Does Social Media Use Promote Political Mass Polarization?

A Structured Literature Review[1]

Katharina Ludwig & Philipp Müller

Abstract

In past years, a large amount of research was conducted to determine whether the use of social media causes political polarization. This research field, however, lacks clear terminological definitions and concepts such as fragmentation and selective exposure are often imprecisely equated with political polarization, which may explain the widespread assumption that social media cause political polarization. With this article, we aim to unravel conceptual confusion and offer distinct definitions of affective, ideological, and partisan polarization. We conducted a structured literature review of 88 studies addressing the potential effects of social media use on polarization. We find the operationalization of relevant concepts to differ significantly between research projects, making the comparability of results difficult and possibly contributing to inconsistent findings. No clear evidence is found to support the generalized perception of strong polarization effects through the use of social media. Implications for future research are proposed.

Since the internet's earliest days, theorists have voiced concerns about the risks of fragmentation and polarization effects (e.g., Dahlberg, 2007; Papacharissi, 2002; Sunstein, 2001). These concerns are amplified by the emergence and growth of social media platforms and algorithmic content-selection mechanisms and their growing importance in political information exposure (Bakshy, Messing, & Adamic, 2015). The so-called fragmentation thesis expresses the idea that discussions about politics are taking place in insulated groups, separated along party or ideological lines, with little or no contact between groups (Bright, 2018). This implies that people are captured in self-selected "echo chambers" (Sunstein, 2001) or algorith-

1 This research was supported by a grant from Baden-Württemberg Stiftung within the research program Responsible Artificial Intelligence.

mically induced "filter bubbles" (Pariser, 2011), communicating only with those who have similar ideological viewpoints and, thus, being exposed only to opinion-confirming information. As empirical studies have shown that, in social media, the fragmentation thesis is, at least partly, in place (e.g., Bright, 2018), theorists worry about the implications for democracy as the democratic formation of a collective will via deliberation requires citizens to be exposed to a range of diverse viewpoints (Gentzkow & Shapiro, 2010). If people, instead, are exposed only to like-minded content and, consequently, constantly reinforced in their beliefs, political polarization and societal disintegration might be the outcomes (e.g., Warner, 2010; Arceneaux & Johnson, 2010).

While the argument that growing segments of the electorate that use social media platforms to become informed might initiate such processes seems convincing at first (and is continuously popularized in public discourse), it remains largely unclear whether this notion is supported by empirical research. The main aim of this chapter is, therefore, to systematically review existing findings on polarization through social media usage and to disentangle different causal mechanisms of social-media-induced effects on polarization effects found in the literature. Before being able to do so, however, we must first clarify extant conceptual confusions that are caused by the frequent interchangeable use of terms such as "fragmentation," "group polarization," or "political polarization" in the literature. In the first part of this chapter, we propose a conceptual framework to disentangle these different types of polarization as well as each one's operationalization.

Moreover, our aim is not to determine whether political polarization *is represented* in social media environments but whether social media environments are *causing political polarization* and if so, to identify those exact mechanisms that play central roles in this process. In other words, if polarization can be documented within social media environments, does this mean that social media technologies can be held responsible for its occurrence? In particular, we are interested in clarifying whether algorithmic selection mechanisms or individual user decisions or predispositions affect political polarization processes. Furthermore, we perform an analysis to identify structural differences between different country contexts and researchers' methodological decisions. As the existing research on this topic lacks clear definitions and distinctions between concepts, operationalization, and methodologies, this is a necessary and important endeavor.

Ours is not the first attempt to provide a systematic overview of the questions discussed up to this point. As we were conducting this study, two literature reviews were published that follow a similar perspective;

they deal with (a.) group polarization in online discussions (Iandoli, Primario, & Zollo, 2021) and (b.) the role of (social) media use in political polarization (Kubin & von Sikorski, 2021). While we see a good amount of merit in these two studies, we believe that at least two arguments justify publishing a third literature review that follows a similar question. First, literature reviews can be seen as meta-empirical research that draws conclusions from a broad overview of the empirical observations of others. As such, the same argument that can be made for single empirical studies has to be made for literature reviews of empirical research; that is, empirical science is based on an accumulation of evidence and, therefore, one research team's observations and interpretations can never be sufficient to draw generalizable conclusions. That being said, reconsidering a question that has previously been investigated by others inevitably adds value to the state of knowledge, if it is only to provide reassurance that previous conclusions can be substantiated.

Second, we see specific limitations of the previous literature reviews that are addressed by our study. More specifically, Iandoli et al. (2021) offer a broad overview of all kinds of research revolving around the themes of social media and polarization. This breadth of focus necessarily restricts the review's ability to answer specific questions precisely. The review study looks at significantly different types of fragmentation and polarization processes without conceptually disentangling them. Furthermore, the review does not focus solely on social media *effects* on polarization but considers, in addition, manifestations of polarization on social media platforms as well as "other online conversational platforms" (p. 1). The second review, by Kubin and von Sikorski (2021), approximates our study in terms of its focus and procedure. Yet the two studies differ in nuance, and most importantly, their corpora vary for several reasons. For instance, we excluded several studies that, in our reading, used the term "polarization" but, instead, investigated what we would call "fragmentation" processes. Contrary to existing literature reviews, we categorize the type of polarization investigated in a study based on the operationalization used rather than on the labeling used by a study's authors. We argue that this process is necessary to achieve comparable results in light of the conceptual vagueness of the field and significantly large discrepancies between studies in terms of labeling and operationalization. At the same time, our literature search resulted in a larger number of studies indicating depolarizing effects of social media use than the review by Kubin and von Sikorski (2021). This leads us to question the conclusion that there is "agreement across studies that social media, in a variety of contexts, can exacerbate both ideological and affective political polarization" (Kubin & von Sikorski, 2021, p. 196).

In the following, the concept of political polarization and its different dimensions will be defined and subsequently distinguished from the concept of fragmentation. Then, we briefly discuss the origins and consequences of political mass polarization and the role that social media technologies might play in this context. Finally, we delve into a systematic review of empirical evidence about social media effects on polarization

Political Mass Polarization: Concept and Overview

Research on political sociology, particularly from the United States, has carved out political polarization as one of the major factors affecting societal and political processes in recent decades (e.g., Fiorina & Abrams, 2008; Baldassarri & Gelman, 2008). As pointed out by DiMaggio et al. (1996), there are two different ways in which time can be considered when defining polarization as a concept: "Polarization is both a state and a process. Polarization as a state refers to the extent to which opinions on an issue are opposed in relation to some theoretical maximum. Polarization as a process refers to the increase in such opposition over time" (p.693). More recent scholarly definitions align with the perspective of polarization as a process—because determining a definite threshold at which topics or groups are polarized seems unrealistic (Fiorina & Abrams, 2008). For this research endeavor, we therefore adopt the definition by McCoy, Rahman, and Somer (2018), who conceptualized polarization as "a process whereby the normal multiplicity of differences in a society increasingly align along a single dimension" (p.16).

Forms and Measurement of Political Mass Polarization

This phenomenon is exactly what we have been witnessing in past decades, not just in the context of the US with its political landscape becoming steadily more polarized (e.g., McCarty, Poole & Rosenthal, 2006; Iyengar, Sood & Lelkes, 2012) and, at the same time, with growing animosities between the parties' electorates (e.g., Abramowitz & Sounders, 2008). These observations are also the main forms of political polarization that are traditionally distinguished: *elite polarization*, respectively *party polarization*, and *mass polarization*. Party polarization describes the polarization between the ruling party and the opposition party at the political system level (Baldassarri & Gelman, 2008). Mass polarization depicts a division

along party lines of the public's attitudes toward political topics, policies, politicians, or opposing political camps within the electorate (e.g., Fiorina & Abrams, 2008; McCarty et al., 2006; Layman, Carsey & Horowitz, 2006). Some researchers assume party polarization to be the main reason for mass polarization, as partisans align with their party's ideals and engage in behaviors that are, seemingly, in line with their party's objectives (e.g., Layman, Carsey, & Horowitz, 2006). Other scholars theorize that the opposite holds true, with party polarization resulting from the publics' separation in opposing camps (see, e.g., Fiorina et al., 2005). While this study focuses on the second form of polarization, the polarization of the electorate, it is important to keep in mind that the strength and particular forms of mass polarization within a society appear to be causally related to party polarization at the system level. This is particularly important to acknowledge for a literature review that attempts to integrate empirical findings from a diverse set of national contexts. Another important theoretical differentiation can be made between the above-mentioned mass polarization, measured on the level of individuals, for example, through surveys or experiments, and group polarization, measured on a group level through, for instance, network or content analyses. While mass polarization studies can make statements about individual polarization effects through the use of social media, studies on group polarization can identify superordinate polarization patterns at the group level.

Recent research has pinpointed the fact that different dimensions of mass polarization have to be disentangled. Some scholars argue that it appears as if US citizens, in particular, are increasingly agreeing on many political issue positions while, at the same time, the strength of partisan identifications and animosities between different political camps have profoundly massively increased (e.g., Baldassarri & Gelman, 2008; Iyengar, Sood, & Lelkes, 2012). Others argue that polarization in both dimensions is still on the rise (e.g., Abramowitz & Sounders, 2008; Abramowitz, 2010). To distinguish these two concepts, researchers have coined the terms *ideological polarization* (DiMaggio, Evans & Bryson, 1996) and *affective polarization* (Iyengar et al., 2019).

Affective and ideological polarization are both characterized by a separation of individuals of different political camps, typically from the ideological left and right, over policy differences (Webster & Abramowitz, 2017). In the case of *affective polarization*, this manifests in a strong liking for one's partisan party and a close attachment to it, accompanied by the simultaneous and equally strong dislike of the opposing party and preference for distance from it or its members. Affective polarization, therefore, is usually measured by surveys and experiments through a "feeling

thermometer" (Stroud, 2010) calculating the participants' warmth toward their preferred party or political camp minus their warmth toward an opposing camp to compare inter- or intra-individual polarity scores. Other modes of operationalization involve measures of trait ratings toward the different camps' partisans, asking respondents, for instance, to rate their intelligence, generosity, and character or asking respondents what aspects they like and dislike about political parties and their voters (e.g., Levendusky & Malhotra, 2016; Garrett et al., 2014). In addition, Iyengar and Westwood (2015) adapted the Implicit Association Test (IAT) to capture unconscious partisan bias. Similarly, studies dealing with group polarization, using for example network, content, and sentiment analysis and a combination of different content features, such as ingroup vs. outgroup references combined with sentiment or other features, e.g., expressions of anxiety, anger, and the use of profanity, can be used to measure affective polarization (Gruzd & Roy, 2014; Bliuc, Smith & Moynihan, 2020; Mentzer, Fallon, Prichard & Yates, 2020).

Such operationalization is easily used in dual-party systems but poses problems for multi-party systems, as coalitions in such political systems are formed temporarily (Sened, 1996) and are characterized by floating affinities and animosities between parties beyond ideology. Therefore, it is generally not possible to identify clear "counterparties" in such systems, which allow using the common affective polarization measures that are considered dyads of political camps. Nevertheless, it is possible to capture affective polarization in multi-party systems by calculating an index of like/dislike scores across different political parties (Wagner, 2020). However, to the best of our knowledge, no studies thus far have used this operationalization.

A simpler approach omits negative sentiment toward an opposing political camp and focuses instead on *partisan polarization*. Studies following this approach usually measure merely the degree of partisans' attachments to their political camps. That is, participants are typically asked about their party identity directly or asked to locate themselves on a left–right or liberal–conservative scale. Some studies have also used profile information to derive the political ideology of users, and network analyses additionally determine the partisanship of social media users through the co-following or co-retweet networks (Grover et al., 2019). This operationalization of political polarization, of course, reduces the concept's explanatory power as it considers only half of the affective polarization process. At the same time, it might be more appropriate to capture polarization dynamics in multi-party systems in which there is not always a clear bipolar relationship between opposing political camps. Another reason to include studies

based on partisan polarization in the present research is that this concept is used by several social-media-oriented polarization studies and, thereby, cannot be omitted from a literature review.

