Xia, Yan; Chen, Ted; Kivelä, Mikko

Spread of tweets in climate discussions

Spread of tweets in climate discussions

A case study of the 2019 Nobel Peace Prize announcement

Yan Xia,[I] Ted Hsuan Yun Chen,[II] & Mikko Kivelä[I]

[I] Department of Computer Science, Aalto University, Finland
[II] Department of Computer Science, Aalto University, Finland; Faculty of Social Sciences, University of Helsinki, Finland

Abstract

Characterising the spreading of ideas within echo chambers is essential for understanding polarisation. In this article, we explore the characteristics of popular and viral content in climate change discussions on Twitter around the 2019 announcement of the Nobel Peace Prize, where we find the retweet network of users to be polarised into two well-separated groups of activists and sceptics. Operationalising popularity as the number of retweets and virality as the spreading probability inferred using an independent cascade model, we find that the viral themes echo and differ from the popular themes in interesting ways. Most importantly, we find that the most viral themes in the two groups reflect different types of bonds that tie the community together, yet both function to enhance ingroup connections while repulsing outgroup engagement. With this, our study sheds light, from an information-spreading perspective, on the formation and upkeep of echo chambers in climate discussions.

Keywords: climate politics, Twitter, virality, polarisation, social networks

Introduction

Climate discussions on social media platforms are often structured into segregated echo chambers of activists and sceptics (Williams et al., 2015). At best, these echo chambers simply demonstrate that different opinions on climate politics exist; worse, they cause vicious cycles of increasing divisiveness (Asikainen et al., 2020; Baumann et al., 2020). Better knowledge of the generative features of polarised echo chambers can facilitate the design of mitigation institutions, especially on social media platforms where the information individuals are exposed to can be algorithmically manipulated (Musco et al., 2018; Nelimarkka et al., 2018). However, while prior work demonstrates the existence of these echo chambers,

we lack a nuanced understanding of the social processes that generate them. Prior studies point to attitude-based homophily whereby individuals limit their interactions to others with whom they share similar attitudes (Williams et al., 2015), or entrenched group cleavages resulting from issue alignment (Chen et al., 2020), as important features of climate discussion echo chambers, but these mechanisms alone cannot explain the complexity of the observed discussion networks.

On social media platforms where information sharing serves as a major channel of communication, a better understanding of the discussion dynamics could be achieved by observing how different types of information spread through the network, accordingly consolidating or attenuating the echo chambers. Thus, in order to better understand the sustained existence of echo chambers in climate discussions on social media platforms, we conduct a study on Twitter that examines climate discussion dynamics through the lens of tweet spreading, particularly what kind of content is most likely to be shared within and across different groups of users. We focus our examination on the cores of the echo chambers that comprise the most popular tweets and users, drawing on research showing that climate politics polarises following elite behaviour (Birch, 2020).

More specifically, given a division of the users into climate activists and climate sceptics, we examine what types of tweets spread within the echo chambers, what cross the boundary, and what characteristics of a tweet lead to more viral spreading on the user network. On the last point, in contrast to previous studies that quantified the virality of an item as the number of times it is shared (e.g., Berger & Milkman, 2012; Hansen et al., 2011), we measure virality while controlling for the underlying complex social network. This allows us to account for the number of times and the context under which a tweet has been seen.

We conduct our study on the climate discussions during the announcement of the 2019 Nobel Peace Prize. The event triggered extensive discussion on climate politics, because Greta Thunberg, the climate activist who mobilised demonstrations around the world, was recognised in the media as a likely winner (Adam, 2019; Carrington, 2019). The fact that Thunberg did not win spurred a host of new discussion points among climate change sceptics. For this reason, the event provides a suitable opportunity for our study, as it injected new information into extant discussion networks. During that period, multiple opinions, attitudes, and emotions from different user groups collided and merged in online social networks, thereby providing a rich context for studying the spreading dynamics of different types of information in climate discussions.

