Echo Chambers without Conversation?

Enriching Research on Polarization and Fragmentation on Twitter with the Analysis of Reciprocal Communication

Felix Gumbert, Ole Pütz, Florian Muhle and Robert Ackland[1]

Contribution to the Ad Hoc Groups »Unlikely Allies? Notwendigkeit, Potentiale und Implikationen der Integration von Computational Social Science und qualitativen Methodologien« and »Interdisziplinäre Erkundung technisch unterstützter (De-)Polarisierung«

Polarization and Social Media

The election of Donald Trump as US president in 2016 led to extensive public and academic debate about the role of social media in processes of societal polarization. Early hopes that the Internet would promote deliberation because it enables global connections have long since given way to fears that social media contribute to the fragmentation of the public sphere and associated polarization effects, thereby undermining the foundations of democratic consensus-building. The discussion focuses on two metaphors in particular: echo chambers and filter bubbles (Rau and Stier 2019; Habermas 2021; Terren and Borge 2021; Ludwig and Müller 2022).

The metaphor of the filter bubble (Pariser 2011) describes the algorithmic curation of content by search engines and social media platforms according to the principle of "more of the same". The metaphor of the echo chamber (Sunstein 2007) refers to users' possibilities to self-select information and connect with other like-minded people, which could result in the formation of homophilic clusters, i.e., a network of people sharing similar beliefs and opinions who all have access to the same information that confirms their beliefs. Both metaphors thus refer to the emergence of (relatively) closed information environments comprising individuals and groups, resulting in fragmented sub-publics with little contact with each other. According to the fragmentation thesis, the lack of confrontation with others and repeated confirmation of one's own opinions and beliefs leads to their solidification or even radicalization. The outcome is increasing divergence in political attitudes, opinions, and beliefs (ideological polarization), as well as an increased aversion towards other groups (affective polarization).

There is considerable academic debate as to whether it is possible to empirically observe fragmented sub-publics with a political and affective polarization of users. In this paper we introduce a novel

[1] This contribution is based on two presentations given at the 41st Congress of the German Sociological Association 2022 and was written in the context of the authors' joint work in the research project „Bots Building Bridges: Theoretical, Empirical, and Technological Foundations for Systems that Monitor and Support Political Deliberation Online", funded by the Volkswagen Foundation.

empirical approach to investigating echo chambers that includes collecting and analyzing threaded conversations instead of hashtag-based corpora. First, we outline what we consider the blind spots to be in the conceptualization of and empirical research on echo chambers. The notion of an echo chamber, or more generally of a homophilic cluster, refers to a structure consisting of interactions between individual users that are mediated and regulated by social media platforms such as Facebook, Instagram, TikTok, Reddit, and Twitter. In addition to particular platform-specific forms of interaction such as likes or friend requests, there are also forms of interaction which support reciprocal communication, i.e., conversations or discussions between two or more users in the narrow sense. However, research on fragmentation and polarization largely focuses on isolated messages in aggregated form and not on messages and the responses they receive. This means that forms of interaction that are potentially important for political opinion-forming processes are not examined, although it is precisely here that it would be possible to determine whether users see their own opinions confirmed by other users, or whether they are confronted with dissenting opinions, and how they respond to them. Accordingly, in the second part of this paper we present an argument that opinion formation should be investigated in the context of threaded reciprocal communication through replies, and not via hashtag-based corpora alone. In the final part of the paper, we illustrate the collection of reciprocal communication during the first debate of the 2020 U.S. presidential election via the Twitter API v2 (the #debate2020 dataset).