Ideological polarization, also referred to as "issue polarization" (Dylko et al., 2017) or "positional polarization" (Yarchi, Baden & Kligler-Vilenchik, 2020), is measured similarly to affective polarization. However, measures are based on issue stances or attitudes toward political topics such as climate change, health care, gay marriage, abortion laws, gun policy, or immigration (e.g., Bail et al. 2018; Cho et al. 2018). Commonly in surveys and experiments, attitudes about polarized or non-polarized topics from both opposing political camps are operationalized as several items, and the aggregation of agreement or disagreement with these statements by the participants results in a polarity score that leans toward, for example, rather liberal or conservative attitudes. However, the measurement of ideological polarization is not limited to surveys. Content analyses can be used to investigate users' issue stances voiced in social media posts or expressions of sentiment toward a particular topic (e.g., Yardi & Boyd, 2010; Yarchi, Baden, & Kligler-Vilenchik, 2020). Another way to measure ideological polarization in social media networks is the so-called modularity approach (e.g., Del Vicario et al., 2017; Zollo, 2019). Here, for example, the balance of a user's likes on social media posts or pages confirming or opposing an issue position is calculated. It is then interpreted as an estimator for the respective user's ideological position on the specific issue (e.g., Vicario et al., 2017).

Beyond that, the literature on social-media-related polarization effects includes a large body of research that applies network analysis methods. Most of this research uses interaction networks, retweet networks, or post-sharing networks as indicators of polarized communities, which are bound by a shared attitude toward a topic. Notably, these analytical network approaches typically do not include negative feedback (such as dislikes), whereby only half of the operationalization of polarization is achieved. Therefore, many network analyses are complemented by additional data, for example from sentiment analyses or external opinion polls. Conceptually, this line of research cannot be clearly allocated to either affective or ideological polarization (even though this may be true for particular studies). This is because these studies usually do not measure individuals' attitudes or feelings toward political camps or issues but, rather, interaction patterns at the group level. These patterns may, of course, mirror the group members' levels of affective or ideological polarization, but they are, at best, coarse indicators for affective or ideological polarization. Consequently, network analyses dealing with social-media-related polariza-

tion apply a variety of labels such as "group polarization" (Yardi & boyd, 2010), "user polarization" (Bessi et al., 2016), "information polarization" (Usui, Yoshida, & Torium, 2013) or "online polarization" (Bliuc, Smith, & Moynihan, 2020). Conceptually, such approaches seem to draw from the idea of fragmentation as well as the above-defined understanding of political mass polarization.

Therefore, before turning to the role of social media technologies as potential drivers of polarization, we need to consider one of the major underlying facilitators of political polarization that is frequently confused with the latter: the phenomenon of political fragmentation.

Political Mass Polarization and Fragmentation

Broadly stated, a society or a network is fragmented if it is separated into or consisting of several parts. In other words, the more fragmented a society or network is, the more divisions between groups can be found (Bright, 2018). On a societal level, this dynamic has been observed in recent years especially in the US context (Arceneaux & Johnson, 2010). Parallel to polarization, elites, parties, media, and societies as a whole can be fragmented. As political fragmentation is accompanied by decreased contact between the fragmented groups, it can reduce group members' abilities to engage in perspective-taking with regard to outgroup individuals. This, in turn, may lead to distancing between social groups or may even promote group-related hostility and, thereby, result ultimately in political mass polarization (Arceneaux & Johnson, 2010).

Empirically, fragmentation is usually assessed at the group level where the degree of social homophily within groups (e.g., McPherson, Smith-Lovin, & Cook, 2001) and group seclusiveness (Bright, 2018) are typical indicators. From a communication perspective, this includes the degree of exposure to diverse political information sources. The technological developments of the past decades, such as the expansion in the numbers of radio and TV stations as well as of newspapers and magazines, have led to broad accessibility of news content (e.g., Arceneaux & Johnson, 2010). This development climaxed in the evolution of the internet as humanity's central communication tool. However, the wide diversification of potential news sources comes with the increased likelihood of decreasing the overlap between the various news repertoires of different members of a society, which in turn makes the fragmentation of information exposure more likely. However, it is important to note that fragmentation research aims at patterns of high social homophily within - and low interaction between

- groups, while mass polarization (as defined above) is typically studied by looking at how the separation of political camps within a society is reflected in individual persons' cognitions and emotions. Thus, the unit of analysis for fragmentation research is group composition and group-level behaviors, while for polarization research, it is individuals' group-oriented cognitions and emotions.

Similarly, political mass polarization can be seen as both a potential driver and outcome of fragmentation processes (Arceneaux & Johnson, 2010). As fragmented communities tend to narrow the scope of available information and reinforce existing beliefs, individual viewpoints might move farther away from more moderate attitudes and toward more extreme ones, and the differences and distance between ideological viewpoints may, in turn, grow (Arceneaux & Johnson, 2010). Of course, the opposite causal pathway almost certainly occurs at the same time, with mass polarization leading to a fragmented social landscape. As a result, mass polarization and fragmentation are mutually dependent. However, they do not refer to the same concept and, therefore, should not be equated in scientific research. Fragmentation and mass polarization are different processes and have different underlying mechanisms. Notably, fragmentation does not necessarily lead to political polarization, but it provides fertile ground for polarization.

Origins and Consequences of Political Mass Polarization

The societal consequences of increasing mass polarization are manifold. Partisan polarization, for instance, appears to strongly affect social relationships. This goes as far as leaning toward hiring staff with congruent partisanship (Iyengar & Westwood, 2015), preferences for romantic relationships, and the selection of friends who are co-partisans (e.g., Huber & Malhotra, 2017; Pew Research Center, 2017; Bakshy, Messing, & Adamic, 2015), and extends even to families becoming increasingly ideologically homogeneous. In 2018, 80% of married couples agreed on party identification; for parents and children, the agreement was 75% (Iyengar, Konitzer, & Tedin, 2018). Furthermore, people prefer living in areas comprised mostly of fellow partisans (Gimpel & Hui 2015). Studies have also identified economic transactions being affected by co-partisanship, with, for example, taxi drivers in Ghana demanding higher prices from counter-partisans (Michelitch, 2015) and US American citizens being willing to pay almost double for a gift card sold by a co-partisan in contrast to one sold by a counter-partisan (McConnell et al., 2018). Polarization also has conse-

quences for political processes. Growing animosities between counter-partisans, for example, make it more difficult to reach consensus; they affect voting decisions (Bartels, 2000) and can lead to growing opinion radicalization (Baldassarri & Gelman, 2008) or even political violence (Jensen et al., 2012). Taking into consideration these various domains of societal life, which are in one way or another being affected by mass polarization, it does not seem overstated to argue that political mass polarization poses a serious threat to social cohesion at the structural level. Sociologically, this means that increasing mass polarization (as documented for a number of countries over the last several decades; see, for example, Boxell et al., 2020) has the potential to endanger the functioning of human coexistence within a society.

In light of these consequences, it is important to investigate the origins of increasing polarization. In the research landscape, three lines of argument are typically emphasized for this purpose: (1.) social-identity-based explanations, (2.) ideology-based explanations, and (3.) information-exposure-based explanations. Notably, while these three lines of reasoning can be distinguished, they are also intertwined in many respects and, therefore, have to be considered complementary rather than competing mechanisms.

The first line of argument, identity-based explanations, underscores how political parties or camps increasingly serve as donors of collective identity for partisans seeking positively charged social entities with which they can identify in order to gain a positive self-image. As a byproduct, this process is also deemed to facilitate outgroup prejudice (Brewer, 1999) and, thereby, increase affective polarization (Mason, 2016). Fundamental to this concept is partisan identity acquired at a young age and frequently expressed in recurring political campaigns. Consequently, partisans build a sense of group identity with their co-partisans that can become more or less central to their self-concept. While outgroup derogation is a potential consequence of all social-identity processes (Brewer, 1999), devaluing opposing partisan groups in a political context appears even more likely since different political camps are, by nature, in opposition to each other (Iyengar, Sood, & Lelkes, 2012).

The second line of argument, ideology-based explanations, asserts that political mass polarization occurs as a consequence of political parties' ideological disparities (Webster & Abramowitz, 2017). They assume that, if the ideological distance between the different parties of a political system grows, this will lead citizens to perceive candidates or parties as polarized. For partisans, this perception of ideological gap formation may induce an urge to reaffirm their own ideological beliefs and partisan identity and corroborate their rejection of diverging ideologies and identities (Ro-

gowski & Sutherland, 2016). However, rather than regarding ideological conflict as a unilateral cause for partisan polarization, there seem to be mutual interrelations between both factors (Lelkes, 2018).

Finally, the third line of argument, information-exposure-based explanations, suggests that exposure to one-sided political content strengthens partisan identities and ideological beliefs, thereby facilitating political mass polarization (Garrett et al., 2014; Lau et al., 2017). While this research is set in the context of traditional mass-media channels, it has been argued that the internet's high-choice media environment (van Aelst et al., 2017) might have, once again, increased media impact on polarization processes.

Social Media Use: A Driver of Political Mass Polarization?

Within the debate about the internet's role in increasing political mass polarization, social media platforms are a crucial factor. When these technologies emerged, their services were predominantly understood as allowing "individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system" (boyd & Ellison, 2007). In the field of political communication research especially, the focus of attention has since shifted from users' abilities to self-present and connect via social media to the content to which they are exposed on these platforms. The infamous "news feed" and its algorithm-driven content selection now play a prominent role in the debate (see, e.g., Bode, 2016). From a business perspective, the central goal of social media's platform architectures is to maximize the time users spend on a platform because this maximizes ad revenues (Cohen, 2018). To achieve this goal, it is often argued that social media platforms' algorithms apply a "more-of-the-same" logic: They identify users' individual content preferences by tracking user behavior within the platform ecosystem (and beyond) and then attempt to serve individual users a content diet that aligns perfectly with their needs and interests. That being said, it is frequently assumed that, in terms of political content, this means users are going to encounter mainly messages that fit their political interests and convictions on social media platforms (which might reinforce their existing attitudes and partisan identities).

This potential mechanism has been popularized in Eli Pariser's (2011) "filter bubble" metaphor, which assumes that algorithmic content selection on social media platforms ultimately promotes political polarization. This is frequently referred to alongside the "echo chamber" metaphor

offered by Cass Sunstein (2001). The latter argues that users' own content-selection choices in high-choice media environments (van Aelst et al., 2017) may lead to homogeneous information environments that could also contribute to mass polarization. However, empirical evidence on whether algorithmic content selection or users' own selection decisions produce such homogeneous information environments in online ecosystems is mixed at best (see, e.g., Bakshy et al., 2015; Bruns, 2019; Flaxman et al., 2016; Möller et al., 2018; Scharkow et al., 2020; Zuiderveen Borgesius et al., 2016), with the occurrence of "echo chambers" appearing somewhat more likely than the emergence of "filter bubbles" (Flaxman et al., 2016). Moreover, even if patterns of homogeneous information environments emerged on a larger scale within online ecosystems, whether or not these "echo chambers," "filter bubbles," or "rabbit holes" actually promoted political mass polarization would still be unclear. For instance, it might very well be that homogeneous information environments have calming instead of radicalizing effects on many individuals since they offer less irritation than exposure to cross-cutting messages (Bor & Petersen, 2021).

Moreover, the entire debate on social media environments potentially contributing to polarization seems somewhat limited to "filter bubble" and "echo chamber" perspectives. Yet various other features and modalities of social media use might also contribute to mass polarization, perhaps to an even greater extent; however, for the most part, these are left untouched in the debate. For instance, it could be argued that the overrepresentation of negative sentiment and hateful expressions of opinion on social media platforms might deepen cleavages between different political camps (Bor & Petersen, 2021; Harel et al., 2020). Or social media self-effects that occur if a person has (semi-)publicly made a political statement might contribute to a radicalization of that person's political convictions and identifications (Valkenburg, 2017). Therefore, the present literature review is not limited to the "filter bubble" or "echo chamber" perspectives but attempts instead to systematically disentangle what is known empirically about the different, potentially causal mechanisms between social media use and political mass polarization.