Our study makes a number of contributions. First, it adds to the body of work describing the content of climate discussions on Twitter (e.g., Cody et al., 2015; Dahal et al., 2019; Jang & Hart, 2015; Kirilenko & Stepchenkova, 2014). Second, it provides a finer picture of climate discussion dynamics through information spreading, furthering our understanding of echo chambers in online discussions (Barberá et al., 2015), especially on climate politics (Williams et al., 2015). Here, it also links more broadly to the literature on the relationship between filter bub-

bles and political polarisation, which is the concept that the feedback between behavioural selective exposure and algorithmically learned filtering of information will lead to societal polarisation (e.g., Dahlgren, 2021; Pariser, 2011). While our study does not consider the algorithmic side of these filter bubbles, we find strong evidence of selective engagement with ingroup members, and to the extent that outgroup information crosses group cleavages, it potentially contributes to a greater degree of inter-group hostility. Finally, our study complements previous work on the virality of online content (Berger & Milkman, 2012; Hansen et al., 2011; Stieglitz & Dang-Xuan, 2013) in adopting a network-aware measure of virality.

The remainder of this article proceeds in two steps. First, we explore the characteristics of popular tweets in each of the echo chambers using an iterative process. After collecting all tweets containing "climate" during the announcement period and identifying the activists and sceptics involved, we examine the characteristics of the discussion within the two groups using an initial reading of the most-popular tweets. Second, using statistical methods, we infer the virality of tweets from our sample and examine which of the previously identified characteristics most strongly predict likeliness to spread. We conclude by discussing the implications of our findings. Most importantly, our results show that the most virality-predicting features of a tweet are also ones that enhance ingroup ties while repulsing outgroup engagement, thus revealing the potential role of tweet spreading in exacerbating the polarisation on climate Twitter.

Exploring climate Twitter

While there have been a number of studies that describe climate discussions on Twitter (e.g., Cody et al., 2015; Dahal et al., 2019; Jang & Hart, 2015; Kirilenko & Stepchenkova, 2014), many of them are more than a few years old. The rapid rate of change of online behaviour means that many of these findings have decreased temporal validity (Munger, 2019). We therefore begin our study with an exploratory approach to understanding the present state of climate discussions. Our primary goal in this exercise is to identify popular themes in current climate discussions such that we can use them to study tweet spreading dynamics. Our approach is an iterative theory-building process, where prior literature sets our expectations for a first-cut exploratory reading of our data, which in turn informs our expectations about tweet virality and our codebook development.

Literature review

We begin our study with a literature review that informs our data exploration, focusing on identifying popular themes in climate discussion, how they differ between activist and sceptic groups, and more generally the determinants of information virality on online social network platforms. Earlier work on Twitter-

specific climate discussions tends to be more descriptive, with studies laying the groundwork by describing the geospatial distribution of tweets and variations in their content (e.g., Kirilenko & Stepchenkova, 2014). Jang and Hart (2015), for example, showed that among four English-speaking countries (the US, the UK, Canada, and Australia), there is considerable variation in how people tweet about climate change. Most relevant to polarisation, they found that within the US, Republican- and Democrat-leaning states exhibit subnational differences, with the former more likely to engage with the "hoax" frame. These results comport with studies showing that the climate issue is generally subject to partisan sorting (e.g., McCright & Dunlap, 2011). Related to this are observed discrepancies in references to "science". In their review of climate-related tweets, Cody and colleagues (2015) found that sceptics tend to use phrases that are less commonly used in conjunction with references to science. On the other hand, Pearce and colleagues (2014) – possibly due to the more science-focused nature of their case study – found that in discussions of IPCC reports, science was highly politicised. Some studies focused explicitly on affective polarisation (e.g., Tyagi et al., 2020), finding large variation between activist and sceptic groups in expressing negativity toward their respective outgroups.

Observed differences between groups are telling, but, with the exception of Williams and colleagues' (2015) work on homophily, these studies do not explicitly address the formation of echo chambers. Drawing on the elite-led polarisation literature (Birch, 2020), we propose that communities resembling echo chambers form when tweets (more generally content) from core members of the group spread to the periphery. This kind of spreading mechanism, coupled with homophily in the sharing network which reduces cross-group ties, should result in the kind of echo chamber structure that has been observed in climate discussions. The implication here is that without the spread of viral tweets from the core, the echo chambers will dissipate, making the study of what gets shared an important area of climate polarisation research.

