

A critical review of filter bubbles and a comparison with selective exposure

Peter M. Dahlgren

Department of Journalism, Media and Communication, University of Gothenburg, Sweden

Abstract

The new high-choice media environment has raised concerns that users of social networking sites primarily select political information that supports their political opinions and avoid information that challenges them. This behaviour is reinforced by personalisation algorithms that create filter bubbles and both narrow the available content and exclude challenging information over time. These concerns have, however, been contested. This article challenges the underlying theoretical assumptions about filter bubbles, and compares filter bubbles to what we already know about selective exposure and human psychology. The article lists nine counterarguments to the filter bubble thesis. In short, I argue that the assumptions of filter bubbles contradict many of the previous findings of selective exposure research. More specifically, when discussing filter bubbles there is a risk of confusing two arguments: one strong – but also trivial – that is about technology (e.g., personalisation leads to different information), and one weak and speculative – but also the most interesting – that is about society (e.g., personalisation increases political polarisation in society).

Keywords: selective exposure, echo chambers, audience fragmentation, confirmation bias, personalisation algorithms

Introduction

With the expansion of the Internet, people now have a wide variety of sources from which to gather information. This has led to two primary concerns: first, users tend to seek information that confirms their existing beliefs, attitudes, or interests (supporting information); and second, Internet services, such as social networking sites and search engines, try to use algorithms to serve up increasingly more supporting information to attract users. These two processes may reinforce each other so that people no longer come into contact with information that challenges their existing beliefs, attitudes, or interests (challenging information). Instead, an individual becomes insulated in a filter bubble: a “personal ecosystem of information that’s been catered by these algorithms to who they think you are” (The Daily Dish, 2010: para. 1), or a “unique universe of information for each of us” (Pariser, 2011: 9).

Dahlgren, P. M. (2021). A critical review of filter bubbles and a comparison with selective exposure. *Nordicom Review*, 42(1), 15–33. <https://doi.org/10.2478/nor-2021-0002>

The existence of filter bubbles has often been asserted and taken for granted among journalists, politicians, and scholars alike. Ironically, both research funding and research papers already exist that address how to break out from a filter bubble (e.g., Bozdag & van den Hoven, 2015; Vinnova, 2017), even though the existence of these bubbles was not demonstrated before this research was done. The filter bubble thesis was primarily discussed on the basis of anecdotes about the different results obtained when two people googled for the ambiguous term “BP”, and these different results were then extrapolated to the whole of society and to the political process (see Pariser, 2011).

The technology underlying many Internet services can change quickly, and today’s findings may be obsolete tomorrow. Human evolution, on the other hand, is much slower.¹ Instead of collecting more empirical data on how and when algorithms supposedly lead to filter bubbles, a different approach is to examine whether the assumptions underlying the filter bubble thesis (and their implications) are consistent with what we already know about selective exposure and human psychology, such as how individuals tend to select information and the causes of political polarisation.

The purpose of this critical review is therefore to scrutinise the assumptions underlying the filter bubble thesis and, more importantly, to challenge the claims that have been pushed too far, by raising several important counterarguments. The following counterarguments are found under the corresponding headings in this article:

1. Filter bubbles can be seen at two levels: technological and societal
2. People often seek supporting information, but seldom avoid challenging information
3. A digital choice does not necessarily reveal an individual’s true preference
4. People prefer like-minded individuals, but interact with many others too
5. Politics is only a small part of people’s lives
6. Different media can fulfil different needs
7. The United States is an outlier in the world
8. Democracy does not require regular input from everyone
9. It is not clear what a filter bubble is

This review makes an important contribution in terms of theory. Filter bubbles may operate, as suggested, at a *technological* or individual level (such that when a user follows recommended videos on a video service, the videos recommended to them the next time are similar). However, there is still a large leap from here to filter bubbles in the broader *societal* level envisioned by Pariser, which result in 1) people seeing no challenging information, and 2) an increase in political polarisation in society. It is this leap from the technological to the societal level that is unsubstantiated. In fact, several studies have shown the complete opposite to what the filter bubble thesis suggests, especially when taking this societal level into account (e.g., Barberá, 2015; Beam et al., 2018; Bruns, 2019a; Davies, 2018; Möller et al., 2018; Nelson & Webster, 2017; see also Boxell et al., 2017; Zuiderveen Borgesius et al., 2016).

Before turning to the counterarguments and reviewing selective exposure, filter bubbles, and human psychology, a brief explanation of the main assumptions behind filter bubbles, as well as their predicted outcomes, is laid out.

What is a filter bubble?

The term filter bubble was coined by the journalist and activist Eli Pariser (2011) in his book *The Filter Bubble*, and made widely known in his TED Talk in 2011.² Although filter bubble is a new term, and arguably the best-known by now, several scholars previously talked about a “daily me” of an online newspaper with a total edition of one (Negroponte, 1995) and of “echo chambers” of only supporting information (Sunstein, 2001). Pariser argued that, to a larger degree than before, Internet services personalise (i.e., customise) information to the specific user with the use of algorithms based on past behaviours (Pariser, 2011), and in an interview, he described a filter bubble as a “personal ecosystem of information that’s been catered by these algorithms to who they think you are” (The Daily Dish, 2010: para. 1). More specifically, he has written,

Internet filters looks at the things you seem to like – actual things you’ve done, or the things people like you like – and tries to extrapolate. They are prediction engines, constantly creating and refining a theory of who you are and what you’ll do and want next. Together, these engines create a unique universe of information for each of us – what I’ve come to call a filter bubble. (Pariser, 2011: 9)

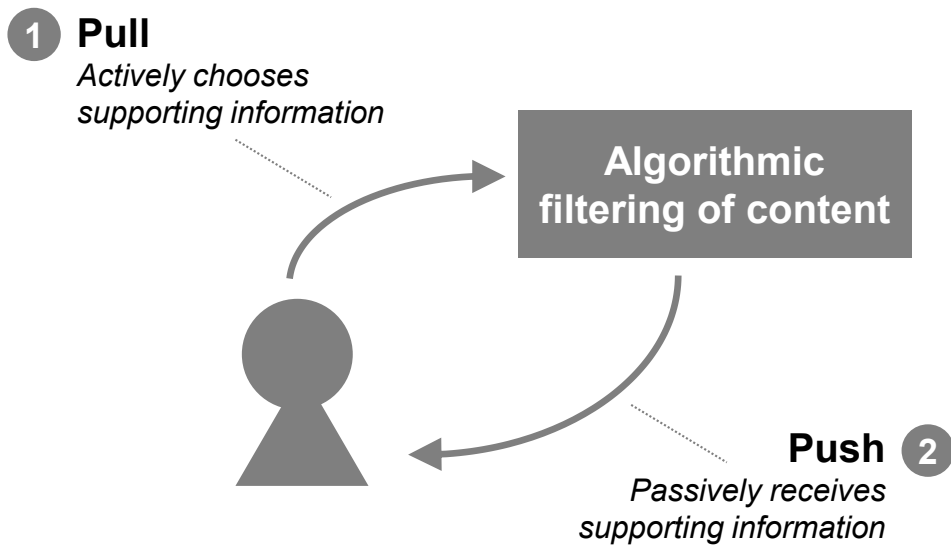
These filter bubbles emerge when users actively seek information and the Internet services learn what users consume. The Internet services then try to provide users with more of the same content during their next visit, based on predictions from past behaviours. As Pariser (2011: 125) puts it,

You click on a link, which signals an interest in something, which means you’re more likely to see articles about that topic in the future, which in turn prime the topic for you. You become trapped in a you loop, and if your identity is misrepresented, strange patterns begin to emerge, like reverb from an amplifier.