Procedure

Literature Selection

Articles for this literature review were selected from EBSCO's *Communication & Mass Media Complete* database as well as *Semantic Scholar*

and had to be published in the period between 2004 and May 2021. To include all relevant articles, we searched for different keyword combinations. We combined the terms "polarization"/"polarisation" with the keyword terms "filter bubble," "echo chamber" or "rabbit hole" and with "social media," in addition to the names of the most common social media platforms ("Twitter," "Facebook, "YouTube," "TikTok," "Instagram," "Reddit," "VKontakte," and "Weibo"). For *Semantic Scholar*, we restricted the search to the fields of sociology, psychology, political science, and computer science and the type of publication to journal articles and conference contributions. For *Communication & Mass Media Complete*, we confined the search to academic journals in the English language. This resulted in a list of roughly 300 articles each from the two databases.

Literature Categorization

After gathering the initial corpus of potentially relevant studies, several selection steps were performed to arrive at a final collection of studies of interest. First, as *Semantic Scholar* also includes preprints, we eliminated studies that were not yet published in peer-reviewed outlets (by the end of May 2021). In the next step, the more-specific eligibility for each publication was determined based on its title, abstract, and—in the case of uncertainty—a full-text read. We narrowed the corpus to a set of empirical articles that dealt explicitly with both social media platforms and political polarization; this meant that research looking at non-political polarization (such as gender or age polarization) was excluded. Likewise, the role of social media platforms had to be an operationalized variable as well. Therefore, either social-media-use variables had to be measured empirically; content had to be posted on social media platforms; relationships between social media users had to be analyzed, or the research had to be embedded in an experimental setting that included social media environments. Studies using social media or political polarization as mere interpretational concepts were eliminated. Furthermore, we excluded studies that relied fully on non-empirical data, such as simulations-based research. This selection step resulted in a total of $n = 88$ studies, for which the full texts were read and will be analyzed in the following.

To gain a better overview of the study results, we categorized them according to their operationalization of polarization: (a.) fragmentation studies, (b.) group polarization studies, and (c.) mass polarization studies.

Review of Studies

Fragmentation or Polarization?

Before going into detail concerning studies that have analyzed group polarization and mass polarization effects, we provide a brief insight into several studies we found through our literature research. We recognized that the operationalization of polarization in several studies was actually one of fragmentation. As previously described, fragmentation and its related concepts as homophily or information diversity are closely connected to polarization and may play a critical role in the polarization process. Nevertheless, a measurement of fragmentation aspects does not necessarily measure polarization (effects); however, almost half of the studies we gathered do not operationalize polarization per se but still frame their research endeavor in this way. This alone is an interesting observation that may help to disentangle the conceptual and operational confusion in this research area. Therefore, although they did not meet our previously defined criteria, we still decided to include a brief overview of these studies and their results. It is important to note, however, that our search does not include a full picture of fragmentation/homophily studies. We report only on studies that frame their research endeavor as a measure of polarization and, thus, were identified through our keyword search.

In this category, we found 41 studies. Their respective operationalization of polarization includes measures of network homophily (n = 19); the density of connections within a network (n = 6), for example, measured through modularity approaches; measures of content diversity (n = 5); and the application of community detection algorithms (n = 4), such as the random walk controversy (RWC). Additionally, we found studies (n = 9) that merely determined the number of partisan users on social media and compared that with poll or election results. These nine studies will not be discussed in greater detail as their results show simply that users on social media are as fragmented as the electorate and, therefore, constitute a reflection of the offline social world. All the aforementioned measures might yield results about polarization processes or effects when combined with other measurements, but used alone, these variants of operationalization cannot illustrate the full polarization process. This is because, as we have argued in the definitions section of this chapter, homophily, content diversity, or network structure alone are not sufficient indicators of political polarization.

It is noteworthy that of the 32 studies we consider as capturing fragmentation rather than polarization, 26 analyzed Twitter, whereas the actual

polarization studies consider a much more balanced variety of social media platforms. The majority of fragmentation studies found what is called "polarization" in their respective arguments (n = 16). Another 12 studies yielded mixed effects—for instance, that a retweet network is "polarized" whereas a mention network is not (Conover et al., 2011), that partisan users formed highly partisan networks, whereas moderate users did not (Kearney, 2019), or that partisanship was less dominant if users had many cross-stance relations (Lai et al., 2019). Additionally, one study found no effects (Garimella, Morales, Gionis, & Mathioudakis, 2017), and another identified a reduction in network homophily over time (Lee & Hahn, 2017). Furthermore, many studies in this category found that, before and during election periods, more fragmentation was present (e.g., Yang et al., 2017; Kearney, 2019; Lai et al., 2019).

The fact that many of these fragmentation studies claim to have documented "polarization" within social media environments might help to explain why the general perception that social media leads to political polarization is so widespread. This argument is further pronounced when comparing these findings with the more inconclusive results found in actual mass polarization studies (see the following sections). The studies discussed here help us observe processes of fragmentation and *potential* signs of political polarization on social media. However, they do not help us to clarify whether political mass polarization is actually enhanced by social media use. For this, we need to take a much closer, in-depth look at the evidence about group polarization and mass polarization effects caused by social media use.

Group Polarization

Analyses of group polarization allow the observation of group dynamics on a larger scale but not of the effects on single individuals. Group polarization occurs when, after participating or being exposed to a discussion or taking part in other group activities, group members are reinforced in their sense of belonging and, consequently, become more extreme in their ideological or affective positions in concordance with their group's collective position (Isenberg, 1986). The difference between this and mass polarization is that, in terms of group polarization, dynamics can be determined on a group level but not traced back to individual polarization processes and effects. These patterns can also be divided into ideological, affective, and partisan (de)polarization, but they should not be misinterpreted as describing *effects* of social media use.

In the category of group polarization, we found 18 studies published in 13 different journals plus five different conference proceedings (for a full overview, see Appendix 1). Of these 18 studies, 10 articles were published between 2019 and 2021 and the remaining eight between 2010 and 2018. In addition, 13 studies conducted network analyses, most in combination with other methods such as content or sentiment analyses; four studies performed qualitative or automated content analyses, and one study conducted an observation. In our sample, we found eight studies that analyzed Twitter, six that analyzed Facebook, one that researched YouTube, and three that compared two or more platforms. Six studies were conducted in the US American context; there were two studies each in Hong Kong and Israel, three in other country contexts (Italy, Canada, Australia), and two studies compared two or more countries. Four studies were conducted during election periods and four during heightened periods of political conflict (the Hong Kong protests and the Israeli–Palestinian conflict). The remaining studies were not conducted during election periods or at least did not specify so. Of these 18 studies, 10 looked at the development of group polarization over time, with time frames ranging from 24 hours to more than seven years. Eight studies were conducted in dual political systems, six in multi-party environments (but four of the six studies examined the dual contexts of Hong Kong and Israel–Palestine), and two studies compared several countries with different party systems. Nine studies in this category measured ideological polarization, five measured affective polarization, one study measured partisan polarization, another measured affective as well as ideological polarization, and two studies analyzed all three types of polarization.

Two studies that combined content analyses and opinion polls found that "cyberbalkanization" (fragmentation on the internet) and ideological polarization were related among young adults in Hong Kong (Chan & Fu, 2015, 2017). Another study that conducted an automated content analysis found that intergroup interactions characterized by direct dissent were drivers of affective polarization (Bliuc, Smith, & Moynihan, 2020). Furthermore, by comparing the two platforms Facebook and YouTube, one observation showed that the content, more than the algorithm, drove ideological polarization (Bessi et al., 2016). Concerning qualitative results from content analyses, one study found that right-wing users voiced a clear demarcation between (as well as the rejection and dehumanization of) the opposing political camp in Israel (Harel, Jameson, & Maoz, 2020). Another study found US Facebook and Twitter users to be polarized along party lines, whereas Dutch users demonstrated less party-related polarization. Instead, they drew a line between ordinary citizens and the elite

(Hameleers, 2020). The third qualitative study found tweets labeled as conservative to contain more negative perceptions toward the US healthcare reform, while tweets considered as liberal suggested the opposite and the majority of all tweets indicated some dislike of "the other" (Mendez, Cosby, & Mohanty, 2017).

Other studies that conducted content and network analyses found similarly differentiated results. For example, US American climate-change disbelievers on Twitter showed higher levels of hostility toward climate-change believers than vice versa (Tyagi, Uyheng, & Carley, 2020). In addition, conservatives in the US tweeted about ingroup candidates more positively and, simultaneously, more negatively about opposing candidates than did liberals (Mentzer, Fallon, Prichard, & Yates, 2020). Moreover, more Twitter users were found to be both positively and negatively polarized toward Hilary Clinton in comparison to Donald Trump (Grover et al., 2019), and men on Twitter appeared to voice less ingroup party support and less dislike of the out-group party than women did (Mentzer, Fallon, Prichard, & Yates, 2020). Furthermore, several studies identified homophily at work (Yardi & boyd, 2010; Gruzd & Roy, 2014), with interactions with like-minded individuals on Twitter strengthening group identity, whereas engagement with different-minded individuals reinforced ingroup and outgroup affiliations (Yardi & boyd, 2010). Moreover, higher engagement seems to have led to a higher number of polarized users (Grover et al., 2019), and users expressing negativity in their tweets were more ideologically polarized, while, surprisingly, negativity in the user's social environment had a depolarizing effect on ideological positions (Buder et al., 2021). Yarchi, Baden, and Kligler-Vilenchik (2020) analyzed Twitter, Facebook, and WhatsApp and found that only Twitter displays clear signs of political-group polarization. They found homophilic interaction patterns present, an increase in ideological polarization, and hostility between users of opposed camps—a sign of affective polarization. For WhatsApp, despite of the heterogeneous composition of the analyzed groups, a shared group identity and common purpose counteracted the polarization dynamics and even led to depolarization of its users. Facebook, in turn, was "found to be the least homophilic platform in terms of interactions, positions, and emotions expressed" (Yarchi, Baden and Kligler-Vilenchik, 2020, p. 1). Furthermore, we encountered four studies using modularity network approaches based on "likes" on Facebook. Contrary to the mixed and rather idiosyncratic results described above, the studies using a modularity approach all found ideological polarization present on Facebook. This might be due to the one-sidedness of the modularity approach: In all four cases, only the positive reactions (likes) were considered, whereas negative reactions and

opinions are not captured. Therefore, only half of the operationalization of ideological polarization, as described above, is included, which might bias results.

To summarize, group polarization studies exhibit significantly differentiated results, often holding only for specific groups of people or certain circumstances. Thus, this line of research is unable to support the idea of strong, generalized group polarization on social media platforms.

Mass Polarization

Quantitative Review of Studies

In the final selection step, we considered only studies that *empirically tested ideological, affective,* or *partisan polarization effects at the individual level*. This means that a larger number of studies focusing on aspects such as network homophily within social media platforms or group polarization were omitted in this step since they do not offer points of comparison that would allow making causal inferences about individual social media effects on political polarization of the public. This selection process yielded 31 studies published between 2014 and 2021, with a large majority (n = 23) published between 2018 and 2021. Articles were published in 23 different journals and the proceedings of one conference; 16 studies conducted surveys, 13 conducted experiments, and two combined surveys with observations. Of these 31 studies, 23 were originally developed for this research purpose, and eight used secondary data provided, for example, by the National Annenberg Election Survey or the Eurobarometer. Sample sizes ranged from n = 21 to n = 37,494, and 14 studies use representative, quota, or stratified samples, five used student samples, and the remaining 12 studies used convenience samples or did not specify their sampling procedure.

The majority of the studies analyzed (the frequency of) social media use in general (n = 10) or news consumption habits in social media environments (n = 6) as predictors of polarization. But some also analyzed polarization effects on specific social media platforms, as follows: YouTube (n = 3), Twitter (n = 2), Facebook (n = 5), Facebook and Twitter (n = 3), Facebook and KakaoTalk (n = 1), and WhatsApp (n = 1). Of these studies, five focused additionally on the influence of algorithmic news recommendations and customization options. The vast majority of studies (n = 20) were conducted in the US; of the remaining studies, three were conducted in Hong Kong, four in South Korea, three in different European countries

(Norway, Denmark, Netherlands), one in multiple European countries simultaneously, and one in Kenya. Thereby, 20 studies were conducted in a dual-party system, and the remaining 11 were conducted in multi-party systems (including three studies from the dual context of Hong Kong and other countries with multi-party systems dominated by two major political parties, i.e., South Korea and Kenya). Seven studies in all were conducted during election periods and the remaining studies were not. Three other studies, nevertheless, were conducted during heightened political conflict in Hong Kong and one study prior to a referendum in the Netherlands.