Empirical Approaches to Investigating Echo Chambers

The empirical research literature on social media and polarization, and on the existence and effects of echo chambers in particular, yields quite mixed results, with some studies finding evidence for the existence of echo chambers and others not. Different findings arise in part from the diversity of the methods applied and the concepts the respective studies are based on. In simplified terms, studies based on network analysis of digital trace data of social media platforms tend to find evidence of fragmentation in the form of homophilic clusters with limited connections to others in the network, which is interpreted as polarization (Conover et al. 2011; Crupi et al. 2022; Gruzd and Roy 2014; Williams et al. 2015). On the other hand, studies based on surveys reveal that social media is usually just one component of the users' larger media ecology; that news consumption is largely based on a broader spectrum of established news outlets; and that people using social media are more likely to be (involuntarily) confronted with different opinions, all of which seems to undermine the notion of digital echo chambers (Bruns 2019; Dubois and Blank 2018; Flaxman et al. 2016; Fletcher and Nielsen 2018; Gentzkow and Shapiro 2011; Newman et al. 2016; Rau and Stier 2019).[2] A large proportion of empirical studies examines the messaging service Twitter using network analytic methods, as is evident from the systematic reviews that have been published in recent years (Arora et al. 2022; Iandoli et al. 2021; Lorenz-Spreen et al. 2022; Ludwig and Müller 2022; Rau and Stier 2019). In many cases, the basic approach in Twitter research is similar: the authors create a list of hashtags or keywords related to specific (mostly) controversial topics or events, which then allows the collection of tweets that contain these words via Twitter's API. The resulting dataset can then be analyzed via a set of methods, which makes it possible to visualize and characterize the formation of more or less isolated clusters and to interpret them as indicators of fragmentation or polarization. To illustrate this kind of research, we provide two examples. These examples

[2] Dubois and Blank (2018) point out that even those who visit politically extreme news websites are more likely than the average user to visit mainstream/centrist websites as well and seem to engage with them the most. Yet, the authors also note that there is a chance for a small portion of the population to be trapped in an echo chamber. However, this small portion is more likely to be politically disinterested.

also show that the findings on fragmentation and polarization depend on the form of interaction that is used to construct networks.

The often-cited study by Barberá et al. (2015) focuses on "ideological polarization" on Twitter. Using a list of keywords on twelve different topics and events, the authors collected about 150 million tweets via Twitter's API. Second, the authors created a list of Twitter accounts with known political leanings such as all members of Congress or certain media outlets (plus more accounts in later stages of the research process). This list allowed them to project the position of users within the dataset onto an "ideological latent space". Depending on which accounts with a known political orientation users follow, they are assigned to the liberal or conservative side of the political spectrum. To measure ideological polarization, the authors analyzed the retweet networks (i.e., sharing content from other accounts with one's own followers via the retweet function). The authors observed polarization when the majority of tweets are retweeted by users who are located at a similar or the same position on the ideological spectrum as the original author (liberals retweet liberals and vice versa). However, such clustering is only possible for political topics such as Barack Obama's State of the Union Address, it is not evident for topics such as the 2014 Winter Olympics. This impression is also confirmed by the visualizations of the retweet networks: For this purpose, users who retweet each other more frequently were placed closer together than users who retweet each other only rarely or not at all. For political topics, two large ideologically homophilic clusters crystallized, with little exchange between them, while there was no larger and clearly identifiable ideologically homophilic clustering for other topics. Barberá et al. (2015) show that fragmentation on Twitter is dependent on the issues at hand. There is an observable degree of fragmentation in the retweet networks when it comes to political issues while the same cannot be said for non-political topics. Furthermore, the authors demonstrate that the degree of fragmentation can be dynamic, i.e., retweet networks on the same topic may become fragmented over time.

In the second example, de Franca et al. (2021) analyzed interactions on Twitter after the Brazilian Ministry of Health authorized the use of chloroquine for the treatment of COVID-19, despite the advice of experts to the contrary and a lack of scientific evidence. The authors collected 314,457 tweets containing the Portuguese keyword *cloroquina* authored within a 26-hour timeframe. After further filtering and processing, they created retweet and reply networks, and labeled the users in the networks manually as Bolsonarist (in agreement with the government), Progressist (in disagreement with the government), or Moderate (in disagreement with the government on this topic) by considering the tweet histories of the users. Interestingly, while the retweet network shows two main clusters of Bolsonarists on the one hand and a mix of Progressists and Moderates on the other, the reply network shows that replies are often made by users with an opposing political stance, i.e., replies to tweets by Bolsonarists were authored predominantly by Progressists or Moderates and vice versa, attacking the other side.

So, what do these mixed findings tell us about fragmentation and polarization? In an early study, Conover et al. (2011) already showed that retweet networks tend to show a higher degree of homophily than mention networks. This is in line with Barberá et al. (2015), who show that the degree of fragmentation in retweet networks is itself dependent on time and the topic at hand, which suggests that users who are part of politically homophilic clusters for political topics can simultaneously be part of heterogeneous clusters for other topics. De Franca et al. (2021) further show that, while the retweet network suggests homophily, the reply network suggests otherwise. Instead of a confirmation of political opinions as the echo chamber hypothesis implies, we find contestation when we take replies into account. It follows that measuring homophily in retweet networks alone is likely to overestimate the degree of fragmentation amongst users on Twitter.