This would lead to a self-reinforcing spiral between, on the one hand, the confirmation bias of human psychology and, on the other hand, the provision of Internet services of giving us more supporting information (see Figure 1): “personalization algorithms can cause identity loops, in which what the code knows about you constructs your media environment, and your media environment helps to shape your future preferences” (Pariser, 2011: 233). Filter bubbles are therefore not just a state in time, but also a process that evolves into increasingly more personalised information, which ultimately makes it impossible to find challenging information. The implications, according to Pariser (2011), are threefold: you become the only person in your bubble, you cannot know how information is filtered out, and you cannot choose whether you want to enter this process.

Pariser (2011) mentions that filter bubbles may have several outcomes for individuals: narrower self-interest; overconfidence; dramatically increased confirmation bias; decreased curiosity; decreased motivation to learn; fewer surprises; decreased creativity and ability to innovate; decreased ability to explore; decreased diversity of ideas and people; decreased understanding of the world; and a skewed picture of the world. Eventually, “You don’t see the things that don’t interest you at all” (Pariser, 2011: 106) because the filter bubble “will often block out the things in our society that are important but complex or unpleasant. It renders them invisible. And it’s not just the issues that disappear. Increasingly, it’s the whole political process” (Pariser, 2011: 151). The only

Figure 1 *More of the same: The self-reinforcing spiral of how a filter bubble emerges*



Comments: 1) A user actively seeks and chooses supporting information, and 2) the user passively receives supporting information from Internet services that have predicted what the user would want next.

way to avoid filter bubbles, Pariser claims, is to turn off our devices. This would only help for a short time, however, because the physical and virtual world will blend together even more in the future, not least through personalised augmented reality (Pariser, 2011).

It has also been said that these bubbles caused both Brexit and the 2016 American election result, because challenging information did not reach the voters (Bruns, 2019a). The filter bubble thesis is therefore interesting precisely because it moves beyond the immediate situation of a single user of technology, and instead focuses on the larger question of societal implications. As Pariser (2011: 12) argues, “when I considered what it might mean for a whole society to be adjusted in this way, [personalisation] started to look more important”, since “the filter bubble is a centrifugal force, pulling us apart” (Pariser, 2011: 9).

Pariser’s book has been highly influential and amassed over 1,200 citations, according to Google Scholar, and filter bubbles have also been referenced by the former American president, Barack Obama. This indicates that many think filter bubbles are something that are important and that may even explain recent developments in society.

In sum, the idea is that filter bubbles mean “search engines and social media, together with their recommendation and personalisation algorithms, are centrally culpable for the societal and ideological polarisation” (Bruns, 2019b: 1).³ To put it briefly, filter bubbles are about the exclusion of the uninteresting at the expense of democracy.

Nine counterarguments to the filter bubble thesis

Nine arguments are now presented, primarily ordered from the micro (technological/individual) to the macro (societal) level, which are intended to counter several of the central claims about filter bubbles. These counterarguments are primarily theoretical,

but they make use of empirical examples when relevant; this is because one can evaluate an argument in two ways – by its foundation or by its implications.⁴

1. Filter bubbles can be seen at two levels: technological and societal

There are two different claims about filter bubbles that must be distinguished. First, filter bubbles can be seen in the immediate technological situation in which any single choice affects the content recommended by personalisation algorithms, thereby narrowing the type of content available over time. This is similar to what happens if a live microphone is placed too close to a loudspeaker: there is an ever-increasing volume when the microphone and speaker are part of the same closed system. Let us call this a filter bubble at the technological (or individual) level. Second, and more broadly, we may see the causes and consequences of these choices and technologies for humans and society, and, more importantly, for the political process and democracy over time. Let us refer to this as filter bubbles at the societal level.

This distinction is not clarified by Pariser (2011), who is concerned with both levels as a chained argument, where one thing leads to another. The distinction between these levels (which should be considered as a continuum, not a dichotomy) would help to reconcile some of the very conflicting claims and research findings about filter bubbles. Some studies at the technological level find evidence of extremely large filter bubbles by using mathematical models or simulations (e.g., Chitra & Musco, 2020). Studies that also take humans and the societal level into account, on the other hand, find virtually no evidence of filter bubbles, but often the complete opposite (for reviews, see Zuiderveen Borgesius et al., 2016; Bruns, 2019a).⁵ For example, a German study of 4,000 users who installed a browser add-on found that between 1 and 3 per cent of users' identical political searches resulted in very different search results on Google (cited in Davies, 2018).

In sum, the distinction between different levels of filter bubbles is important, especially when evaluating mathematical models that do not take humans into account but nonetheless generalise the results to humans or make policy recommendations.

2. People often seek supporting information, but seldom avoid challenging information

It is impossible for a human to consume all existing information. We have to choose, and all information consumption is therefore inevitably selective. When people have a pre-existing attitude or belief, and are given a choice, they are expected (according to selective exposure theory) to seek supporting information and avoid challenging information. This combination of seeking and avoiding has been the primary focus of selective exposure research (Frey, 1986). Similar arguments are made regarding filter bubbles: “identity drives media. But the personalizers haven't fully grappled with a parallel fact: Media also shapes identity” (Pariser, 2011: 124). The media could therefore become “a perfect reflection of our interests and desires” (Pariser, 2011: 12), since confirmation bias will drastically increase because of filter bubbles (Pariser, 2011: 86).

Selective exposure research has shown that people, on average, prefer supporting information to challenging information (Cotton, 1985; D'Alessio & Allen, 2006; Frey, 1986; Hart et al., 2009; Lazarsfeld et al., 1948; Smith et al., 2008), which is a typical case of confirmation bias. This is a well-documented phenomenon, but the difference is

not particularly large, with an average effect size of Cohen's $d = .36$ for selecting supporting over challenging information in general (Hart et al., 2009), and with a somewhat larger effect size for political information in particular (Cohen's $d = .46$).⁶

The evidence for the claim that people avoid challenging information, on the other hand, is much more bleak. This is because people have two motivations: to seek information (which is a moderately strong motivation) and to avoid information (which is a comparatively weak motivation) (Garrett, 2009; see also Fischer et al., 2011; Frey, 1986; Munson & Resnick, 2010). This means that people do not necessarily avoid challenging information.