Eight of the studies measured affective polarization; 12 analyzed ideological polarization; and five selected partisan polarization as their dependent variable of interest. In addition, three studies incorporated measures of both ideological and affective or partisan polarization, and three analyzed all three types of polarization.

Of the 31 studies, eight identified ideological polarization patterns through social media use; three studies found affective polarization effects and three found partisan polarization effects. Three studies found only depolarization effects (affective and ideological, which includes one study where depolarization could be observed after deactivating Facebook), and two studies identified depolarization effects and no polarization simultaneously (affective and ideological). Seven studies found no polarization effects at all (affective, ideological, or partisan). The remaining six studies found mixed results, such as both polarization effects and depolarization effects or no polarization.

Topics analyzed in terms of ideological polarization are, on one hand, commonly discussed issues such as immigration, the economy, education, crime, health care, taxes, same-sex marriage, and feelings and attitudes toward candidates. On the other hand, more specific topics, such as North Korea, relations between the US and China, or investigations regarding Russian interference in elections are discussed. Several studies included both polarized and less-polarized topics.

Qualitative Review of Studies

Across all types of political polarization (affective, ideological, and partisan), our analysis indicates that there are several groups of main factors that appear to influence individual political polarization and depolarization processes.

The first factor found to be politically polarizing is the *frequency of social media use* or reliance on social media for news and political information.

This factor was considered mostly in studies based on survey designs, including surveys using longitudinal data as well as those dependent on cross-sectional data. Although only longitudinal data can provide causal inferences, we found no structural differences in the results between these two designs; thus, all findings will be presented together. In such research, the reliance on social media for political information was shown to affectively polarize users (Johnson, Kaye, & Lee, 2017), and time spent on social media indirectly heightened ideological polarization, especially for those users who frequently encountered like-minded information (Lu et al., 2020). Similarly, when users deactivated Facebook, they encountered less opinion-confirming partisan information, which, in turn, led to a decrease in all three types of polarization (Allcott et al., 2020). Likewise, it was found that active social media users had a higher likelihood of becoming engaged in political processes, which led them, in turn, to become ideologically more polarized than non-users (Lee, Shin, & Hong, 2018). Furthermore, users of social media and partisans were shown to become ideologically more polarized, whereas people using traditional media did not (Ohme, 2021; Suk et al., 2020). In contrast, another study found that the use of partisan mass media, as well as demographic factors (e.g., gender, age), had a stronger influence on ideological polarization than the use of social media (Lee et al., 2018). In line with this, Nguyen and Vu (2019) showed that reliance on social media did not ideologically polarize users more than participants relying on traditional media for political information. In total, in this category, we find evidence focused almost exclusively on ideological polarization with significantly mixed results, which might stem from the very general operationalization of social media use as a frequency measure.

Second, the strength of *partisanship and party ties* was found to play a crucial role in the process of polarization (e.g., Min & Yun, 2018). Party ties seemed to be strengthened by the use of social media (Cho et al., 2018), with stronger ties enhancing selective exposure, which led, in turn, to ideological polarization (Johnson, Kaye, & Lee, 2017). Nevertheless, political orientation had a stronger effect on ideological polarization than the use of social media (Lee et al., 2018), and social media use was not related to partisan polarization for moderate partisans (Lee, Shin, & Hong, 2018). This evidence becomes most clear by comparing single identifiers with dual identifiers, which showed that people who identified with only one political camp become more polarized through the use of social media (for all three types of polarization), whereas depolarization was observed for people identifying with both political ideologies (Kobayashi, 2020). The influence of partisanship was found almost exclusively in dual politi-

cal contexts. Therefore, it seems appropriate to conclude that people with strong party ties and a strong partisan identity in countries with clear opposing camps become more polarized through the use of social media, without generalizing these findings for all contexts and population groups.

Another group of factors that we discovered involves the *content* to which social media users are exposed or with which they engage. Here we differentiate between (a.) pro-attitudinal exposure, (b.) counter-attitudinal exposure, and (c.) pro- and counter-attitudinal expression.

Concerning pro-attitudinal exposure, research yielded highly mixed results. Twitter was found to heighten partisan polarization through the display of mostly opinion-confirming information (Hahn, Ryu, & Park, 2015). Facebook was also shown to reduce the likelihood of encountering counter-attitudinal news content, which increased affective polarization in comparison to counter-attitudinal news exposure (Levy, 2020), and Min and Yin (2018) found selective exposure toward political information to heighten affective polarization on KakaoTalk and, to a lesser extent, on Facebook. Similarly, the amount of time spent on social media indirectly heightened ideological polarization, especially for users who frequently encountered like-minded information (Lu et al., 2020). In contrast, Johnson et al. (2020) found that ideological polarization was not heightened through exposure to either pro- or counter-attitudinal information on Facebook. Likewise, Kim and Kim (2019) demonstrated that exposure to opinion-confirming comments did not affect ideological polarization.

Studies that looked at counter-attitudinal news exposure also found contradictory results. Beam, Hutchens, and Hmielowski (2018), for example, found counter-attitudinal news exposure on Facebook to increase over time, leading to a modest affective depolarization, whereas Levy (2020) found Facebook to decrease users' counter-attitudinal news exposure and, conversely, to increase pro-attitudinal news exposure, which heightened affective and ideological polarization. Furthermore, Bail et al. (2018) identified a backfire effect and an increase of ideological polarization through counter-attitudinal exposure for Republicans on Twitter. The latter study, however, forced users to expose themselves to counter-attitudinal news, which might have led to a negative predisposition and aversion toward the presented content beforehand.

Next to the causal dimensions of exposure to news content, behavioral components in the context of potentially polarizing content were also considered in past research. This includes pro- and counter-attitudinal expression, for example in the form of sharing news content as well as commenting on news or discussing it with other users. Turning to these studies investigating pro- and counter-attitudinal expression, Johnson et

al. (2020) found that sharing pro-attitudinal news articles on Facebook led to ideological polarization, whereas sharing counter-attitudinal news articles reduced ideological polarization. Kibet and Ward (2018) found higher levels of political discussion on WhatsApp to increase ideological and affective polarization, whereas, somewhat contradictorily, for respondents commenting more frequently on news, a reduction of both kinds of polarization was observed. This is contrasted by Cho et al.'s (2018) study that found YouTube users who express opinions about election campaigns to be strengthened in their initial opinion and to be affectively polarized. These contradictions might be explained by Karlsen et al.'s (2017) finding showing that discussions with both opponents and supporters on Facebook or Twitter might reinforce the preexisting attitude, possibly because of the aforementioned backfire effect. They also found that these effects were stronger for individuals with strong attitudes compared to those with moderate attitudes. Additionally, Shmargad and Klar (2019) demonstrated that those who are aware of their social surroundings share more moderate news articles when confronted with an out-group environment, whereas those previously enclosed by echo chambers share their preexisting (and more extreme) views independently of their social environment in the context of social networks.

Connected to the latter is the factor of *network heterogeneity***.** Here again, we find very mixed results. Network heterogeneity on social media in general was shown to decrease ideological polarization (Lee & Choi, 2020), whereas, in the case of WhatsApp in Kenya, higher levels of network heterogeneity increased ideological and affective polarization (Kibet & Ward, 2018). Representing greater differentiation, Lee et al. (2014) found higher levels of social network diversity to increase partisan polarization for individuals participating in more political discussions, whereas almost no effect was observed for those joining fewer political discussions.

The last group of factors comprises studies dealing with the role of **recommendation systems** or customization options implemented in social media. One study found that, in an experimental setup, customization on social media led to selective exposure, which heightened ideological polarization (Dylko et al., 2017). Similarly, affective polarization was heightened through YouTube's recommendation system in an experimental setup on the platform itself by providing opinion-confirming information (Hilbert et al., 2018). Other studies, by contrast, did not find affective polarization to be increased by users' customization preferences, and social preferences, i.e., the preferences of the users' social environment, were found in an experimental setup on YouTube as well as based on survey results to even reduce affective polarization (Cho et al., 2020; Feezell,

Wagner, & Conroy, 2021). Furthermore, sorting articles by popularity did not increase partisan polarization (Shmargad & Klar, 2020).

Nevertheless, probably the most-overlooked category in polarization research is that with null findings concerning polarization effects. Our analysis, however, demonstrates a considerable share of published empirical studies yielding null effects. For instance, ideological polarization was not affected in one study when users were exposed to uncivil commentary attacking the other side of an issue on YouTube (Hwang, Kim, & Huh, 2014). According to Munger et al. (2020), affective polarization was also not increased through partisan clickbait headlines on Facebook and Twitter. Lee and Choi (2020) demonstrated that individuals who fear others with opposing views and those who feel disadvantaged or excluded from dominant positions might adhere stick to their initial viewpoints; thus, ideological polarization was neither reduced nor heightened in this case either.

Overall, we find many mixed and often contradictory results. Therefore, in the following, we discuss structural differences between studies conducted in multi-party and dual-party contexts, studies that analyzed different social media platforms, and polarization and fragmentation studies—and the extent to which these differences might have affected the studies' results.

Evidence from Dual-Party Systems and Multi-Party Systems

Almost two-thirds of the studies analyzing mass polarization effects at the individual level were conducted in the US context; for studies analyzing group polarization, the US focus was slightly less dominant, yet still about half were conducted in this dual-party system. Another quarter of all studies concerning group and mass polarization effects were conducted in other countries with dual-party or multi-party systems dominated by two major political parties (e.g., the UK, South Korea, Kenya, Australia) or in countries with heightened political conflict between two groups (Hong Kong, Israel). This might be caused by polarization being a more severe problem in these contexts, but it may also be that measures are more easily operationalized if two clearly antagonistic groups contribute to this pattern. Both reasons may also help explain why, with very few exemptions, all studies considering affective polarization as an outcome of social media use were conducted in these dual contexts, whereas ideological and partisan polarization were analyzed almost exclusively in multi-party environments.

Specific Social Media Platforms and Method Choices

Most studies analyzing specific social media platforms focused on Facebook and Twitter, and a few each on YouTube, WhatsApp, and KakaoTalk. Although researchers have complained about the dominance of Twitter studies (Kubin & von Sikorski, 2021), we found such research to be dominant only in the area of fragmentation studies and only using "polarization" as a label, whereas for actual group or mass polarization studies, we saw a more balanced focus on different social media platforms. The difference between studies researching group polarization and those analyzing mass polarization effects is interesting: While studies analyzing mass polarization effects on Facebook found an exceptionally high number of depolarization and null effects of social media use, studies analyzing group polarization on Facebook found polarization effects (with the exemption of Buder et al., 2021). Additionally, studies of group polarization on Twitter found numerous patterns of polarization, whereas studies of mass polarization effects on Twitter and YouTube returned mixed results. This suggests that differences in operationalization resulted in this disparity. While group polarization was analyzed using content and network analyses, mass polarization effects were detected through surveys and experiments. This means that content and network analyses appear to be more prone to identifying patterns of polarization mirrored on social media platforms at the group level, whereas surveys and experiments at the individual level show few actual polarization effects of using specific social media platforms.

That the operationalization of social media use plays a critical role is also suggested by examining studies that employ modularity approaches. All these found polarization effects due to their one-sided approach; as described above, only the positive reactions (likes) were considered, whereas negative reactions and opinions were not captured. This omits half of the theoretical concept of polarization. Another methodological decision stands out regarding the group of studies operationalizing the frequency of generalized social media use. These studies found comparatively few depolarization effects or null effects. This might be due to the broad operationalization of "frequency of" or "reliance on" social media use used in most of the surveys, through which it is not possible to fully capture the depth and facets of social media usage as participants' self-disclosure is vulnerable to forgetfulness, social desirability, and other distortions. Furthermore, only some surveys analyzed longitudinal data; hence, causality might not always be assumed. In total, 19 of 31 research projects conducted studies with designs that allowed causal inferences, such as analyzing longitudinal data or conducting experiments. As previously noted, no clear patterns within

this group of studies and no differences between them were found; these studies, in addition, found both polarization and depolarization effects as well as no polarization effects for all three forms of political polarization (affective, ideological, and partisan).