Reciprocal Communication in Homophilic Networks

Existing research often speaks of 'discussion' or 'conversation', which evokes associations of dialogic events in which the participants exchange their (political) views and alternately assume the roles of speakers and listeners over a sequence of reciprocal messages. When retweet networks are studied, however, the focus is not on the analysis of reciprocal communication, but on processes of one-way information diffusion (Muhle et al. 2018). We argue that the technical affordances of different forms of interaction on Twitter need to be taken seriously; by definition a retweet does not allow a user to comment on the message that they disseminate to their followers; a retweet does not support reciprocal references that are necessary for conversations or discussions in a narrow sense; reciprocal references to original tweets are only supported by quote tweets, replies, or by mentions by other users. Stated differently, it is not surprising that users retweet content from other users with similar political views because retweets do not support the expression of dissent. Retweets may be understood as affirmation of a point of view (Crupi et al. 2022; Garimella et al. 2016) and thus can be understood as an indicator of homophily, but they are not the only form of interaction on Twitter.[3]

In our opinion, it does not seem very useful to study homophily without considering different forms of interaction on Twitter. The problem of the echo chamber is not primarily a problem of the dissemination of information, it is foremost a problem of the formation and solidification of uniform positions, which close themselves off from deviations. Taking this process seriously as something that requires investigation, it becomes important to consider how tweets refer to each other to examine a) whether patterns of mutual affirmation emerge and solidify and b) whether dissenting positions are rejected and delegitimized. Methodologically, this requires a shift from the focus on isolated messages and processes of information diffusion to sequences of related messages that allow us to observe discussions.

From Hashtag-based Corpora to Threaded Conversations: The Example of the #debate2020 Dataset

We have argued that our understanding of fragmentation and polarization on social media would benefit from an analysis of reciprocal communication. On Twitter, where every message is in principle visible to everybody across the globe, there are basically two ways to construct a dataset (that can also be combined): beginning with a set of users and who interacts with them, or beginning with a set of keywords and who uses them.[4] The majority of network analytical research on Twitter is based on large datasets of tweets that must contain at least one of a list of prespecified keywords or hashtags. Thus, the relevance of the collected tweets to the selected topic is given. However, it is at the same time unclear how many users use other hashtags to refer to the chosen topic, or do not use hashtags at all (Moon et al. 2016). This means that hashtag-based data collections omit an unknown number of potentially relevant posts and the discussions that are captured through replies or mentions are only partially observable if not all participants use hashtags in all of their tweets. With the introduction of the *conversation_id* by Twitter in 2020 as part of a tweet's metadata, it is now possible to assess how much reciprocal communication in the form of replies is missed by hashtag-based collections.

[3] According to a study by the PEW research center (2022) among adult Americans, 33% of tweets are about politics. Of these political tweets, 4% are original tweets, 9% quote tweets, 25% replies and 62% retweets.

[4] There are also projects that attempt to capture all traffic where a specific language is used. But this is only possible for languages where the number of Twitter users that use the language is relatively small, such as German.

Twitter defines "conversation threads" as consisting of an original tweet, all replies to that original tweet, as well as all replies to those replies. This can be understood as a tree structure that is instantiated when users use the reply function to respond to an original tweet, a tree that grows with every additional reply. The original tweet acts as the root of this tree, and all replies as well as replies to replies act as branches of this tree, all sharing one conversation_id. The advantage over a purely hashtag-based collection is that not all the tweets collected in this way need to contain one of the hashtags from the list; it is sufficient that just one tweet within the entire tree meets this criterion. Before the introduction of the conversation_id, conversations had to be constructed through other means, such as *mentions* (Muhle et al. 2018) or the *reply_to_tweet_id*, which only links pairs of tweets (Scheffler 2017; Moon et al. 2016). A method that relies on the conversation_id is not only easier to implement, it also provides us with a broader and more complete picture of reciprocal communication on Twitter, thus shedding light on the blind spots of existing research.