Nonetheless, one might suspect that individuals with extreme views avoid challenging information. One study found that people who visited sites with white supremacist content were "twice as likely as visitors to Yahoo! News to visit nytimes.com in the same month" (Gentzkow & Shapiro, 2011: 1823). As Bruns suggests, "they must monitor what they hate" (2019a: 97). Meta-analytic findings similarly suggest that the more confident an individual is in their belief, attitude or behaviour, the more exposure they have to challenging information (Hart et al., 2009). Furthermore, relatively few respondents across eight countries reported that they saw news on social networking sites that supports their beliefs or attitudes (Matsa et al., 2018). In the Reuters Institute's annual survey of the populations of 36 countries, 40 per cent of social networking site users agreed with the statement that they were exposed to news they were not interested in, or news sources they did not typically use, while only 27 per cent disagreed with this statement (Newman et al., 2017). Those who used Facebook for news consumption during the American presidential election in 2016 were *more* exposed to news that both confirmed and challenged their beliefs and attitudes, and polarisation *declined* in one study (Beam et al., 2018). Taken together, one can suspect that "there is simply no evidence for selective avoidance or partisan 'echo chambers'" (Weeks et al., 2016: 263).

This has profound implications for the arguments about filter bubbles, because it leaves a lot of room for incidental exposure to challenging information. This may be especially relevant for those who use social networking sites on a habitual basis. In fact, one of the major reasons why people select challenging information is that the information is useful, or they are curious or interested (Hart et al., 2009; Kahan et al., 2017; Knobloch-Westerwick, 2014). In other words, people actually do consume challenging information, but the proportion of information they consume is (on average) slightly tilted toward supporting information.

In sum, people can have a strong motivation to seek supporting information without having a motivation to avoid challenging information. One of the basic tenets of the filter bubble thesis (i.e., that filter bubbles create a drastically increased confirmation bias) is therefore very unlikely.

3. A digital choice does not necessarily reveal an individual's true preference

The central and most important causal argument underlying the existence of filter bubbles is that people's preferences guide their choice of content: "identity drives media" (Pariser, 2011: 124). Personalisation algorithms learn what people prefer and give them more of similar content, which they then happily consume. This means that the diversity of available content is expected to constantly decrease, and, more importantly, that the

previous counterargument (that people do not avoid challenging information) may not be relevant because there would be no challenging information to begin with. People's interests could therefore become increasingly narrower over time, since the media also shapes individuals (Pariser, 2011). As Pariser puts it, "what the code knows about you constructs your media environment, and your media environment helps to shape your future preferences" (Pariser, 2011: 233). Given enough time, Internet services are ultimately expected to become "a perfect reflection of our interests and desires" (Pariser, 2011: 12), which can lead to "information determinism, in which our past clickstreams entirely decide our future" (Pariser, 2011: 135).

The argument claims that an algorithm can detect the content an individual user prefers from the choices and selections that he or she has made, and that this content is then returned to the user, which can lead to a feedback loop over time. This type of argument is a baked-together form of technological determinism (Davies, 2018) and strong behaviourism (Graham, 2017), and there are several problems with arguments of this kind. For example, we should not necessarily assume that people have an active agency when they select content, but become passive and malleable when they receive information. There are also several theoretical reasons why preferences and choices should be kept separate (Hansson & Grüne-Yanoff, 2012; see also Webster, 2017). First, preferences are subjective comparative evaluations and are therefore states of mind, whereas choices are actions. We (or the algorithm) can directly observe what a person selects, but we can never directly observe what a person prefers. The fact that an individual chooses a movie about journalism is therefore not necessarily an indication that the individual prefers journalism (the movie could have good actors). Second, people can make choices that are consistent with their commitments rather than with their preferences (an atheist can visit a religious web site in order to find counterarguments in a debate). Third, choices that have greater consequences may also weaken preferences; for example, choosing to watch a sitcom is likely to satisfy an immediate preference, but choosing to study is more about the likely *outcome* of studying (getting a satisfying job). These are just a few of the reasons why preferences are distinct from choices, as previous research has indicated (Hansson & Grüne-Yanoff, 2012).

These reasons imply that people can have a preference that does not correspond to their choice: they can have competing preferences that are disclosed in different situations (e.g., social desirability or preference falsification); they can make a choice without a preference (e.g., social influence or heuristics); or they can create a new preference from their choice of content (e.g., mere exposure effect). For instance, clicking "like" on something on a social networking site is not only about clicking on supporting information, but is also about what that particular "like" will communicate to others (Hart et al., 2019). As the name implies, *social* networking sites are places where people try to show themselves in a socially favourable light. This tells us that there may be several competing preferences and that different media may reveal different cultures depending on their goals and their users. People also click on news items that have many likes, regardless of whether the news source contains politically challenging information (Messing & Westwood, 2014). This indicates that people do not want to miss out on popular news. Furthermore, not all media consumption is necessarily preference-driven. Media companies may also try to push content towards individuals through advertisements (Webster, 2017).

Until now, I have been considering the theoretical objection. It is also questionable whether preferences can be empirically predicted from digital trace data (e.g., from what people like or share on social networking sites). Meta-analyses have shown that the relationship between our attitudes and our behaviours is very strong, with an average effect size of $r = .79$ (Kim & Hunter, 1993). Nonetheless, this is far from a perfect correlation. Predicting Big Five personality traits from digital traces is even further away from perfect, with an average effect size of $r = .33$ (Settanniet al., 2018). However, one could argue, as Pariser does, that all that is needed is more data and more programmers to “solve” this problem (Pariser, 2011:118). Here we can object by citing the bias–variance trade-off (i.e., underfitting contra overfitting), which means that an algorithm or model may perform very well when it is trained on one set of data, but then perform poorly when applied (generalised) to new unseen data. These are well-known limitations of all theoretical or statistical models, and they will never go away unless one makes a custom model for each and every individual (this should not be confused with personalisation algorithms, which are not necessarily tailored for each and every individual – rather, it is often the *output* of personalisation algorithms that is tailored for each and every individual).

However, some algorithms use collaborative filtering when recommending content to individuals (e.g., recommender systems). This means that the content of the output from the algorithm is determined by relatively large datasets of the behaviours of many users (e.g., big data). This, by definition, is not the same as predicting the content a specific individual will prefer (based on that user’s preference); rather, it means taking the average (or similar) of multiple individuals. This is the opposite of what is suggested by the filter bubble thesis, which is that people are alone in their filter bubble (Pariser, 2011). Instead, collaborative filtering can make use of the wisdom of the crowd, and inject *collectively* gathered content into the feeds of individual users.

In sum, an algorithm can only work on what is observable, which means that there are limits on the ability of an algorithm to predict an individual person’s preferences (this is similar to the critique of behaviourism stemming from cognitive science; see Graham, 2017). Furthermore, a theoretical objection is that people can have multiple preferences that lead to similar choices, and the empirical evidence suggests that the ability to predict an individual’s personality from the content they post online is far removed from the deterministic claims of the filter bubble thesis. In addition, collaborative filtering, for example, stresses the role of the collective rather than the individual. For these reasons, filter bubbles are very unlikely at the technological level.