Fragmentation vs. Polarization Studies

Despite the general belief that before and during elections political polarization increases, we could not find any systematic evidence in this structured literature review that an election taking place during the study period heightened any kind of political polarization. Nevertheless, this was a recurring finding in the fragmentation studies found by our literature search since they used the label "polarization." This supports the assumption that inconsistencies in operationalization and concepts distort conclusions about polarization effects. Conclusively, it seems that fragmentation is heightened on social media platforms before and during electoral campaigns but not necessarily political mass polarization.

Another structural difference identified between the fragmentation and polarization studies in this review is the strong concentration of fragmentation studies on the homophily of users. This might stem from the methodological dominance of network analyses in this category, which inherently have a focus on the compilation of users in different clusters, whereas in studies conducting experiments and surveys, this aspect is more difficult to measure. While the homogeneity of users within clusters is seen as evidence of political polarization in the fragmentation studies, the studies on mass polarization present a more nuanced picture.

Furthermore, fragmentation studies find "polarization effects" almost exclusively, while studies analyzing group polarization and individual polarization effects also find many depolarization effects and more-differentiated results, with heightened polarization being identified only for a certain group of social media users, for example. Again, this supports the assumption that conceptual unclarities and different ways of operationalization in this research field resulted in an overstatement of the role social media plays in the political polarization process.

Takeaways and Research Desiderata

Overall, we found significantly heterogeneous findings, conceptually overlapping constructs, and an inconclusive empirical research landscape. As polarization research has gained increasingly more attention in the past decade, the term "polarization" seems to be used frequently as a catchword rather than being an actual essential concept in a research endeavor. Frequently, the concepts of fragmentation and polarization appear to be equated, and distinctions between group and individual polarization effects are often not clarified. Definitions of and differentiations between the different dimensions of political polarization may be lacking or not applied. Therefore, first and foremost, in future polarization research, we plead for conceptual clarity and the provision of definitions of relevant concepts. Our literature review has proposed a typology of patterns of fragmentation, group, and mass polarization that may help ensure greater precision in the research landscape.

Concerning the role social media plays in the political polarization process, it is difficult to make universal statements based on the empirical findings generated thus far. Nevertheless, one unambiguous statement we can make based on our systematic review of empirical literature is that people with strong party ties and a strong partisan identity in countries with clear opposing camps become more polarized through the use of social media. Therefore, partisanship seems to play a major role in the polarization process and should be an essential component of future research in this area.

The same applies to the content that users consume and interact with on social media platforms. A substantial amount of research already concentrates on this aspect, but findings are inconclusive. (De)polarization effects have been observed both for exposure to and interaction with attitude-confirming and attitude-opposing content. Future research should, therefore, focus on disentangling these effects by applying comparable definitions and operationalization.

Furthermore, research analyzing algorithms or including actual running algorithms remains scarce. Thus far, most studies have used proxies for the role played by algorithmic recommendation systems, such as experimental setups with mock recommendations or survey designs (sometimes combined with behavioral web-tracking data). This low external validity leads to disparities between real-world social media use and research results. Nevertheless, it seems that these few studies agree on the finding that attitude-congruent content exposure evoked by a recommendation system heightened polarization.

Finally, our review of extant research indicates many other influences on political polarization, such as the strength of partisanship, polarized contexts, use of traditional media, personal conversations, age, or gender. These variables are often included as controls in empirical studies focusing on the effects of social media use. However, in many studies these control variables proved to produce a much stronger impact on polarization than social media use did. This indicates that social media is not as polarizing as popular discourse assumes; rather, a combination of different factors has to come into play to create strong polarization effects. As social media use itself is co-varying with many of the aforementioned third variables, it is crucial for future research on social-media-induced polarization effects to include a multitude of control variables to avoid producing false positive results as a result of omitted variable bias (Clarke, 2009).

Turning toward the methodological decisions and their implementations, we found a strong bias for studies dealing with political polarization to be conducted in dual-party contexts, especially in the US. Regarding the mass polarization effects of social media use in particular, we did not observe systematic differences between dual-party and multi-party contexts. It appears that the same mechanisms play a central role both in multi- and dual-party contexts. However, methodologically, we see that, with very few exceptions, all studies interested in affective polarization were conducted in dual contexts, whereas in multi-party environments, ideological and partisan polarization were analyzed almost exclusively. Therefore, we plead for researchers to also analyze affective polarization in multi-party contexts, based for example on Wagner's (2020) like–dislike scoring, and to conduct more internationally comparative research. Furthermore, most studies have focused on single social media platforms. As different platforms are expected to have different effects on their users, more comparative research analyzing multiple platforms in direct comparison is also needed.

Other methodological implications stand out as well. It seems that studies analyzing group polarization on Facebook and Twitter have found many more polarization effects than studies analyzing mass polarization effects at the individual level on the same platforms. This suggests that differences in accessing polarization on social media, either through observing societal group dynamics, in the case of group polarization, or individual polarization effects, in the case of mass polarization, yielded a disparity of findings. In the latter group, it is further noticeable that many studies based their analyses on the self-disclosed "frequency of" or "reliance on" social media use. These studies find comparatively few depolarization effects or no polarization effects, which might be because these types of

operationalization cannot possibly fully capture the depth and facets of social media usage and do not allow conclusions about the content to which participants were exposed or which they shared., Looking at the bigger picture, these findings show that the choice of how to access the concept of polarization and the choice of measurement play critical roles role in which polarization effects are found or if any can be found at all.

To put the present analysis into perspective, we may ask how our findings correspond to or differ from insights gleaned from other recent literature reviews. In line with Kubin and von Sikorski (2021), we found an increase in research over the past 10 years and a strong focus on the US context, but an increasing number of studies from other country contexts appearing in recent years. Likewise, our findings also correspond to the authors' insights that "political polarization is not consistently discussed, or measured, across the literature" (p. 197), that "ideological and affective polarization are not clearly defined, nor consistently measured" (p. 188), and that there is "a lack of research exploring ways (social) media can depolarize" (Kubin & von Sikorski, 2021, p. 188). However, contrary to the literature reviews conducted by Kubin and von Sikorski (2021) as well as Iandoli et al. (2021), we did not find a strong dominance of polarization studies that analyzed Twitter. What we have found is a hyperfocus on Twitter for those studies that have actually analyzed fragmentation. This difference in findings results again from the lack of conceptual differentiation between polarization and fragmentation studies discussed above. Moreover, contrary to Kubin and von Sikorski (2021), we did not find that pro-attitudinal media clearly exacerbates polarization (see, e.g., Johnson et al., 2020; Kim & Kim, 2019). Furthermore, also in contrast to Kubin and von Sikorski (2021), our literature review included several experiments that provided "insight into ways social media can decrease (or have no effect) on ideological [and affective] polarization" (e.g., Cho et al., 2020; Munger et al., 2020).

Overall, we can say that the landscape of political polarization research needs more conceptual clarity and more inclusion of and comparison across different political and national contexts—and that, in general, the causal role of social media in the process of political polarization seems overstated or can be, at least, strongly disputed.

References

Abramowitz, A. I. (2010). *The Disappearing Center. Engaged Citizens, Polarization, and American Democracy.* Yale University Press.

Abramowitz, A. I., & Saunders, K. L. (2008). Is polarization a myth? *The Journal of Politics*, *70*(2), 542–555. https://doi.org/10.1017/S0022381608080493

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, *110*(3), 629–676. https://doi.org/10.1257/aer.20190658

Arceneaux, K. & Johnson, M. (2010). Does Media Fragmentation Produce Mass Polarization? Selective Exposure and a New Era of Minimal Effects. *APSA 2010 Annual Meeting Paper*. Retrieved from https://ssrn.com/abstract=1642723

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, *115*(37), 9216–9221. https://doi.org/10.1073/pnas.1804840115

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

Baldassarri, D., & Gelman, A. (2008). Partisans without constraint: Political polarization and trends in American public opinion. *American Journal of Sociology*, *114*(2), 408–446. https://doi.org/10.1086/590649

Bartels, L. M. (2000). Partisanship and voting behavior, 1952-1996. *American Journal of Political Science*, *44*(1), 35. https://doi.org/10.2307/2669291

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

Bessi, A., Zollo, F., Del Vicario, M., Puliga, M., Scala, A., Caldarelli, G., Uzzi, B., & Quattrociocchi, W. (2016). Users' polarization on Facebook and YouTube. *PLOS ONE*, *11*(8), e0159641. https://doi.org/10.1371/journal.pone.0159641

Bliuc, A.-M., Smith, L. G. E., & Moynihan, T. (2020). "You wouldn't celebrate September 11": Testing online polarisation between opposing ideological camps on YouTube. *Group Processes & Intergroup Relations*, *23*(6), 827–844. https://doi.org/10.1177/1368430220942567

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

Bor, A., & Petersen, M. B. (2021). The psychology of online political hostility: A comprehensive, cross-national test of the mismatch hypothesis. *American Political Science Review*, 1–18. https://doi.org/10.1017/S0003055421000885

Boxell, L., Gentzkow, M., & Shapiro, J. (2020). Cross-country trends in affective polarization (Nr. w26669; S. w26669). *National Bureau of Economic Research*. https://doi.org/10.3386/w26669

boyd, danah m., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210–230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues*, *55*(3), 429–444. https://doi.org/10.1111/0022-4537.00126

Bright, J. (2018). Explaining the emergence of political fragmentation on social media: The role of ideology and extremism. *Journal of Computer-Mediated Communication*, *23*(1), 17–33. https://doi.org/10.1093/jcmc/zmx002

Bruns, A. (2019). *Are Filter Bubbles Real?* Polity Press.

Buder, J., Rabl, L., Feiks, M., Badermann, M., & Zurstiege, G. (2020). *Does negatively toned language use on social media lead to attitude polarization?* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/dx9ws

Cambridge Dictionary. (n.d.). Polarization. *dictionary.cambridge.org*. Retrieved December 4, 2021, from https://dictionary.cambridge.org/dictionary/english/polarization

Merriam-Webster. (n.d.). Polarization. *Merriam-Webster.com Dictionary*. Retrieved December 4, 2021, from https://www.merriam-webster.com/dictionary/polarization

Chan, C., & Fu, K. (2015). Predicting political polarization from cyberbalkanization: Time series analysis of facebook pages and opinion poll during the hong kong occupy movement. Proceedings of the ACM Web Science Conference, 1–2. https://doi.org/10.1145/2786451.2786509

Chan, C., & Fu, K. (2017). The relationship between cyberbalkanization and opinion polarization: Time-series analysis on Facebook pages and opinion polls during the Hong Kong occupy movement and the associated debate on political reform: cyberbalkanization and opinion polarization. *Journal of Computer-Mediated Communication*, *22*(5), 266–283. https://doi.org/10.1111/jcc4.12192

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

Clarke, K. A. (2009). Return of the phantom menace: Omitted variable bias in political research. *Conflict Management and Peace Science*, *26*(1), 46–66. https://doi.org/10.1177/0738894208097666

Cohen, J. N. (2018). Exploring echo-systems: How algorithms shape immersive media environments. *Journal of Media Literacy Education*, *10*(2), 139–151. https://doi.org/10.23860/JMLE-2018-10-2-8

Conover, M., Ratkiewicz, J., Francisco, M., Goncalves, B., Menczer, F., & Flammini, A. (2021). Political polarization on Twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, *5*(1), 89-96. Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/14126

Dahlberg, L. (2007). Rethinking the fragmentation of the cyberpublic: From consensus to contestation. *New Media & Society*, *9*(5), 827–847. https://doi.org/10.1177/1461444807081228

Del Vicario, M., Zollo, F., Caldarelli, G., Scala, A., & Quattrociocchi, W. (2017). Mapping social dynamics on Facebook: The Brexit debate. *Social Networks*, *50*, 6–16. https://doi.org/10.1016/j.socnet.2017.02.002

DiMaggio, P., Evans, J., & Bryson, B. (1996). Have American's social attitudes become more polarized? *American Journal of Sociology*, 102(3), 690–755.

Dylko, I., Dolgov, I., Hoffman, W., Eckhart, N., Molina, M., & Aaziz, O. (2017). The dark side of technology: An experimental investigation of the influence of customizability technology on online political selective exposure. *Computers in Human Behavior*, *73*, 181–190. https://doi.org/10.1016/j.chb.2017.03.031

Feezell, J. T., Wagner, J. K., & Conroy, M. (2021). Exploring the effects of algorithm-driven news sources on political behavior and polarization. *Computers in Human Behavior, 116*, 106626. https://doi.org/10.1016/j.chb.2020.106626

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

Fiorina, M. P., Abrams, S. J., & Pope, J. C. (2005). *Culture War? The Myth of a Polarized America*. Pearson Longman.