To elaborate this point by practical example, we use a dataset we collected during the 2020 US presidential election, the #debate2020 dataset. The #debate2020 dataset aims to provide insight into the discussions on Twitter that took place during the first TV debate between President Donald Trump and his contender Joe Biden in September 2020. The starting point of data collection was a list of target hashtags related to the TV debate.[5] Using Twitter's Search API, tweets authored during the debate period (plus 15 minutes before and after the debate) that contained at least one of the target hashtags were collected, resulting in a dataset containing 2,387,587 tweets, 28,562 of which were reply tweets (Gertzel 2021a). The Twitter API v2 was then used to identify the conversation_id for each of these reply tweets, and we further collected all reply tweets for each conversation_id, resulting in a set of *reply trees.*[6] This step resulted in an additional 2,590,102 reply tweets. By using the conversation_id, we were able to increase the total number of tweets by a factor of 2 from roughly 2.3 million tweets (hashtag-based collection) to nearly 5 million tweets (hashtag-based plus conversation_id-based collection).[7] More importantly, we were able to increase the number of replies by a factor of 92 from 28,562 to a total of 2,618,664. This means that our initial hashtag-based collection only includes 1.1% of replies that were written by users as part of Twitter conversations pertaining to the TV debate. We stated above that we do not know how many users do not use hashtags, and we can now provide an answer to this question by suggesting that a large majority of debate-related Twitter reply activity did not involve the use of our target hashtags; up to 98.9% of this kind of reciprocal communication would have been missed by a collection based on our target hashtags alone.[8] Although these numbers are based on one example and concern a topic from US politics that most likely incites more discussion than others, the discrepancy between 98.9% and 1.1% is so large that it seems obvious that replies are both a largely unexplored phenomenon and more common than previous research assumed (Muhle et al. 2018, p. 623).

5 #PresidentialDebates, #Election2020, #Debates2020, #Debates, #DebateNight, #Biden, #Biden2020, #BidenHarris2020, #Trump, #Trump2020, #PresidentialDebate.

6 The Twitter reply trees were collected using the voson.tcn software (via the Twitter API v2), as described in Gertzel (2021b).

7 By comparison, Moon et al. (2016), who collected data on discussions by Australian twitter users about the ride-sharing company Uber, were able to increase their dataset by 17% (or a factor of 0.17) using the reply_to_tweet_id.

8 We note that it is possible that during the TV debate a user authored a reply tweet containing one of our target hashtags, but that this tweet was in fact a reply to a tweet authored before the TV debate. Yet, our expectation is that most of the replies collected using the conversation_id approach are part of conversations that were taking place during the TV debate or are otherwise related to the debate as at least one person contributed to the reply tree during the debate. We also note that the data set could be further extended by using all conversation_ids from the initial hashtag-based collection. This means that we would miss out on reply trees if no reply contained one of the target hashtags.

So, what is it that we can learn about reciprocal communication on Twitter from replies? A reply is a response to the immediately preceding message and is visualized as such by Twitter. Every reply is a form of reciprocal communication, and every reply is part of a reply tree. Our dataset contains 13,119 reply trees ranging in size from 2 to 55,636 (median=20, mean=243), with the number of unique users authoring replies ranging from 1 to 47,082 (median=16, mean=191). We assume that the branches of a tree are relatively independent of each other. This means that a reply tree does not form a large conversation where everybody talks with everybody, as suggested by Twitter's naming of the conversation_id, but rather, starting from an original tweet, dialogical interactions split off with each reply, and this is where conversations or discussions take place. We call these interactions *reply chains.* Each reply chain is formed by the shortest path from the stem (original tweet) to the outermost leaf, which is a reply that did not receive a further response (root-to-leaf approach).[9]

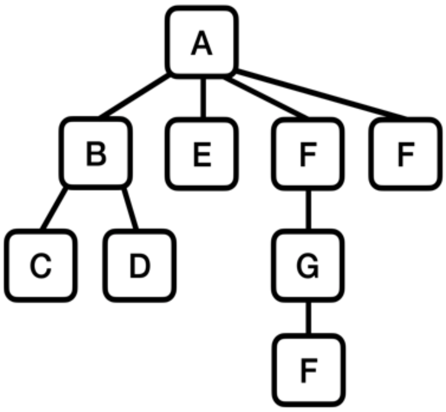

Figure 1: Exemplary reply tree

In the example of a reply tree in Figure 1, five reply chains could be formed in this way, each starting with the original tweet by A. From the entire #debate2020 dataset, 2.1 million reply chains can be reconstructed. The majority of the reply chains (87.4%) consist of only 2 turns, i.e., an original tweet and a direct reply to it from another user. The longest reply chain in our dataset consists of 75 turns.[10]