4. People prefer like-minded individuals, but interact with many others too

If we focus on people rather than technology, users on Twitter, for instance, tend to disseminate (i.e., retweet) messages from like-minded individuals, while messages from those of opposing opinions are not shared to the same extent (Conover et al., 2011). Measured in this way, two distinct political groups can be visualised who are barely in contact with each other. However, if instead we study who is talking to whom, the picture is different. People are regularly exposed to crosscutting messages during political discussions (Conover et al., 2011), as there is essentially more to talk about when we disagree. Nevertheless, there is a case to be made that sharing and disseminat-

ing information may have larger consequences than simply commenting on the information, not least because information reaches more eyeballs when it is disseminated. However, as studies on literal eyeballs, using eye-tracking methods, have shown, people do not always remember what they see on social networking sites even immediately after exposure (Vraga et al., 2016). People may also read content more if the content and its replies demonstrate different opinions (Sülflow et al., 2018).

Social networking sites also tend to become more heterogeneous over time. The more we use social networking sites, the greater the diversity of the people we talk to (Choi & Lee, 2015; Lee et al., 2014), although political networks may become polarised at the same time (Kearney, 2019). People also tend to block others on social networking sites, and one reason for this is that unknown people tend to pop up in discussions (Skoric et al., 2018).

An important reason why polarisation does not arise when we discuss things with others – for example, in the workplace where we are typically exposed to challenging information (Mutz & Mondak, 2006) – is that we prefer to avoid conflict and thus do not reveal what we actually believe or think (Cowan & Baldassarri, 2018). This can lead us to think that our beliefs are shared by others to a greater extent than they actually are. If we are more exposed to challenging information through social networking sites, polarisation can thus occur more easily because (not despite the fact that) we are exposed to challenging information that we find morally objectionable (e.g., Crockett, 2017).

The research results that show whether or not we consume challenging information also depend partly on the method used. For instance, when Internet services passively personalise the newsfeed on Facebook, users tend to be exposed to more challenging information, but when users are in control, they tend to be exposed to less challenging information (Bakshy et al., 2015; note that this study was authored by Facebook, which may have a vested interest in the study's conclusion). Others have found the opposite result (Beam, 2014). This speaks for and against the filter bubbles thesis: *for* because users sometimes select more supporting information, and *against* because personalisation sometimes increases the amount of challenging information. However, it is worth noting that this is a *slight* increase or decrease over time, which is still far from the large (or total) exclusion of challenging information as suggested by the filter bubble thesis.

In sum, audience fragmentation is likely to be an accidental, rather than an essential, property of social networking sites. That is, audience fragmentation may occur for reasons other than algorithms that steer people into fragmentation and isolation. Algorithms aside, people might actually prefer interacting with people who are different from themselves, at least sometimes.

5. Politics is only a small part of people's lives

The filter bubble thesis focuses almost exclusively on political discussions and information, and on the way in which personalisation algorithms filter out people and information with opposing political viewpoints. This implies that politics is a large part of people's lives. However, there is a risk of overstating the problem of filter bubbles by explicitly focusing on only one topic, such as politics, without acknowledging the role of multiple overlapping networks in people's lives. Duggan and Smith (2016: 9) remarked that “a notable proportion of users simply don't pay much attention to the

political characteristics of the people in their networks”. Similarly, we typically interact with about 150 people offline (Hill & Dunbar, 2003), and of those, we tend to discuss politics with only about 9 (Eveland & Hively, 2009).

Many articles on filter bubbles study the network structure of users and estimate the degree of polarisation by looking at either controversial political topics or the participants’ self-reported ideology (e.g., Bakshy et al., 2015; Conover et al., 2011). However, sampling these users does not necessarily tell us how the technology affects users, or how users interact with the technology. It does, however, tell us about discussions between already active and politically engaged users. This sampling bias makes it difficult to infer conclusions about the personalisation algorithm itself. For this reason, the opposite causal direction should also be investigated. Arceneaux and Johnson (2015: 309) argued that “the emergence of partisan news media is more a symptom of a polarized political system than a source”. In other words, it is not necessarily partisan political content that causes people to become polarised; it can also be already polarised people who create and disseminate political content.

Even if people are not interested in politics, the fact that they are connected on a large world-wide network can make it difficult to avoid certain types of information, even if they try to do so. This means that even if an individual is only seeking supporting information, it can become difficult to avoid challenging political information because of the sheer size of the network, the number of sources, and the number of overlapping networks of different kinds (for a discussion of the network aspect of filter bubbles, see Bruns, 2019a).

In sum, the filter bubble thesis focuses primarily on politics. This is natural, given that it predicts negative consequences for democracy, but it also means that it is easy to overstate the problem by sampling of the dependent variable, so to speak. That is, when we sample people who are already extreme political users and discussions that are already politically extreme, we may then infer that the extremeness is a result of filter bubbles. This kind of reasoning can very easily become circular, with the phenomenon that is to be demonstrated simply being assumed.

6. Different media can fulfil different needs

An important point of departure for the filter bubble thesis is that, more and more, we consume news exclusively through Internet services that are personalised for us. According to Pariser, this is inevitable, because the Internet is dominated by major American companies whose interests are motivated primarily by money (Pariser, 2011). Even if one wishes to consume other kinds of information, it is simply not possible to do so without some sort of filtering through the Internet services. The filter bubble thesis assumes that traditional mass media are seldom or never used, and that they have been more or less replaced by algorithms that preselect media content. As Pariser puts it, “you have all of the top sites on the Web customizing themselves to you, then your information environment starts to look very different from anyone else’s” (The Daily Dish, 2010: para. 1). At a societal level, this could have severe consequences. Those in filter bubbles would no longer find out about major news events at all (because we rarely click “like” on news about terrorist attacks or natural disasters, which are consequently filtered out from our bubbles), which means that social networking sites, such as Facebook, would not show major news events in the future (Pariser, 2011).

However, studying one medium at a time, without taking into account the fact that different mediums can fulfil different needs, risks overstating the claims of filter bubbles and their effects. For example, the mass media could be used for general news consumption, and social networking sites for special interests. Studies that take people's entire media diet into account show that users tend to consume from many different media and are consequently exposed to a lot of challenging information (Fletcher & Nielsen, 2017a, 2017b). People still use mass media quite frequently – including through mobile phone push notifications – and those who consume content from smaller niche media outlets that appeal to their interests also consume content from the mass media, as has been seen in all the countries surveyed (Fletcher & Nielsen, 2017b).

Pariser later acknowledged in an interview that mass media may actually still play an important role in the American society:

In fact it's still the case in 2016 that most Americans get their news from local TV news, according to Pew. So I actually think it's very hard to attribute the results of this election to social media generally or the filter bubble in particular. (Newton, 2016: ques. 2)

Fake news has often been discussed in relation to social networking sites and filter bubbles, but the evidence suggests that fake news does not spread particularly far on social networking sites and that the problem has been overstated (Guess et al., 2020). One recent review also suggested that “more people learn about these [fake news] stories from mainstream news media than from social media” (Tsfati et al., 2020: 168). The majority of news consumption does not take place via social networking sites, either, although the share has increased over time. People instead go directly to the homepages of news sites (Flaxman et al., 2016), and those who do not receive supporting information via news do not usually look at news reports at all (Arceneaux & Johnson, 2013).