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

Garimella, K., De Francisci Morales, G., Gionis, A., & Mathioudakis, M. (2017). The effect of collective attention on controversial debates on social media. *Proceedings of the 2017 ACM on Web Science Conference*, 43–52. https://doi.org/10.1145/3091478.3091486

Garrett, R. K., Gvirsman, S. D., Johnson, B. K., Tsfati, Y., Neo, R., & Dal, A. (2014). Implications of pro- and counterattitudinal information exposure for affective polarization: Partisan media exposure and affective polarization. *Human Communication Research, 40*(3), 309–332. https://doi.org/10.1111/hcre.12028

Gentzkow, M. & Shapiro, J.M. (2010). What drives media slant? Evidence from u. S. Daily newspapers. (2010). *Econometrica*, *78*(1), 35–71. https://doi.org/10.3982/ECTA7195

Gimpel, J. G., & Hui, I. S. (2015). Seeking politically compatible neighbors? The role of neighborhood partisan composition in residential sorting. *Political Geography, 48*, 130–142. https://doi.org/10.1016/j.polgeo.2014.11.003

Grover, P., Kar, A. K., Dwivedi, Y. K., & Janssen, M. (2019). Polarization and acculturation in US Election 2016 outcomes – Can twitter analytics predict changes in voting preferences. *Technological Forecasting and Social Change*, *145*, 438–460. https://doi.org/10.1016/j.techfore.2018.09.009

Gruzd, A., & Roy, J. (2014). Investigating political polarization on twitter: A Canadian perspective: investigating political polarization on twitter. *Policy & Internet, 6*(1), 28–45. https://doi.org/10.1002/1944-2866.POI354

Hahn, K. S., Ryu, S., & Park, S. (2015). Fragmentation in the twitter following of news outlets: The representation of South Korean users' ideological and generational cleavage. *Journalism & Mass Communication Quarterly, 92*(1), 56–76. https://doi.org/10.1177/1077699014559499

Hameleers, M. (2020). Augmenting polarization via social media? A comparative analysis of Trump's and Wilders' online populist communication and the electorate's interpretations surrounding the elections. *Acta Politica*, *55*(3), 331–350. https://doi.org/10.1057/s41269-018-0119-8

Harel, T. O., Jameson, J. K., & Maoz, I. (2020). The normalization of hatred: Identity, affective polarization, and dehumanization on Facebook in the context of intractable political conflict. *Social Media + Society*, *6*(2), 205630512091398. https://doi.org/10.1177/2056305120913983

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

Huber, G. A., & Malhotra, N. (2017). Political homophily in social relationships: Evidence from online dating behavior. *The Journal of Politics*, *79*(1), 269–283. https://doi.org/10.1086/687533

Hwang, H., Kim, Y., & Huh, C. U. (2014). Seeing is believing: Effects of uncivil online debate on political polarization and expectations of deliberation. *Journal of Broadcasting & Electronic Media*, *58*(4), 621–633. https://doi.org/10.1080/08838151.2014.966365

Iandoli, L., Primario, S., & Zollo, G. (2021). The impact of group polarization on the quality of online debate in social media: A systematic literature review. *Technological Forecasting and Social Change*, *170*, 120924. https://doi.org/10.1016/j.techfore.2021.120924

Isenberg, D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology*, *50*(6), 1141–1151. https://doi.org/10.1037/0022-3514.50.6.1141

Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization: fear and loathing across party lines. *American Journal of Political Science*, *59*(3), 690–707. https://doi.org/10.1111/ajps.12152

Iyengar, S., Konitzer, T., & Tedin, K. (2018). The home as a political fortress: Family agreement in an era of polarization. *The Journal of Politics*, *80*(4), 1326–1338. https://doi.org/10.1086/698929

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the united states. *Annual Review of Political Science*, *22*(1), 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034

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

Jackson, T. W., & Farzaneh, P. (2012). Theory-based model of factors affecting information overload. *International Journal of Information Management*, *32*(6), 523–532. https://doi.org/10.1016/j.ijinfomgt.2012.04.006

Jensen, J., Kaplan, E., Naidu, S., & Wilse-Samson, L. (2012). Political polarization and the dynamics of political language: Evidence from 130 years of partisan speech. *Brookings Papers on Economic Activity*, (1), 1–81. https://doi.org/10.1353/eca.2012.0017

Johnson, B. K., Neo, R. L., Heijnen, M. E. M., Smits, L., & van Veen, C. (2020). Issues, involvement, and influence: Effects of selective exposure and sharing on polarization and participation. *Computers in Human Behavior*, *104*, 106155. https://doi.org/10.1016/j.chb.2019.09.031

Johnson, T. J., Kaye, B. K., & Lee, A. M. (2017). Blinded by the spite? Path model of political attitudes, selectivity, and social media. *Atlantic Journal of Communication*, *25*(3), 181–196. https://doi.org/10.1080/15456870.2017.1324454

Karlsen, R., Steen-Johnsen, K., Wollebæk, D., & Enjolras, B. (2017). Echo chamber and trench warfare dynamics in online debates. *European Journal of Communication*, *32*(3), 257–273. https://doi.org/10.1177/0267323117695734

Kearney, M. W. (2019). Analyzing change in network polarization. *New Media & Society*, *21*(6), 1380–1402. https://doi.org/10.1177/1461444818822813

Kibet, A., & Ward, S. (2018). Socially networked heterogeneity: The influence of Whatsapp as a social networking site on polarisation in Kenya. *African Journalism Studies*, *39*(4), 42–66. https://doi.org/10.1080/23743670.2018.1537979

Kim, Y., & Kim, Y. (2019). Incivility on Facebook and political polarization: The mediating role of seeking further comments and negative emotion. *Computers in Human Behavior*, *99*, 219–227. https://doi.org/10.1016/j.chb.2019.05.022

Kobayashi, T. (2020). Depolarization through social media use: Evidence from dual identifiers in Hong Kong. *New Media & Society*, *22*(8), 1339–1358. https://doi.org/10.1177/1461444820910124

Lai, M., Tambuscio, M., Patti, V., Ruffo, G., & Rosso, P. (2019). Stance polarity in political debates: A diachronic perspective of network homophily and conversations on Twitter. *Data & Knowledge Engineering*, *124*, 101738. https://doi.org/10.1016/j.datak.2019.101738

Lau, R. R., Andersen, D. J., Ditonto, T. M., Kleinberg, M. S., & Redlawsk, D. P. (2017). Effect of media environment diversity and advertising tone on information search, selective exposure, and affective polarization. *Political Behavior*, *39*(1), 231–255. https://doi.org/10.1007/s11109-016-9354-8

Layman, G. C., Carsey, T. M., & Horowitz, J. M. (2006). Party polarization in American politics: Characteristics, causes, and consequences. *Annual Review of Political Science*, *9*(1), 83–110. https://doi.org/10.1146/annurev.polisci.9.070204.105138

Lee, J., & Choi, Y. (2020). Effects of network heterogeneity on social media on opinion polarization among South Koreans: Focusing on fear and political orientation. *International Communication Gazette*, *82*(2), 119–139. https://doi.org/10.1177/1748048518820499

Lee, H., & Hahn, K. S. (2018). Partisan selective following on Twitter over time: Polarization or depolarization? *Asian Journal of Communication*, *28*(3), 227–246. https://doi.org/10.1080/01292986.2017.1384845

Lee, F. L. F. (2016). Impact of social media on opinion polarization in varying times. *Communication and the Public, 1*(1), 56–71. https://doi.org/10.1177/2057047315617763

Lee, P. S. N., So, C. Y. K., Lee, F., Leung, L., & Chan, M. (2018). Social media and political partisanship – A subaltern public sphere's role in democracy. *Telematics and Informatics, 35*(7), 1949–1957. https://doi.org/10.1016/j.tele.2018.06.007

Lee, C., Shin, J., & Hong, A. (2018). Does social media use really make people politically polarized? Direct and indirect effects of social media use on political polarization in South Korea. *Telematics and Informatics, 35*(1), 245–254. https://doi.org/10.1016/j.tele.2017.11.005

Lelkes, Y. (2018). Affective polarization and ideological sorting: A reciprocal, albeit weak, relationship. *The Forum, 16*(1), 67–79. https://doi.org/10.1515/for-2018-0005

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

Levy, R. (2021). Social media, news consumption, and polarization: Evidence from a field experiment. *American Economic Review, 111*(3), 831–870. https://doi.org/10.1257/aer.20191777

Lu, Y., Ray, R., Ha, L., & Chen, P. (2020). Social media news consumption and opinion polarization on China's trade practices: Evidence from a U.S. national survey. *International Journal of Communication*, *14*, 3478–3495.

Mason, L. (2016). A cross-cutting calm: How social sorting drives affective polarization. *Public Opinion Quarterly, 80*(S1), 351–377. https://doi.org/10.1093/poq/nfw001

McCarty, N. M., Poole, K. T., & Rosenthal, H. (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. MIT Press.

McConnell, C., Margalit, Y., Malhotra, N., & Levendusky, M. (2018). The economic consequences of partisanship in a polarized era. *American Journal of Political Science, 62*(1), 5–18. https://doi.org/10.1111/ajps.12330

McCoy, J., Rahman, T., & Somer, M. (2018). Polarization and the global crisis of democracy: Common patterns, dynamics, and pernicious consequences for democratic politics. *American Behavioral Scientist, 62*(1), 16–42. https://doi.org/10.1177/0002764218759576

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

Mendez, G.P.R, Cosby, A.G. & Mohanty, S.D. (2017). Obamacare and political polarization on Twitter: An application of machine learning and social network analysis. *Teorija in Praksa, 55*(2).

Mentzer, K., Fallon, K., Prichard, J. J. & Yates, D. (2020): Measuring and unpacking affective polarization on Twitter: The role of party and gender in the 2018 Senate races. *Proceedings of the 53rd Hawaii International Conference on System Sciences (HICSS)*. 1-10

Michelitch, K. (2015). Does electoral competition exacerbate interethnic or interpartisan economic discrimination? Evidence from a field experiment in market price bargaining. *American Political Science Review*, *109*(1), 43–61. https://doi.org/10.1017/S0003055414000628

Min, H., & Yun, S. (2018). Selective exposure and political polarization of public opinion on the presidential impeachment in south korea: Facebook vs. Kakaotalk. *Korea Observer - Institute of Korean Studies*, *49*(1), 137–159. https://doi.org/10.29152/KOIKS. 2018.49.1.137

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

Munger, K., Egan, P. J., Nagler, J., Ronen, J., & Tucker, J. (2020). Political knowledge and misinformation in the era of social media: Evidence from the 2015 uk election. *British Journal of Political Science*, 1–21. https://doi.org/10.1017/S0007123420000198

Nguyen, A., & Vu, H. T. (2019). Testing popular news discourse on the "echo chamber" effect: Does political polarisation occur among those relying on social media as their primary politics news source? *First Monday*. https://doi.org/10.5210/fm.v24i6.9632

Ohme, J. (2021). Algorithmic social media use and its relationship to attitude reinforcement and issue-specific political participation – The case of the 2015 European immigration movements. *Journal of Information Technology & Politics*, 18(1), 36–54. https://doi.org/10.1080/19331681.2020.1805085

Papacharissi, Z. (2002). The virtual sphere: The internet as a public sphere. *New Media & Society*, *4*(1), 9–27. https://doi.org/10.1177/14614440222226244

Pariser, E. (2011). *The Filter Bubble: What the Internet is Hiding from You*. Penguin Press.

Pew Research Center. (2017). The partisan divide on political values grows even wider. Retrieved April 10 2021, from https://www.people-press.org/2017/10/05/the-partisan-divide-on-political-values-grows-even-wider/.