Table 1: Grouped distribution of reply chain lengths by turn

Turn length	2	3	4–6	7–15	16–35	>35
Frequency	1,852,940	177,157	72,577	14,715	1,720	144
	87.4 %	8.36 %	3.42 %	0.69 %	0.08 %	0.01 %

Outlook

In this contribution, we argued that the dominant network analytic approaches concerned with fragmentation and polarization focus on isolated messages in aggregated form. Findings partially depend on the form of interaction that is used to construct networks and here there is a common blind spot: The analysis of homophilic clusters through retweet networks does not take reciprocal communication via replies into account. Reciprocal communication describes the possibility for users to directly address one another, articulate (political) opinions, and express agreement or disagreement with them. A retweet does not allow a user to comment on the message that is being disseminated and thereby does not support the reciprocity that is necessary for conversations or discussions; reciprocal references to original tweets are only supported by quote tweets, replies, or by mentions by other users. It follows that we cannot rely on a single form of interaction such as retweets, but we must consider other forms

[9] Users that participate in one reply chain may be aware of the existence of further chains, but references to other reply chains are unlikely. It is of course possible that users participate in multiple reply chains, as in the example in Figure 1 where F responds twice to A. It is even possible that the same users participate in different chains in parallel, but these cases appear to be rare.

[10] In order not to distort the degree of reciprocity in reply trees, we calculate the length of reply chains not by the number of replies but by the number of turns at talk, which means that multiple continuous reply tweets from the same user are combined into a single turn, so that a turn can be made up of multiple replies where users reply to themselves.

of interaction such as replies. For this purpose, we introduced our approach for the construction of the #debate2020 dataset, which allows us to study discursive dynamics by systematically focusing on replies.

Data collection often relies on keywords or hashtags, which raises the question how many users do not use hashtags when they reply to a message. For the #debate2020 dataset we discovered that only a fraction of replies included the target hashtags. Our initial hashtag-based collection only included 1.1% of replies that can be deemed as debate-related, and 98.9% of reciprocal communication through replies would have been missed by a collection based on hashtags alone. This makes it important to complement a hashtag-based collection with a collection based on Twitter's conversation_id. The conversation_id yields all replies to an original tweet in a tree structure, and we suggested that reciprocal communication takes place in the reply chains formed by the shortest path from the stem (original tweet) to the outermost leaf (the last reply).

As a concept, echo chambers can be understood as fragmented sub-publics that are characterized by a repeated confirmation of existing political opinions and a lack of confrontation with other opinions, a process that leads to the solidification and radicalization of beliefs (Sunstein 2007). In the light of the mixed empirical results, some authors suggest that it is not sufficient to analyze echo chambers through the absence of opinion-contradicting information alone. Rather, we must consider how people react to other opinions and interact with those that articulate them (Bright et al. 2021; Karlsen et al. 2017; Nguyen 2020; Oswald and Bright 2022; Törnberg and Törnberg 2022). While this is a conceptual improvement from our perspective, studies that consider reciprocal communication empirically are still lacking. What requires investigation is a) whether patterns of mutual affirmation emerge and solidify and b) whether dissenting positions are rejected and delegitimized. It is not only that our understanding of fragmentation and polarization would be improved by an analysis of these communication patterns, but such an analysis would also allow us to conclude that a cluster which does not exhibit such patterns empirically should not be considered an echo chamber.