Another problem regarding relevance, which is related to this point of different media, is that the filter bubble thesis assumes a journalistic lens when evaluating different media, which risks creating problems where none exist. For instance, Facebook is primarily governed by some sort of social ideal by connecting people. This may be a problem if Facebook is considered as a news site and its personalisation algorithms are considered as a type of editor. Why should we make this comparison? It would be strange for anyone to demand from his or her own friends an impartial and balanced news distribution service from which an algorithmic selection could then be made. A requirement like that could barely be satisfied even before social networking sites (or personalisation) existed.

In sum, even if an individual was completely insulated in a filter bubble that excluded challenging information on one site, they could still consume challenging information on other sites. In other words, audience fragmentation happens not only within a medium, but between media as well.

7. The United States is an outlier in the world

The filter bubble thesis originates from the US. This is a country that, over time, has had the most significant decline of trust in both the press and political institutions among some 50 countries (Hanitzsch et al., 2018), has a weak public service broadcaster (public

service broadcasters can have a dampening effect on political polarisation; see Bos et al., 2016; Trilling et al., 2017), and has a two-party system where the Senate and Congress have been heavily polarised since the 1970s (Arceneaux & Johnson, 2015). Therefore, filter bubbles may risk being mistaken for a cause of polarisation, simply from the fact that some parts of the US were already heavily polarised (this is the *cum hoc, ergo propter hoc* fallacy).

Even though this polarisation started before the Internet became mainstream in the US, there is a possibility that these trends have been exacerbated by the Internet. One way to examine this issue is to compare the US with other countries and to carry out analyses over time in order to decide whether the US is an outlier. As it turns out, whether polarisation increases due to the Internet is also highly questionable.

A study on Twitter across 16 countries, for instance, found higher levels of polarisation in two-party systems compared to multi-party systems (Urman, 2019). A recent working paper, using several different measures of polarisation, found that the US had the largest increase in polarisation out of four OECD countries over the past four decades, while five other OECD countries were found to have had a decrease in polarisation over the same period (Boxell et al., 2020). Similarly, another study based on surveys of American citizens found that “polarization has increased the most among the groups least likely to use the Internet and social media” (Boxell et al., 2017: 10612), which is precisely the opposite of what the filter bubble thesis would have predicted.

Studies from within selective exposure research, however, show that media messages may lead to polarisation in some cases. For instance, a two-wave panel survey indicated that “selective exposure leads to polarization” (Stroud, 2010: 556) although the evidence also supported the reverse causal direction – that polarisation leads to selective exposure. At the same time, selective exposure can also weaken over time (Fischer et al., 2011). Taken together, these examples illustrate one of the reasons why a recent review on the effects of media exposure on polarisation summarised the evidence as “mixed at best” (Prior, 2013: 101).

In sum, there is a risk that too much emphasis is put on the Internet and social networking sites when explaining polarisation in a society, especially when a more viable explanation exists: the US may be an exception in the world when it comes to polarisation, not the rule.

8. Democracy does not require regular input from everyone

The normative assumption of filter bubbles is that “democracy only works if we citizens are capable of thinking beyond our narrow self-interest” (Pariser, 2011: 164) because “democracy requires citizens to see things from one another’s point of view” as well as having “a reliance on shared facts” (Pariser, 2011: 5). Even if these are Pariser’s words, similar normative assumptions are commonly found in the literature on selective exposure (Althaus, 2006).

However, it is not self-evident that democracy only works under these conditions, since many democracies delegate the responsibilities of citizens to representatives:

Certainly the interests of many citizens will be at stake in any policy decision, but it is another thing to presume that democracy requires citizens to exercise vigilance over every interest they might have. The institutions of representative

as opposed to direct democracy are designed precisely to avoid encumbering citizens with such an onerous responsibility. (Althaus, 2006: 83)

Consequently, even if filter bubbles give people less exposure (or no exposure at all) to challenging information, it does not follow that democracy stops working, or even diminishes. People can still follow democratic processes (vote, raise concerns with representatives, protest, demonstrate, etc.) even if they only consume supporting information. In fact, filter bubbles with people having a certain interest could just as well be one democratic way of forming interest groups that exert influence on societal institutions and the mass media (Althaus, 2006).

In sum, it may be true that citizens need to consume challenging information regularly, and to think in a certain way, given *some* normative views on democracy. But, this does not mean that this is the only way for democracies to work, empirically speaking, or that democracy will stop working if we do not adhere to these principles.

9. It is not clear what a filter bubble is

Last, but certainly not least, is the definition and description of a filter bubble. To reiterate from what was said earlier, a filter bubble is a “personal ecosystem of information that’s been catered by these algorithms to who they think you are” (The Daily Dish, 2010: para. 1), or a “unique universe of information for each of us” (Pariser, 2011: 9). Although there is nothing inherently wrong with definitions or descriptions of this kind, they nonetheless lack clarity (for a similar point, see Bruns, 2019a).

This article has extensively cited Pariser’s own claims regarding filter bubbles, and that is for an important philosophical reason. When filter bubbles are discussed, there is a risk of confusing two arguments: the first strong but also trivial and uninteresting, and the second, weak and speculative but also the most interesting. This general phenomenon has been called the Motte and Bailey doctrine (Shackel, 2005), and it can be explained something like this. A person is making a big and broad claim using vague terms that raise many ambiguous implications (the Bailey). When that claim is challenged, however, the person retreats and instead uses strict terms, obvious truths, and rigorous reasoning that no one could disagree with (the Motte). When the Motte is successfully defended, the person goes back to claiming the Bailey. This is a problem because the Motte and the Bailey are two different arguments.

The filter bubble thesis put forward by Pariser warns about the negative consequences, such as political polarisation, of technology for democracy (the Bailey). When filter bubbles are challenged, however, one can always retreat to the true but rather trivial claim that a site gives two different results to two users who search for the same information (the Motte). In other words, the retreat discusses quite factual matters, but is nonetheless far from the original and more interesting claim of the negative democratic consequences. This is why the discussion of filter bubbles at the technological level (e.g., whether personalisation leads to different information) should be identified as separate from the discussion of filter bubbles at the societal level (e.g., whether personalisation insulates people from challenging information and increases political polarisation in society).

It should also be noted that there is a trade-off between extensively citing the original source (Pariser) and citing new empirical and theoretical literature that might have pushed the research in another direction beyond Pariser. I have tried to stay as close to

the source as possible in this article.⁷ This can be risky, since today's technology could quickly become obsolete. However, one should also consider the opposite side of the problem. Constantly adapting a theory to the newest empirical data means that the theory can never be criticised, since there is no agreed theory at any given point in time; the theory is much like a floating signifier, where the point of the theory becomes to summarise the current state of the empirical evidence, rather than to predict new phenomena in the world that should be revealed by empirical evidence (Lakatos, 1999).

Consequently, there is a need to use more precise terms and to interlink concrete causal predictions in a nomological network if filter bubbles are to be made into a coherent theory. As enumerated at the beginning of this article, there have been several predictions of what will happen to individuals as a result of algorithmic personalisation (decreased curiosity, increased confirmation bias, etc.), which builds the foundation for filter bubbles at the societal level. Are we to say that filter bubbles emerge when *all* of these predictions come true, or is it sufficient that only *one* of them does? Which are necessary and which are sufficient? What is the relationship between them? These questions are not answered by summarising more empirical studies and invoking filter bubbles as an explanation, since filter bubbles would then only become a circular re-statement of the empirical data.