Previous work in this area, beyond climate discussions, has shown that in online social networks, the spreading of content is strictly connected to the characteristics of the content itself (Guerini et al., 2012; Jenders et al., 2013), including types of content (Nagarajan et al., 2010), URLs, hashtags (Suh et al., 2010), and emotions (Berger & Milkman, 2012; Hansen et al., 2011; Stieglitz & Dang-Xuan, 2013). Specifically, the textual features of content largely predict its spreading on Digg (a news-based Internet sharing platform, Guerini et al., 2011), and the type of tweet (e.g., calling for action, information sharing) plays a vital role in shaping its retweet network on Twitter (Nagarajan et al., 2010). Meanwhile, others find URLs and hashtags to be the most important content features correlated with retweetability on Twitter (Suh et al., 2010). Hashtags have also been identified as a way individuals engage with a broader imagined community on social media platforms (Hanteer & Rossi, 2019). Various emotions (i.e., positive or negative sentiments, and more specifically awe, anger, and anxiety) in the

content also positively predict sharing behaviour of users on both *The New York Times* (Berger & Milkman, 2012) and Twitter (Hansen et al., 2011; Stieglitz & Dang-Xuan, 2013). Other than the content features, author features including account age, follower count, and followee count also affect how likely content is shared on Twitter (Jenders et al., 2013; Suh et al., 2010). Finally, information diffusion on social networks is influenced by their underlying network structures (Weng et al., 2013).

Data

To further explore the concepts identified in our literature review, we collected Twitter data on climate discussions. As noted, we focused on the climate discussion surrounding the 2019 Nobel Peace Prize announcements, because Greta Thunberg featured prominently as a likely winner. Using the Twitter streaming application programming interface (API) and a list of relevant keywords, we collected every tweet related to Greta Thunberg, climate politics, and the Nobel Peace Prize from 10–22 October 2019, which covers a short interval before and after the announcement of the 2019 Nobel Peace Prize on 11 October 2019. A total of 5,422,617 tweet records were collected, including records of both original tweets and retweets, and each record contains information of the text, author, and status of the tweet (e.g., whether it is an original tweet or a retweet). A total of 2,011,410 users were involved in the dataset.

Since our study aims to explore discussion dynamics with respect to climate change, we focused our examination on tweets including the substring "climate" (case-insensitive) in the text. We further restricted our examination to tweets written in English, and excluded replies to other tweets. Based on trends identified from prior work (Jang & Hart, 2015), we also selected and focused on the subset of users who have either posted or retweeted any tweet with hashtags that included "climate" in combination with either "crisis" or "hoax" (case-insensitive). Using the resulting set of 317,243 tweet records involving 24,770 users, we built a retweet network of users, with each user corresponding to a single node, and a single link connecting two nodes if there exists any retweet record going between the corresponding users (i.e., an undirected link exists between node A and node B if either user A retweeted any tweet posted by user B at least once, or user B retweeted any tweet posted by user A at least once).

Following prior work in this area (e.g., Chen et al., 2020; Garimella et al., 2018), we focused on only the largest connected component of the retweet network, which consists of 20,628 nodes (i.e., users) and 73,381 links, as visually shown in Figure 1; the disconnected components were excluded because there was not sufficient information for inferring their stances on the topic. Colours of the user nodes indicate climate activism or scepticism, classified using a network partitioning method. Following prior work (e.g., Garimella et al., 2018), we used the METIS partitioning algorithm, which finds the two groups with the

lowest intergroup connections (Karypis & Kumar, 1998). After the partitioning, the green bubble consists of 14,812 users classified as climate activists, and the orange bubble consists of 5,816 users classified as climate sceptics. A total of only 223 links going between the two bubbles indicates that our procedure yielded two echo chambers of climate activists and sceptics.