An important step in this direction is the analysis of reciprocal communication in reply chains with a focus on discussions. Here, we can observe a) whether there is mutual affirmation of opinions: If political opinions that conform with a user's opinion are accepted without discussion, this would indicate the existence of solidified beliefs within a sub-public. At the same time, if there are discussions observable among users with similar beliefs, this would serve as an indicator contradicting such fragmentation. Furthermore, the analysis of reply chains can also shed light on (b) the confrontation with other opinions: If political opinions that deviate from a user's opinions are rejected outright, this would serve as an indicator of polarization. At the same time, if discussions among users with divergent opinions are observable, this would serve as an indicator contradicting polarization. In such discussions, we would expect that opinions are assessed, arguments countered, and that we then can observe whether original opinions are altered or whether users respond with further arguments. Such conflictual discussions need not end with mutual agreement, their sheer existence would speak against the formation of an echo chamber.[11]

[11] There is much to learn about reciprocal communication on social media. Empirically, we approach the analysis of reply chains by applying a neutral understanding of conflictual discussions based on conversation analysis. Here, conflicts are understood as a sequence of at least two disagreements by at least two participants (Vuchinich 1990; Messmer 2003). A conflict emerges when users mutually articulate that they are "not on the same side" on a given issue. This can be done in various ways: by formulating an argument, providing personal experiences or opinions, or by accusing, attacking, or insulting each other. We describe this as a neutral understanding of conflictual discussions because it does not assume that arguments have to be of a certain quality or that participants remain civil when they interact, while also allowing us to distinguish and compare different types of conflicts or conflict trajectories.

Bibliography

Arora, Swapan Deep, Guninder Pal Singh, Anirban Chakraborty, and Moutusy Maity. 2022. Polarization and social media: A systematic review and research agenda. *Technological Forecasting & Social Change* 183:1–17.

Barberá, Pablo, John T. Jost, Jonathan Nagler, Joshua A. Tucker, and Richard Bonneau. 2015. Tweeting from Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science* 26(10):1–12.

Bright, Jonathan, Nahema Marchal, Bharath Ganesh, and Stevan Rudinac. 2021. How Do Individuals in a Radical Echo Chamber React to Opposing Views? Evidence from a Content Analysis of Stormfront. *Human Communication Research* 48(1):1–31.

Bruns, Axel. 2019. Filter bubble. *Internet Policy Review* 8(4):1–14.

Conover, Michael D., Jacob Ratkiewicz, Matthew Francisco, Bruno Goncalves, Filippo Menczer, and Alessandro Flammini. 2011. Political polarization on Twitter. *Fifth International Conference on Weblogs and Social media* 5(1):89–96.

Crupi, Giuseppe, Yelena Mejova, Michele Tizzani, Daniela Paolotti, and André Panisson. 2022. Echoes through Time: Evolution of the Italian COVID-19 Vaccination Debate. *Proceedings of the Sixteenth International AAAI Conference on Web and Social Media* 16(1):102–113.

de Franca, Fabricio, Carlos Dos Santos, Claudio Penteado, Denise Goya, Lucas Mazim, Carlos Kamienski, Daniel Di Genova, and Diogo Ramos. 2021. Dynamics of Conflicts on the Twitter Social Network: A case study on the use of chloroquine in Brazil. *Proceedings of the 54th Hawaii International Conference on System Sciences*: 2976–2985.

Dubois, Elizabeth, and Grant Blank. 2018. The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, Communication & Society* 21(5):729–745.

Flaxman, Seth, and Sharad Goel Justin M. Rao. 2016. Filter bubbles, echo chambers, and online news. *Public Opinion Quarterly* 80 (Special Issue):298–320.

Fletcher, Richard, and Rasmus Kleis Nielsen. 2018. Are people incidentally exposed to news on social media? A comparative analysis. *new media & society* 20(7):2450–2468.

Garimella, Kiran, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. 2016. Quantifying Controversy in Social Media. *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining* (WSDM '16):33–42.

Gentzkow, Matthew, and Jesse M. Shapiro. 2011. Ideological Segregation Online and Offline. *The Quarterly Journal of Economics* 126(4):1799–1839.

Gertzel, Bryan. 2021a. #DebateNight 2020: Hashtag Twitter Collection of the US Presidential Debates. VOSON Lab (Australian National University) report. https://vosonlab.github.io/posts/2021-06-03-us-presidential-debates-2020-twitter-collection/ (accessed: 30 January 2023).

Gertzel, Bryan. 2021b. Twitter Conversation Networks. VOSON Lab (Australian National University) report. https://vosonlab.github.io/posts/2021-03-23-twitter-conversation-networks/ (accessed: 30 January 2023).