Conclusions

Should we be worried that social networking sites and other sites provide us with more and more personalised information that increasingly leads us to insulate ourselves from viewpoints other than those we already have? And, in the worst-case scenario, that we become excluded from the whole political process? These predictions do not seem to be very well grounded in what we already know about humans and technology, as previous reviews on filter bubbles have pointed out (e.g., Bruns, 2019a; Zuiderveen Borgesius et al., 2016).

It is easy to point to technology and to say something provocative about its power to destabilise democracy. It is much more difficult to substantiate that claim, especially when humans can interfere with the technology – in both good and bad ways. For example, it may be true that personalisation algorithms lead to filter bubbles from the purely technical standpoint of feedback loops (as happens when a microphone is too close to a loudspeaker). The filter bubble thesis nonetheless goes far beyond the technology and extrapolates with additional claims about the long-term effects on humans, society, and ultimately, democracy. These additional claims are the most interesting, but also the most unsubstantiated, and the evidence often goes in the opposite direction than to what is predicted by the filter bubble thesis (see, e.g., Barberá, 2015; Beam et al., 2018; Boxell et al., 2017; Bruns, 2019a; Davies, 2018; Dubois & Blank, 2018; Möller et al., 2018; Nelson & Webster, 2017; Zuiderveen Borgesius et al., 2016). In short, a filter bubble “is a misunderstanding especially about the power of technology – and in particular, of algorithms – over human communication” (Bruns, 2019a: 93).

This critical review, with its focus on counterarguments and the assumptions of filter bubbles, emphasises that the filter bubble thesis often posits a special kind of political human who has opinions that are strong, but at the same time highly malleable, according to the information given. It is an inherent paradox that people have an active

agency when they select content, but are passive receivers once they are exposed to the algorithmically curated content recommended to them (see Figure 1). The filter bubble thesis therefore resurrects and brings together several more or less outdated theories: technological determinism (Davies, 2018), strong behaviourism (Graham, 2017), and the hypodermic needle model (Jensen & Rosengren, 1990). While it is certainly true that the individual and the environment affect each other reciprocally, this reciprocity is comparatively weak, indirect, and temporary. The filter bubble thesis, on the other hand, posits that these effects are strong, permanent, and inescapable, since they can only be exacerbated over time – not alleviated – with detrimental consequences for society.

The metaphor of an enclosing bubble that one cannot escape is powerful in its persuasiveness and simplicity. However, the metaphor is misleading since it helps solidify the bubble as something singular. As pointed out previously in this article, people have multiple competing preferences that can be in conflict with each other (e.g., studying versus watching a sitcom). People also have overlapping social networks in dimensions other than political ideology (e.g., people may share the same workplace but disagree ideologically). For these reasons, there are two seemingly contradictory accounts: people can now – unlike in times before the Internet existed, be exposed to a large amount of challenging information online – while at the same time, they can also be relatively much more exposed to supporting information (see Malinas & Bigelow, 2016). This is not a contradiction, because people can have different motivations for media use.

The notion that only *some* people are insulated in filter bubbles (or that filter bubbles merely decrease the amount of challenging information by a small amount) might seem to be a tempting compromise, or a more nuanced approach to filter bubbles. But why give so much emphasis to technology and its negative democratic outcomes in the first place, if technology only plays a minor role? Does this thesis mean that technology (or, more precisely, personalisation algorithms) is simply a moderator or mediator of selective exposure?⁸ Let us say, for the sake of argument, that only some individuals are actually insulated in filter bubbles. What, then, is the democratic problem if there are only a few of them?

This article ends with a suggestion for future research. There are sufficient grounds to study different aspects highlighted by the filter bubble thesis. For example, the effects of personalisation algorithms on content diversity are highly relevant, as more content is filtered through computer systems than through human editors. These can be thoroughly studied without invoking the theoretical baggage of filter bubbles.

To conclude, the causes (and consequences) of increased choice is a recurring theme throughout media history. When the printing press was introduced to the masses around the sixteenth century, Erasmus of Rotterdam warned that all the new books would distract people from what was really important in life – reading Aristotle (Erasmus & Barker, 2001). Similarly, the previous literature provides some grounds for caution about the supposedly negative effects of new media:

For each new medium, there has been widespread fear that its effects might be deleterious, especially to supposedly weak minds, such as those of children, women and uneducated people. “Moral panics” of this type accompanied the introduction of film, comics, TV and video. (Jensen & Rosengren, 1990: 209)

Perhaps we will look back at the debate about filter bubbles in the same way?

Notes

1. Although humans can change their behaviour when confronted with a new situation, this does not necessarily mean that they are easily (or rapidly) shaped by the environment: “Human behavior is characterized by situational variance around discernible central tendencies. Consequently, identification of intraindividual differences in behavior across situations does not mean that the possible significance of traits should be dismissed” (Mondak, 2010: 10).
2. See https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles
3. Although there are many types of polarisation, such as divergent (e.g., two persons moving further away from each other in a belief) and sorting (e.g., two persons becoming more consistent in several of their beliefs) – it is not clear what type of polarisation is referred to in the idea of filter bubbles.
4. For instance, in a syllogism the premises can be false, or the inference can be invalid, which thereby undermines the soundness of a deductive argument (or undermines the strength of an inductive or abductive argument). An argument can also be shown to be unsound when it leads to an absurd conclusion (e.g., a conclusion that contradicts a claim known to be true). One can then criticise the argument by *reductio ad absurdum*.
5. The apparent paradox between a positive trend at the micro level (large filter bubbles) and a negative trend at the macro level (the opposite of filter bubbles) is a nice example of the Simpson’s paradox (Malinas & Bigelow, 2016).
6. Also note that the meta-analytic effect size for selective exposure has a broad range, from a Cohen’s *d* of -1.5 to +3.3 (Hart et al., 2009), indicating that there are also situations in which people select challenging information (minus sign) as well as situations in which they select supporting information (plus sign).
7. Even if Pariser’s book is journalistic in nature, there is nothing journalistic about a causal claim or a prediction. Instead, a causal claim or prediction can be more or less substantiated.
8. Compare this with contextualism: the idea that all theories are true, at least under certain circumstances (Perry, 1988).