Rogowski, J. C., & Sutherland, J. L. (2016). How ideology fuels affective polarization. *Political Behavior*, *38*(2), 485–508. https://doi.org/10.1007/s11109-015-9323-7

Scharkow, M., Mangold, F., Stier, S., & Breuer, J. (2020). How social network sites and other online intermediaries increase exposure to news. *Proceedings of the National Academy of Sciences*, *117*(6), 2761–2763. https://doi.org/10.1073/pnas.1918279117

Sened, I. (1996). A model of coalition formation: Theory and evidence. *The Journal of Politics, 58*(2), 350–372. https://doi.org/10.2307/2960230

Shmargad, Y., & Klar, S. (2019). How Partisan Online Environments Shape Communication with Political Outgroups. *International Journal of Communication, 13*, 27.

Shmargad, Y., & Klar, S. (2020). Sorting the news: How ranking by popularity polarizes our politics. *Political Communication, 37*(3), 423–446. https://doi.org/10.1080/10584609.2020.1713267

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

Suk, J., Shah, D. V., Wells, C., Wagner, M. W., Friedland, L. A., Cramer, K. J., Hughes, C., & Franklin, C. (2020). Do improving conditions harden partisan preferences? Lived experiences, imagined communities, and polarized evaluations. *International Journal of Public Opinion Research, 32*(4), 750–768. https://doi.org/10.1093/ijpor/edz051

Sunstein, C. R. (2001). *Echo Chambers: Bush v. Gore, Impeachment, and Beyond*. Princeton University Press.

Tyagi, A., Uyheng, J., & Carley, K. M. (2020). Affective polarization in online climate change discourse on twitter. *arXiv:2008.13051* [cs]. http://arxiv.org/abs/2008.13051

Usui, S., Yoshida, M., & Toriumi, F. (2018). Analysis of information polarization during japan's 2017 election. *2018 IEEE International Conference on Big Data (Big Data)*, 4383–4386. https://doi.org/10.1109/BigData.2018.8622143

Valkenburg, P. M. (2017). Understanding self-effects in social media. *Human Communication Research*, *43*(4), 477–490. https://doi.org/10.1111/hcre.12113

van Aelst, P., Strömbäck, J., Aalberg, T., Esser, F., de Vreese, C., Matthes, J., Hopmann, D., Salgado, S., Hubé, N., Stępińska, A., Papathanassopoulos, S., Berganza, R., Legnante, G., Reinemann, C., Sheafer, T., & Stanyer, J. (2017). Political communication in a high-choice media environment: A challenge for democracy? *Annals of the International Communication Association, 41*(1), 3–27. https://doi.org/10.1080/23808985.2017.1288551

Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral Studies, 69*, 102199. https://doi.org/10.1016/j.electstud.2020.102199

Warner, B. R. (2010). Segmenting the electorate: The effects of exposure to political extremism online. *Communication Studies, 61*(4), 430–444. https://doi.org/10.1080/10510974.2010.497069

Webster, S. W., & Abramowitz, A. I. (2017). The ideological foundations of affective polarization in the U. S. electorate. *American Politics Research, 45*(4), 621–647. https://doi.org/10.1177/1532673X17703132

Yang, M., Wen, X., Lin, Y.-R., & Deng, L. (2017). Quantifying content polarization on Twitter. *2017 IEEE 3rd International Conference on Collaboration and Internet Computing (CIC)*, 299–308. https://doi.org/10.1109/CIC.2017.00047

Yarchi, M., Baden, C., & Kligler-Vilenchik, N. (2021). Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media. *Political Communication, 38*(1–2), 98–139. https://doi.org/10.1080/10584609.2020.1785067

Yardi, S., & boyd, D. (2010). Dynamic debates: An analysis of group polarization over time on Twitter. *Bulletin of Science, Technology & Society, 30*(5), 316–327. https://doi.org/10.1177/0270467610380011

Zollo, F. (2019). Dealing with digital misinformation: A polarised context of narratives and tribes. *Proceedings of the Third EFSA Scientific Conference: Science, Food and Society.* https://doi.org/10.2903/j.efsa.2019.e170720

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

Katharina Ludwig *(M.A., U of Mannhheim, 2020) is Research Associate in the Mannheim Centre for European Social Research (MZES) and the Institute for Media and Communication Studies, University of Mannheim, Germany. Her research interests concern the perception, usage and effects of political communication in news media and social networks with special emphasis on (1) polarization, fragmentation and extremism, (2) migration, flight, and racism/discrimination, (3) quantification and effects of (self-transcendent) emotions.*

Philipp Müller *(Dr. phil., LMU Munich, 2015) is Senior Lecturer in the Institute for Media and Communication Studies, University of Mannheim, Germany, and Project Director in the Mannheim Centre for European Social Research (MZES). His research deals with questions of political communication and media change, with a special focus on digital news consumption and its effects on democracy and societal cohesion. Wolfram Peiser was his PhD advisor in Munich. Philipp worked as research and teaching associate with his chair from 2010 to 2014.*

Appendix 1: Overview of Group Polarization Studies

Title	Country	Method	S.M. Platform	Type of Polarization	Results on Polarization
Bessi, A., Zollo, F., Vicario, M. D., Puliga, M., Scala, A., Caldarelli, G., Uzzi, B., & Quattrociocchi, W. (2016): Users' polarization on Facebook and Youtube. doi: 10.1371/journal.pone.0159641	USA	Observation	Facebook & Youtube	ideological	The content more than the algorithm drove polarization
Bliuc, A., Smith, L. G.E., & Moynihan, T. (2020): "You wouldn't celebrate September 11": Testing online polarisation between opposing ideological camps on YouTube doi: 10.1177/1368430220942567	Australia	Automated Content Analysis	Youtube	affective	Mostly intergroup interaction, when direct dissent is expressed, drives polarisation.
Brugnoli, E., Cinelli, M., Quattrociocchi, W. & Scala, A. (2019): Recursive patterns in online echo chambers doi: 10.1038/s41598-019-56191-7	Italy	Network Analysis - Modularity Approach	Facebook	ideological	Polarized communities: polarized users tend to remain confined within groups of very few pages and reinforce their preexisting beliefs by leveraging the activity of their like-minded neighbors
Buder, J., Rabl, L., Feiks, M., Badermann, M., & Zurstiege, G. (2021): Does negatively toned language use on social media lead to attitude polarization? doi: 10.1016/j.chb.2020.106663	/	Network Analysis & Sentiment Analysis	Twitter	ideological	Negativity in tweets is linked to a polarized attitude. Negativity in a user's social environment had a slightly depolarizing effect on attitude extremity.
Chan, C. & Fu, K. (2015): Predicting political polarization from cyberbalkanization: Time series analysis of Facebook pages and opinion polls during the Hong Kong Occupy Movement doi: 10.1145/2786451.2786509	Hong-Kong	Network Analysis & Opinion Polls	Twitter	ideological	Cyberbalkanization and opinion polarization are connected for young adults.

Title	Country	Method	S.M. Platform	Type of Polarization	Results on Polarization
Chan, C. & Fu, K. (2017): The Relationship between cyberbalkanization and opinion polarization: Time-series analysis on Facebook pages and opinion polls during the Hong Kong occupy movement and the associated debate on political reform. doi: 10.1111/jcc4.12192	Hong-Kong	Network Analysis & Opinion Polls	Facebook	ideological	Cyberbalkanization and opinion polarization are connected for young adults.
Del Vicario, M., Zollo, F., Caldarelli, G.,Scala, A. & Quattrociocchi, W. (2017): Mapping social dynamics on Facebook: The Brexit debate. doi: 10.1016/j.socnet.2017.02.002	UK	Network Analysis - Modularity Approach	Facebook	ideological (towards news pages)	Polarized communities: users are divided into two main distinct groups and confine their attention on specific pages.
Grover, P., Kar, A. K., Dwivedi, Y. K., & Janssen, M. (2019): Polarization and acculturation in US Election 2016 outcomes – Can twitter analytics predict changes in voting preferences. doi: 10.1016/j.techfore.2018.09.009	USA	Content, Network & Sentiment Analysis	Twitter	partisan	Higher engagement leads to a higher number of polarized users.
Gruzd, A., & Roy, J. (2014): Investigating political polarization on Twitter: A Canadian perspective. doi: 10.1002/1944-2866.POI354	Canada	Network Analysis & Content Analysis	Twitter	affective	Clustering effect around shared political views; Evidence of cross-ideological connections and exchanges characterized by hostility.
Hameleers, M.(2020): Augmenting polarization via social media? A comparative analysis of Trump's and Wilders' online populist communication and the electorate's interpretations surrounding the elections. doi: 10.1057/s41269-018-0119-8	USA, Netherlands	Qualitative Content Analysis	Facebook & Twitter	affective	U.S. citizens find themselves fragmented along partisan lines. In the Netherlands, the divide between the ordinary people and the others (elite) is more central.
Harel, T., Jameson, J., & Maoz, I., (2020): The Normalization of Hatred:Identity, Affective polarization, and dehumanization on	Israel	Qualitative Content Analysis	Facebook	affective	Individuals on the right-wing FB-page seem polarized.

Title	Country	Method	S.M. Platform	Type of Polarization	Results on Polarization
Facebook in the context of intractable political conflict. doi: 10.1177/2056305120913983					
Mendez, G., Cosby, A.G., & Mohanty, S.D. (2017): Obamacare and political polarization on Twitter: An application of machine learning and social network analysis.	USA	Qualitative Content Analysis & Sentiment Analysis	Twitter	partisan & ideological & affective	Users are polarized along party lines; Influencers have a growing capacity to polarise the public's views and opinions; Ideologically opposing tweets express opposing perceptions about health care; Majority of the tweets indicate they are against 'the other'.
Mentzer, K., Fallon,K., Prichard, J. J. & Yates, D. (2020): Measuring and Unpacking affective polarization on Twitter: The role of party and gender in the 2018 Senate races. doi: 10.24251/hicss.2020.301	USA	Network Analysis & Sentiment Analysis	Twitter	affective	greater level of polarization, and larger fluctuations in polarization, among Conservatives over Liberals; women and Conservatives expressed stronger in-group party support and greater dislike of out-group party than men or Liberals
Schmidt, A.L., Zollo, F., Scala, A., C. Betsch, C. & Quattrociocchi, W. (2018): Polarization of the vaccination debate on Facebook. doi: 10.1016/j.vaccine.2018.05.040	/	Network Analysis - Modularity Approach	Facebook	ideological	polarized communities
Tyagi, A., Uyheng, J., & Carley, K. M. (2020): Affective polarization in online climate change discourse on Twitter. arXiv:2008.13051v1	USA	Network Analysis & Sentiment Analysis	Twitter	affective	Climate change Disbelievers tended to exhibit high levels of hostility toward climate change Believers; Disbelievers had similarly valenced interactions toward in-group and out-group members
Williams, H.T.P., McMurray, J.R., Kurz, T. & Lambert, F. H. (2015: Network analysis reveals open forums and echo chambers in soci-	/	Network Analysis	Twitter	ideological	High ideological polarisation; social media discussion of climate change is characterised by strong attitude-

Title	Country	Method	S.M. Platform	Type of Polarization	Results on Polarization
al media discussions of climate change. doi: 10.1016/j.gloenvcha.2015.03.006					based homophily and widespread segregation into like-minded communities. Users exposed to diverse views in mixed attitude communities were less likely to hold a polarised view; polarized members were more likely to express negative sentiment towards others with differing views.
Yarchi, M., Baden, C., & Kligler-Vilenchik, N. (2020): Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on Ssocial media. doi: 10.1080/10584609.2020.1785067	Israel	Content Analysis & Network Analysis	Facebook, Twitter, WhatsApp	affective, ideological	Dependent on platform: Twitter is polarized; WhatsApp less, Facebook not.
Yardi, S., & Boyd, D. (2010): Dynamic debates: An analysis of group polarization over time on Twitter. doi: 10.1177/0270467610380011	USA	Content Analysis & Network Analysis	Twitter	ideological	Users were more likely to interact within their like-minded group, which strengthend group identity, but also actively engaged with ideologically competing groups, which reinforced in-group and out-group affiliation

Title	Country	Method	S.M. Platform	Type of Polarization	Results on Polarization
Zollo, F. (2019): Dealing with digital misinformation: Apolarised context of narratives and tribes. doi: 10.2903/j.efsa.2019.e170720	Italy, USA, UK	Network, Content - Modularity Approach & Sentiment Analysis	Facebook	ideological	polarized communities: reinforcement of world view by opinion-confirming content and interaction with like-minded people who; sentiment of polarised users tend to be more negative than general ones

Appendix 2: Overview of Mass Polarization Studies

Title	Country	Method	S.M. Platform	Type of Polarization	Results on Polarization
Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. doi: 10.1257/aer.20190658	USA	Experiment	Facebook	affective, ideological, partisan	Deactivation of Facebook reduced political polarization.
Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. doi: 10.1073/pnas.1804840115	USA	Experiment	Twitter	ideological	No evidence that exposing Twitter users to opposing views reduces polarization. Evidence for backfire effects and increase of polarization through counter-attitudinal views.
Beam, M. A., Hutchens, M. J., & Hmielowski, J. D. (2018). Facebook news and (De)polarization: Reinforcing spirals in the 2016 US election. doi: 10.1080/1369118X.2018.1444783	USA	Survey	Facebook	affective	Facebook news use was related to a modest over-time spiral of depolarization. Counter-attitudinal news exposure increased over time, which resulted in depolarization. We found no evidence of a parallel model, where pro-attitudinal exposure stemming from Facebook news use resulted in greater affective polarization.
Cho, J., Ahmed, S., Hilbert, M., Liu, B. & Luu, J. (2020). Do search algorithms endanger democracy? An experimental investigation of algorithm effects on political polarization. doi: 10.1080/08838151.2020.1757365	USA	Experiment	Youtube	affective	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.