Gruzd, Anatoliy, and Jefferey Roy. 2014. Investigating political polarization on Twitter: A Canadian perspective. *Policy & Internet* 6(1):28–45.

Habermas, Jürgen. 2021. Überlegungen und Hypothesen zu einem erneuten Strukturwandel der politischen Öffentlichkeit. *Leviathan* 49(37):470–500.

Iandoli, Luca, Simonetta Primario, and Giuseppe Zollo. 2021. The impact of group polarization on the quality of online debate in social media: A systematic literature review. *Technological Forecasting & Social Change* 170:1–12.

Karlsen, Rune, Kari Steen-Johnsen, Dag Wollebæk, and Bernard Enjolras. 2017. Echo chamber and trench warfare dynamics in online debates. *European Journal of Communication* 32(3):257–273.

Lorenz-Spreen, Philipp, Lisa Oswald, Stephan Lewandowsky, and Ralph Hertwig. 2023. Systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature Human Behavior* 7:74–101.

Ludwig, Katharina, and Philipp Müller. 2022. Does Social Media Use Promote Political Mass Polarization. In *Questions of Communicative Change and Continuity. In Memory of Wolfram Peiser*, eds. Benjamin Krämer and Philipp Müller, 118–166. Baden-Baden: Nomos.

Messmer, Heinz. 2003. Konflikt und Konfliktepisode. Prozesse, Strukturen und Funktionen einer sozialen Form. *Zeitschrift für Soziologie* 32(2):98–122.

Moon, Brenda, Nicolas P. Suzor, and Ariadna Matamoros-Fernandez. 2016. Beyond hashtags: Collecting and analysing conversations on Twitter. *Association of Internet Researchers Annual Conference* 6 (AoIR 2016).

Muhle, Florian, Robert Ackland, and Timothy Graham. 2018. Socialbots in politischen Online-Konversationen. Eine (überschätzte) Gefahr für demokratische Willensbildung im Internet? *Zeitschrift für Parlamentsfragen* 49(3):618–638.

Newman, Nic, Richard Fletcher, David Levy, and Rasmus Kleis Nielsen. 2016. Reuters Institute Digital News Report 2016. https://ssrn.com/abstract=2796534 (accessed: 30 January 2023).

Nguyen, C. Thi. 2020. Echo chambers and epistemic bubbles. *Episteme* 17(2):141–61.

Oswald, Lisa, and Jonathan Bright. 2022. How Do Climate Change Skeptics Engage with Opposing Views Online? Evidence from a Major Climate Change Skeptic Forum on Reddit. *Environmental Communication* 16(6):805–821.

Pariser, Eli. 2011. *The filter bubble: What the Internet is hiding from you*. New York, NY: Penguin Press.

Pew Research Center. 2022. 5 facts about political tweets shared by U.S. adults. https://www.pewresearch.org/fact-tank/2022/06/30/5-facts-about-political-tweets-shared-by-u-s-adults/ (accessed: 30 January 2023).

Rau, Jan Philipp, and Sebastian Stier. 2019. Die Echokammer-Hypothese: Fragmentierung der Öffentlichkeit und politische Polarisierung durch digitale Medien? *Zeitschrift für Vergleichende Politikwissenschaft* 13:399–417.

Scheffler, Tatjana. 2017. Conversations on Twitter. In *Investigating computer-mediated communication: Corpus-based approaches to language in the digital world*, eds. Darja Fišer and Michael Beißwenger, 124–144. Ljubljana: University Press.

Sunstein, Cass R. 2007. *Republic.com 2.0*. Princeton, NJ: Princeton University Press.

Terren, Ludovic, and Rosa Borge. 2021. Echo Chambers on Social Media: A Systematic Review of the Literature. *Review of Communication Research* 9:99–118.

Törnberg, Petter, and Anton Törnberg. 2022. Inside a White Power echo chamber: Why fringe digital spaces are polarizing politics. *new media & society 0(0):1–23.*

Vuchinich, Samuel. 1990. The sequential organization of closing in verbal family conflict. In *Conflict talk. Sociolinguistic investigations of arguments in conversations*, ed. Allen D. Grimshaw, 118–138. Cambridge: University Press.

Williams, Hywel T.P., James R. McMurray, Tim Kurz, and F. Hugo Lambert. 2015. Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Environmental Change* 32:126–138.