References

- Althaus, S. L. (2006). False starts, dead ends, and new opportunities in public opinion research. *Critical Review*, 18(1–3), 75–104. <https://doi.org/10.1080/08913810608443651>
- Arceneaux, K., & Johnson, M. (2013). *Changing minds or changing channels? Partisan news in an age of choice*. Chicago: University of Chicago Press.
- Arceneaux, K., & Johnson, M. (2015). More a symptom than a cause: Polarization and partisan news media in America. In J. A. Thurber, & A. Yoshinaka (Eds.), *American Gridlock: The sources, character, and impact of political polarization* (pp. 309–336). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781316287002.016>
- 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>
- Barberá, P. (2015, September 3–6). *How social media reduces mass political polarization. Evidence from Germany, Spain, and the U.S.* [Conference presentation]. Annual Meeting of the 2015 American Political Science Association, San Francisco, California. http://pablobarbera.com/static/barbera_polarization_APSA.pdf
- 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>
- Beam, M. A., Hutchens, M. J., & Hmielowski, J. D. (2018). Facebook news and (de)polarization: Reinforcing spirals in the 2016 US election. *Information, Communication & Society*, 21(7), 940–958. <https://doi.org/10.1080/1369118X.2018.1444783>
- Bos, L., Kruikemeier, S., & de Vreese, C. H. (2016). Nation binding: How public service broadcasting mitigates political selective exposure. *PLoS ONE*, 11(5), e0155112. <https://doi.org/10.1371/journal.pone.0155112>
- Boxell, L., Gentzkow, M., & Shapiro, J. M. (2017). Greater Internet use is not associated with faster growth in political polarization among US demographic groups. *Proceedings of the National Academy of Sciences*, 114(40), 10612–10617. <https://doi.org/10.1073/pnas.1706588114>
- Boxell, L., Gentzkow, M., & Shapiro, J. M. (2020). *Cross-country trends in affective polarization* (No. w26669). Cambridge, Massachusetts: National Bureau of Economic Research. <https://doi.org/10.3386/w26669>
- Bozdag, E., & van den Hoven, J. (2015). Breaking the filter bubble: Democracy and design. *Ethics and Information Technology*, 17, 249–265. <https://doi.org/10.1007/s10676-015-9380-y>
- Bruns, A. (2019a). *Are filter bubbles real?* Oxford: Polity Press.
- Bruns, A. (2019b). Filter bubble. *Internet Policy Review*, 8(4), 1–14. <https://doi.org/10.14763/2019.4.1426>
- Chitra, U., & Musco, C. (2020). Analyzing the impact of filter bubbles on social network polarization. *Proceedings of the 13th International Conference on Web Search and Data Mining* (pp. 115–123). Houston, Texas: Association for Computing Machinery. <https://doi.org/10.1145/3336191.3371825>

- Choi, J., & Lee, J. K. (2015). Investigating the effects of news sharing and political interest on social media network heterogeneity. *Computers in Human Behavior, 44*, 258–266. <https://doi.org/10.1016/j.chb.2014.11.029>
- Conover, M. D., Ratkiewicz, J., Francisco, M., Goncalves, B., Menczer, F., & Flammini, A. (2011, July 17–21). *Political polarization on Twitter* [Conference presentation]. Fifth International AAAI Conference on Weblogs and Social Media, Barcelona, Spain. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2847/3275>
- Cotton, J. L. (1985). Cognitive dissonance in selective exposure. In D. Zillmann, & J. Bryant (Eds.), *Selective exposure to communication* (pp. 11–33). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Cowan, S. K., & Baldassarri, D. (2018). “It could turn ugly”: Selective disclosure of attitudes in political discussion networks. *Social Networks, 52*, 1–17. <https://doi.org/10.1016/j.socnet.2017.04.002>
- Crockett, M. J. (2017). Moral outrage in the digital age. *Nature Human Behaviour, 1*, 769–771. <https://doi.org/10.1038/s41562-017-0213-3>
- D’Alessio, D., & Allen, M. (2006). The selective exposure hypothesis and media choice processes. In R. W. Preiss (Ed.), *Mass media effects research: Advances through meta-analysis* (pp. 103–118). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Davies, H. C. (2018). Redefining filter bubbles as (escapable) socio-technical recursion. *Sociological Research Online, 23*(3), 637–654. <https://doi.org/10.1177/1360780418763824>
- Dubois, E., & Blank, G. (2018). The echo chamber is overstated: The moderating effect of political interest and diverse media. *Information, Communication & Society, 21*(5), 1–17. <https://doi.org/10.1080/1369118X.2018.1428656>
- Duggan, M., & Smith, A. (2016). *The political environment on social media*. Washington, D.C.: Pew Research Center. <https://www.pewinternet.org/2016/10/25/the-political-environment-on-social-media/>
- Erasmus, D., & Barker, W. W. (2001). *The adages of Erasmus*. Toronto: University of Toronto Press.
- Eveland, W. P., Jr., & Hively, M. H. (2009). Political discussion frequency, network size, and heterogeneity of discussion as predictors of political knowledge and participation. *Journal of Communication, 59*(2), 205–224. <https://doi.org/10.1111/j.1460-2466.2009.01412.x>
- Fischer, P., Lea, S., Kastenmüller, A., Greitemeyer, T., Fischer, J., & Frey, D. (2011). The process of selective exposure: Why confirmatory information search weakens over time. *Organizational behavior and human decision processes, 114*(1), 37–48. <https://doi.org/10.1016/j.obhdp.2010.09.001>
- 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>
- Fletcher, R., & Nielsen, R. K. (2017a). Are news audiences increasingly fragmented? A cross-national comparative analysis of cross-platform news audience fragmentation and duplication. *Journal of Communication, 67*(4), 476–498. <https://doi.org/10.1111/jcom.12315>
- Fletcher, R., & Nielsen, R. K. (2017b). Are people incidentally exposed to news on social media? A comparative analysis. *New Media & Society, 20*(7), 2450–2468. <https://doi.org/10.1177/1461444817724170>
- Frey, D. (1986). Recent research on selective exposure to information. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (vol.19, pp. 41–80). San Diego, California: Academic Press.
- Garrett, R. K. (2009). Politically motivated reinforcement seeking: Reframing the selective exposure debate. *Journal of Communication, 59*(4), 676–699. <https://doi.org/10.1111/j.1460-2466.2009.01452.x>
- Genzkow, 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>
- Graham, G. (2017). Behaviorism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2017 ed.). <https://plato.stanford.edu/archives/spr2017/entries/behaviorism/>
- Guess, A. M., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. *Nature Human Behaviour, 4*, 472–480. <https://doi.org/10.1038/s41562-020-0833-x>
- Hanitzsch, T., Van Dalen, A., & Steindl, N. (2018). Caught in the nexus: A comparative and longitudinal analysis of public trust in the press. *The International Journal of Press/Politics, 23*(1), 3–23. <https://doi.org/10.1177/1940161217740695>
- Hansson, S. O., & Grüne-Yanoff, T. (2012). Preferences. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2012 ed.). <https://plato.stanford.edu/archives/win2012/entries/preferences/>
- 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>
- Hart, W., Richardson, K., Tortoriello, G. K., & Earl, A. (2019). “You are what you read:” Is selective exposure a way people tell us who they are? *British Journal of Psychology, 111*(3), 417–442. <https://doi.org/10.1111/bjop.12414>
- Hill, R. A., & Dunbar, R. I. M. (2003). Social network size in humans. *Human Nature, 14*(1), 53–72. <https://doi.org/10.1007/s12110-003-1016-y>

