466.2008.01402.x

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, *76*(3), 405–431. https://doi.org/10.1093/poq/nfs038

Kinder, D. R., & Kam, C. D. (2009). *Us against them: Ethnocentric foundations of American opinion*. University of Chicago Press.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Kwak, H., Lee, C., Park, H., & Moon, S., (2010). What is Twitter, a social network or a news media? *Proceedings of the 19th International Conference on World wide web, April 26–30, 2010*, Raleigh, NC. https://doi.org/10.1145/1772690.1772751

Lazer, D. (2015). The rise of the social algorithm. *Science*, *348*(6239), 1090–1091. https://doi.org/10.1126/science.aab1422

Levendusky, M. S. (2018). Americans, not partisans: Can priming American national identity reduce affective polarization? *The Journal of Politics*, *80*(1), 59–70. https://doi.org/10.1086/693987
Lodge, M., & Stroh, P. (1993). Inside the mental voting booth: An impression-driven model. In S. Iyengar & W. McGuire (Eds.), *Explorations in political psychology* (pp. 225–263). Duke University Press.
Marcus, G. E. (2000). Emotions in politics. *Annual Review of Political Science*, *3*(1), 221–250. https://doi.org/10.1146/annurev.polisci.3.1.221
Marcus, G. E., Neuman, W. R., & MacKuen, M. (2000). *Affective intelligence and political judgment*. University of Chicago Press.
Messing, S., & Westwood, S. J. (2014). Selective exposure in the age of social media: endorsements trump partisan source affiliation when selecting news online. *Communication Research*, *41*(8), 1042-1063. https://doi.org/10.1177/0093650212466406
Miller, A. G., McHoskey, J. W., Bane, C. M., & Dowd, T. G. (1993). The attitude polarization phenomenon: Role of response measure, attitude extremity, and behavioral consequences of reported attitude change. *Journal of Personality and Social Psychology*, *64*(4), 561–574. https://doi.org/10.1037/0022-3514.64.4.561
Mutz, D. C. (2006). *Hearing the other side: Deliberative versus participatory democracy*. Cambridge University Press.
Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Prentice-Hall.
Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin.
Pfeil, U., Arjan, R., & Zaphiris, P. (2009). Age differences in online social networking: A study of user profiles and the social capital divide among teenagers and older users in MySpace. *Computers in Human Behavior*, *25*(3), 643–654. https://doi.org/10.1016/j.chb.2008.08.015
Redlawsk, D. (2006). Motivated reasoning, affect, and the role of memory in voter decision making. In D. Redlawsk (Ed.), *Feeling politics: Emotion in political information processing* (pp. 87–107). Palgrave Macmillan.
Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, *40*(3), 56–58. https://doi.org/10.1145/245108.245121
Ricci, F., Rokach, L., Shapira, B., & Kantor, P. B. (2011). *Recommender systems handbook*. Springer.
Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In M. J. Rosenberg & C. I. Hovland (Eds.), *Attitude organization and change* (pp. 1–14). Yale University Press.
Song, H., Cho, J., & Benefield, G. A. (2020). The Dynamics of Message Selection in Online Political Discussion Forums: Self-segregation or Diverse Exposure? *Communication Research*. *47*(1), 125–152. https://doi.org/10.1177/0093650218790144
Stroud, N. J. (2011). *Niche news: The politics of news choice*. Oxford University Press.
Sunstein, C. R. (2001). *Echo chambers: Bush v. Gore, impeachment, and beyond*. Princeton University Press.
Tewksbury, D., & Riles, J. M. (2015). Polarization as a function of citizen predispositions and exposure to news on the Internet. *Journal of Broadcasting & Electronic Media*, *59*(3), 381–398. https://doi.org/10.1080/08838151.2015.1054996
Thorson, K., & Wells, C. (2015). Curated flows: A framework for mapping media exposure in the digital age. *Communication Theory*, *26*(3), 309–328. https://doi.org/10.1111/comt.12087

Tufekci, Z. (2018, March 10). *YouTube, the great radicalizer*. *The New York Times*, Retrieved from https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html

Webster, J. G. (2011). The duality of media: A structurational theory of public attention. *Communication Theory*, *21*(1), 43–66. https://doi.org/10.1111/j.1468-2885.2010.01375.x

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

Appendix 1 Search terms/statements

(1) **Clean up corruption** and special interest collusion in **Washington** DC
(2) **Renegotiate NAFTA** and international trade deals
(3) **Remove** two million criminal illegal **immigrants**
(4) Suspend immigration from terror regions and **extreme vetting**
(5) Grow the economy and create millions of new jobs through **tax reduction**
(6) **Repeal** and replace **Obamacare** Act
(7) **Build a wall** against illegal immigration at the Mexican border
(8) **Renegotiate** nuclear **Iran deal**
(9) **Crooked Hillary**
(10) **Lock her up**
(11) **Clinton Foundation corruption**
(12) **Investigation Clinton E-Mails**
(13) **Pray for Trump**
(14) Making sure the wealthy, Wall Street, and corporations pay their **fair share of taxes**
(15) No discrimination of **lesbian, gay, bisexual, and transgender** Americans
(16) Making **college debt-free** and reducing student debt
(17) Undertake a **campaign finance reform**
(18) Taking on the **threat of climate change** and making America clean energy superpower
(19) **Women's rights** and break down barriers that hold women back
(20) **Wall Street reform** to make Wall Street must work for Main Street
(21) Immigration reform with a **path to citizenship**
(22) **Trump lies**
(23) **Trump's sexual assault**
(24) **Trump bankruptcies**
(25) **Lawsuit Trump University**
(26) **How to impeach a President**

Notes:

(1) 1–13 for Trump campaign; 14–26 for Clinton campaign
(2) 1–8 and 14–21 were selected from the two candidates' official campaign websites; the rest from Google Trend.
(3) Words in bold with underline were used as search terms to train the search/recommender algorithm.