Title	Country	Method	S.M. Platform	Type of Polarization	Results on Polarization
Cho, J., Ahmed, S., Keum, H., Choi, Y. J., & Lee, J. H. (2018). Influencing myself: Self-reinforcement through online political expression. doi: 10.1177/0093650216644020	USA	Survey	Social Media	ideological, partisan	Results show that expressing opinions about election campaigns strengthened the expresser's initially held opinions.
Dylko, I., Dolgov, I., Hoffman, W., Eckhart, N., Molina, M., & Aaziz, O. (2017). The dark side of technology: An experimental investigation of the influence of customizability technology on online political selective exposure. doi: 10.1016/j.chb.2017.03.031	USA	Experiment	Social Media	ideological	Customizability increased political polarization indirectly, via its effect on political selective exposure.
Feezell, J. T., Wagner, J. K., & Conroy, M. (2021). Exploring the effects of algorithm-driven news sources on political behavior and polarization. doi: 10.1016/j.chb.2020.106626	USA	Survey	Facebook & Twitter	affective	Neither non-algorithmic media nor user-driven or socially-driven algorithms influenced political polarization
Hahn, K. S., Ryu, S., & Park, S. (2015). Fragmentation in the twitter following of news outlets: The representation of south korean users' ideological and generational cleavage. doi: 10.1177/1077699014559499	South Korea	Survey & Webbrousing	Twitter	partisan	Partisan and generational selectivity sharply polarizes news following on Twitter; Results imply that the network of Twitter following mirrors the landscape of offline political polarization.
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. doi: 10.1080/19312458.2018.1479843	USA	Experiment	Youtube	affective	Besides user selectivity and homogeneous networking, algorithm-based recommender systems seem to function as a structural factor promoting polarization by providing confirmatory information and thus reinforcing prior predispositions.
Hwang, H., Kim, Y., & Huh, C. U. (2014). Seeing is believing: Effects of uncivil online debate on political polarization and expectations of deliberation. doi: 10.1080/08838151.2014.966365	USA	Experiment	Youtube	ideological	Exposure to uncivil online discussion in which commenters uncivilly attacked the other side of the issue did not affect participants' attitude polarization.

Title	Country	Method	S.M. Platform	Type of Polarization	Results on Polarization
Johnson, B. K., Neo, R. L., Heijnen, M. E. M., Smits, L., & van Veen, C. (2020). Issues, involvement, and influence: Effects of selective exposure and sharing on polarization and participation. 10.1016/j.chb.2019.09.031	Netherlands	Experiment	Facebook	ideological	Pro- and counter-attitudinal information exposure did not have any main effects on opinion polarization and political participation. Sharing of pro-attitudinal news articles about refugees and equal pay had positive effects on opinion polarization. Also, the sharing of counter-attitudinal news articles about refugees and equal pay were negatively associated with opinion polarization.
Johnson, T. J., Kaye, B. K., & Lee, A. M. (2017). Blinded by the spite? Path model of political attitudes, selectivity, and social media. doi: 10.1080/15456870.2017.1324454	USA	Survey	Social Media	affective	Social media influence selective exposure and selective avoidance and political polarization.The only exception is that reliance on blogs does not lead to political polarization. The indirect effects model shows that strength of party ties and reliance influence confidence in the president and Congress indirectly through polarization and selective exposure through social media use.
Karlsen, R., Steen-Johnsen, K., Wollebæk, D., & Enjolras, B. (2017). Echo chamber and trench warfare dynamics in online debates. doi: 10.1177/0267323117695734	Norway	Experiment	Facebook, Twitter	ideological	Both discussing with opponents and supporters might lead to a reinforcement of the original opinion. Effects are stronger for individuals with strong attitudes than individuals with moderate attitudes.
Kibet, A., & Ward, S. (2018). Socially networked heterogeneity: The influence of whatsapp as a social networking site on polarisation in kenya. doi: 10.1080/23743670.2018.1537979	Kenya	Survey	Whats App	ideological, affective	Respondent's age and commenting on news reduces all forms of polarization, whereas, Ethnic identity, Political discussion, Class ideology, and Social network heterogeneity increase the three types of polarization.

Title	Country	Method	S.M. Platform	Type of Polarization	Results on Polarization
Kim, Y., & Kim, Y. (2019). Incivility on Facebook and political polarization: The mediating role of seeking further comments and negative emotion. doi: 10.1016/j.chb.2019.05.022	USA	Experiment	Facebook	ideological	Whether supporting evidence is provided or not in comments did not have any significant effect on respondents' ideological polarization.
Kobayashi, T. (2020). Depolarization through social media use: Evidence from dual identifiers in Hong Kong. doi: 10.1177/1461444820910124	Hong Kong	Survey & behavioral data	Social Media	affective, ideological, partisan	The political use of social media polarizes the attitudes and affects of single identifiers, whereas it has depolarizing effects on dual identifiers.
Lee, C., Shin, J., & Hong, A. (2018). Does social media use really make people politically polarized? Direct and indirect effects of social media use on political polarization in South Korea. doi: 10.1016/j.tele.2017.11.005	South Korea	Survey	Social Media	partisan	Social media did not directly push users into political extremes, whether they were neutrals or moderate partisans but indirectly affected polarization through increased political engagement. Social media use was negatively associated with a shift of neutrals toward a conservative view and positively associated with a shift toward a liberal view. Social media was not related to polarization in any direction for moderate partisans.
Lee, F. L. F. (2016). Impact of social media on opinion polarization in varying times. doi: 10.1177/2057047315617763	Hong-Kong	Survey (3)	Social Media	ideological	Political communication via social media can contribute to political polarization when the context itself is polarizing.
Lee, J., & Choi, Y. (2020). Effects of network heterogeneity on social media on opinion polarization among South Koreans: Focusing on fear and political orientation. doi: 10.1177/1748048518820499	South Korea	Survey	Social Media	ideological	Exposure to diverse opinions on social media hat potential to mitigate polarization. Individuals who feel disadvantaged or excluded from dominant position and those who are afraid of others with opposing views may stick to their initial viewpoints.

Title	Country	Method	S.M. Platform	Type of Polarization	Results on Polarization
Lee, J.K., Choi, J., Kim, C. & Kim, Y. (2014). Social media, network heterogeneity, and opinion polarization. doi: 10.1111/jcom.12077	USA	Survey	Social Media	affective, ideological, partisan	The higher level of SNS diversity led to more partisanship polarization for individuals participating in more political discussions whereas it had almost no effects on partisanship polarization for those joining fewer political discussions. It showed a similar pattern for ideological polarization.
Lee, P. S. N., So, C. Y. K., Lee, F., Leung, L., & Chan, M. (2018). Social media and political partisanship – A subaltern public sphere's role in democracy. doi: 10.1016/j.tele.2018.06.007	Hong-Kong	Survey	Social Media	ideological	The frequent use of social media contributed to a positive assessment of "the importance of democratic development in Hong Kong" but a negative view of "the importance of national interest," "trust in the central government," "social situation of Hong Kong two decades after returning to China," and "the future of Hong Kong". Political orientation, demographic factors and mass media had stronger relationships with the stance toward political values and social issues than social media use.
Levy, R. (2021). Social media, news consumption, and polarization: Evidence from a field experiment. doi: 10.1257/aer.20191777	USA	Experiment	Facebook	ideological, affective	Exposure to counter-attitudinal news decreases negative attitudes toward the opposing political party. No evidence that the political leanings of news outlets affect political opinions; Exposure to pro-attitudinal news increases affective polarization compared to counter-attitudinal news.
Lu, Y., Ray, R., Ha, L., & Chen, P. (2020). Social media news consumption and opinion	USA	Survey	Social Media	ideological	Time spent on social media is indirectly associated with opinion polarization on China's trade practice. Social media had

Title	Country	Method	S.M. Platform	Type of Polarization	Results on Polarization
polarization on China's trade practices: Evidence from a U.S. national survey.					a mediationg effect among those who frequently encounter like-minded information.
Min, H., & Yun, S. (2018). Selective exposure and political polarization of public opinion on the presidential impeachment in south korea: Facebook vs. Kakaotalk. doi: 10.29152/KOIKS. 2018.49.1.137	South Korea	Survey	Facbook, Kakao Talk	affective	It shows that politically motivated selective exposure predicted a significant increase in political polarization. In addition, education, political interest, and political ideology strength were also positively related to political polarization. The online activity on KakaoTalk has more of an impact on the increase of political polarization than online activity on Facebook.
Munger, K., Egan, P. J., Nagler, J., Ronen, J., & Tucker, J. (2020). Political knowledge and misinformation in the era of social media: Evidence from the 2015 uk election. doi: 10.1017/S0007123420000198	USA	Experiments	Facebook, Twitter	affective	No polarization through partisan cklickbait headlines.
Nguyen, A., & Vu, H. T. (2019). Testing popular news discourse on the "echo chamber" effect: Does political polarisation occur among those relying on social media as their primary politics news source? doi: 10.5210/fm.v24i6.9632	EU	Survey	Social Media	ideological	There is little (no) evidence for an increase in polarisation of EU related attitudes through social media compared to traditional media.
Ohme, J. (2021). Algorithmic social media use and its relationship to attitude reinforcement and issue-specific political participation – The case of the 2015 European immigration movements. doi: 10.1080/19331681.2020.1805085	Denmark	Survey	Social Media	ideological	The likelihood to report attitude reinforcement was twice as high for citizens who strongly rely on social media than for those who strongly rely on nonalgorithmic or offline media. Algorithmic news selection in combination with the homophily of social media can affect po-

Title	Country	Method	S.M. Platform	Type of Polarization	Results on Polarization
					litical attitudes of citizens in a reinforcing way.
Semaan B. C. Robertson, S., Douglas, S. & Maruyama, M. (2014). Social media supporting political deliberation across multiple public spheres: towards depolarization. doi: 10.1145/2531602.2531605	USA	Survey – qualitative	Social Media	/	We found that people's broader online interactions were depolarized in nature.
Shmargad, Y., & Klar, S. (2019). How partisan online environments shape communication with political outgroups.	USA	Experiment	News in social online contexts	partisan	Those who are attentive to their social surroundings learn over time how to moderate in the face of diversity. When they are subsequently exposed to our outgroup treatment, they choose to share more moderate news articles. Those who are surrounded by echo chambers respond to political outgroups by clinging to their preexisting views and sharing these views with a disagreeing audience.
Shmargad, Y., & Klar, S. (2020). Sorting the news: How ranking by popularity polarizes our politics. doi: 10.1080/10584609.2020.1713267	USA	Experiment (2)	News in social online contexts	partisan	No evidence that sorting news articles by popularity increases affective polarization.
Suk, J., Shah, D. V., Wells, C., Wagner, M. W., Friedland, L. A., Cramer, K. J., Hughes, C., & Franklin, C. (2020). Do improving conditions harden partisan preferences? Lived experiences, imagined communities, and polarized evaluations. doi: 10.1093/ijpor/edz051	USA	Survey	Social Media	partisan	Partisans' use of digital media strengthened their polarizing attitudes toward Obama evaluations.