- Jensen, K. B., & Rosengren, K. E. (1990). Five traditions in search of the audience. *European Journal of Communication*, 5, 207–238. <https://doi.org/10.1177/0267323190005002005>
- Kahan, D. M., Landrum, A., Carpenter, K., Helft, L., & Jamieson, K. H. (2017). Science curiosity and political information processing. *Political Psychology*, 38, 179–199. <https://doi.org/10.1111/pops.12396>
- Kearney, M. W. (2019). Analyzing change in network polarization. *New Media & Society*, 21(6), 1380–1402. <https://doi.org/10.1177/1461444818822813>
- Kim, M.-S., & Hunter, J. E. (1993). Relationships among attitudes, behavioral intentions, and behavior: A meta-analysis of past research, part 2. *Communication Research*, 20(3), 331–364. <https://doi.org/10.1177/009365093020003001>
- Knobloch-Westerwick, S. (2014). *Choice and preference in media use: Advances in selective exposure theory and research*. New York: Routledge. <https://doi.org/10.4324/9781315771359>
- Lakatos, I. (1999). *The methodology of scientific research programmes*. Philosophical papers volume 1. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511621123>
- Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1948). *The people's choice: How the voter makes up his mind in a presidential campaign*. New York: Columbia University Press.
- Lee, J. K., Choi, J., Kim, C., & Kim, Y. (2014). Social media, network heterogeneity, and opinion polarization. *Journal of Communication*, 64(4), 702–722. <https://doi.org/10.1111/jcom.12077>
- Malinas, G., & Bigelow, J. (2016). Simpson's paradox. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2016 ed.). <https://plato.stanford.edu/archives/fall2016/entries/paradox-simpson/>
- Matsa, K. E., Silver, L., Shearer, E., & Walker, M. (2018, October). A minority in all eight countries say the news they see on social media reflects their own political views. *Pew Research Center's journalism project*. https://www.journalism.org/2018/10/30/younger-europeans-are-far-more-likely-to-get-news-from-social-media/pj_2018-10-30_europe-age_0-08/
- 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>
- Mondak, J. J. (2010). *Personality and the foundations of political behavior*. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511761515>
- Munson, S. A., & Resnick, P. (2010). Presenting diverse political opinions: How and how much. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1457–1466). New York: ACM. <https://doi.org/10.1145/1753326.1753543>
- Mutz, D. C., & Mondak, J. J. (2006). The workplace as a context for cross-cutting political discourse. *The Journal of Politics*, 68(1), 140–155. <https://doi.org/10.1111/j.1468-2508.2006.00376.x>
- Möller, J., Trilling, D., Helberger, N., & van Es, B. (2018). Do not blame it on the algorithm: An empirical assessment of multiple recommender systems and their impact on content diversity. *Information, Communication & Society*, 21(7), 959–977. <https://doi.org/10.1080/1369118X.2018.1444076>
- Negroponte, N. (1995). *Being digital*. New York: Knopf.
- Nelson, J. L., & Webster, J. G. (2017). The myth of partisan selective exposure: A portrait of the online political news audience. *Social Media + Society*, 3(3), 1–13. <https://doi.org/10.1177/2056305117729314>
- Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A. L., & Nielsen, R. K. (2017). *Reuters Institute digital news report 2017*. Oxford: Reuters Institute for the Study of Journalism, University of Oxford.
- Newton, C. (2016, November 16). The author of The Filter Bubble on how fake news is eroding trust in journalism. *The Verge*. <https://www.theverge.com/2016/11/16/13653026/filter-bubble-facebook-election-eli-pariser-interview>
- Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. New York: Penguin Press.
- Perry, D. K. (1988). Implications of a contextualist approach to media-effects research. *Communication Research*, 15(3), 246–264. <https://doi.org/10.1177/009365088015003002>
- Prior, M. (2013). Media and political polarization. *Annual Review of Political Science*, 16(1), 101–127. <https://doi.org/10.1146/annurev-polisci-100711-135242>
- Settanni, M., Azucar, D., & Marengo, D. (2018). Predicting individual characteristics from digital traces on social media: A meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, 21(4), 217–228. <https://doi.org/10.1089/cyber.2017.0384>
- Shackel, N. (2005). The vacuity of postmodernist methodology. *Metaphilosophy*, 36(3), 295–320. <https://doi.org/10.1111/j.1467-9973.2005.00370.x>
- Skoric, M. M., Zhu, Q., & Lin, J.-H. T. (2018). What predicts selective avoidance on social media? A study of political unfriending in Hong Kong and Taiwan. *American Behavioral Scientist*, 62(8), 1097–1115. <https://doi.org/10.1177/0002764218764251>
- Smith, S. M., Fabrigar, L. R., & Norris, M. E. (2008). Reflecting on six decades of selective exposure research: Progress, challenges, and opportunities. *Social and Personality Psychology Compass*, 2(1), 464–493. <https://doi.org/10.1111/j.1751-9004.2007.00060.x>

- Stroud, N. J. (2010). Polarization and partisan selective exposure. *Journal of Communication*, 60(3), 556–576. <https://doi.org/10.1111/j.1460-2466.2010.01497.x>
- Sunstein, C. R. (2001). *Republic.Com*. Princeton, New Jersey: Princeton University Press.
- Süllflow, M., Schäfer, S., & Winter, S. (2018). Selective attention in the news feed: An eye-tracking study on the perception and selection of political news posts on Facebook. *New Media & Society*, 21(1), 168–190. <https://doi.org/10.1177/1461444818791520>
- The Daily Dish. (2010, October). The filter bubble. *The Atlantic*. <https://www.theatlantic.com/daily-dish/archive/2010/10/the-filter-bubble/181427/>
- Trilling, D., Klingeren, M. van, & Tsifti, Y. (2017). Selective exposure, political polarization, and possible mediators: Evidence from the Netherlands. *International Journal of Public Opinion Research*, 29(2), 189–213. <https://doi.org/10.1093/ijpor/edw003>
- Tsifti, Y., Boomgaarden, H. G., Strömbäck, J., Vliegthart, R., Damstra, A., & Lindgren, E. (2020). Causes and consequences of mainstream media dissemination of fake news: Literature review and synthesis. *Annals of the International Communication Association*, 44(2), 157–173. <https://doi.org/10.1080/23808985.2020.1759443>
- Urman, A. (2019). Context matters: Political polarization on Twitter from a comparative perspective. *Media, Culture & Society*, 42(6), 857–879. <https://doi.org/10.1177/0163443719876541>
- Vinnova. (2017, October 26). Unikt samarbete ska motverka falska nyheter inför valet [A unique collaboration will counteract fake news prior to the election]. [press release]. *Vinnova*. <https://www.vinnova.se/nyheter/2017/10/unikt-samarbete-ska-motverka-falska-nyheter-infor-valet/>
- Vraga, E., Bode, L., & Troller-Renfree, S. (2016). Beyond self-reports: Using eye tracking to measure topic and style differences in attention to social media content. *Communication Methods and Measures*, 10(2–3), 149–164. <https://doi.org/10.1080/19312458.2016.1150443>
- Webster, J. G. (2017). Three myths of digital media. *Convergence*, 23(4), 352–361. <https://doi.org/10.1177/1354856517700385>
- Weeks, B. E., Ksiazek, T. B., & Holbert, R. L. (2016). Partisan enclaves or shared media experiences? A network approach to understanding citizens' political news environments. *Journal of Broadcasting & Electronic Media*, 60(2), 248–268. <https://doi.org/10.1080/08838151.2016.1164170>
- Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review*, 5(1). <https://doi.org/10.14763/2016.1.401>