Figure 1   Partitioned retweet network
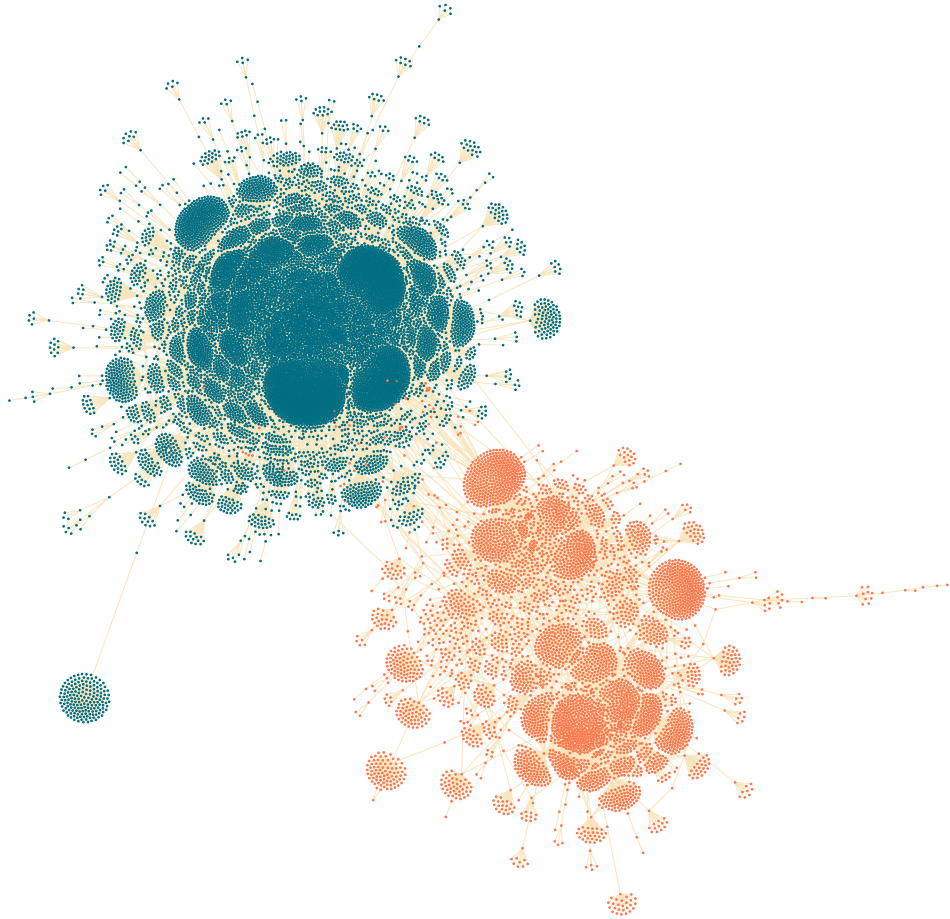
Comments: The partitioned retweet network of selected users and tweets, where the green bubble consists of 14, 812 users classified as climate activists, and the orange bubble consists of 5, 816 users classified as climate sceptics.

Within-group spreading tweets and popular themes

From the data described above, we prepared a subset of tweets for study. Because we are interested in understanding how tweets from popular users spread, we focused on the top 500 most-retweeted items in each group. These have ingroup retweet counts ranging from 66 to 1,762 in the activist group, and from 26 to 520 in the sceptic group.

We began our examination by identifying the most commonly occurring words in each group (counted maximum once per tweet), adjusting for how often they appear in the other group. These words are presented in Figure 2. There seems to be a clear theme of taking action among the activist discussions, with "crisis", "action", "#actonclimate", and "act" among the ten most common words. In this group, Greta Thunberg's Twitter account is frequently mentioned. In the sceptic group, there is frequent mention of actors and entities from the other side (e.g., "greta", "aoc", "protester", "un"). Our later reading shows that these mentions are evidence of outgroup attacks. The United Nations, for example, is often discussed from a negative, anti-internationalist perspective. The sceptics also commonly used "jet", "private", and "fly" to support their argument that activists are hypocritical about environmental protection. It is also interesting that the words "climate" and "change" appear much more in the sceptic group than in the activist group, because activists tend to use "climate crisis" and "#climatechange" instead of "climate change".

Figure 2   Most characteristic words, activist and sceptic groups
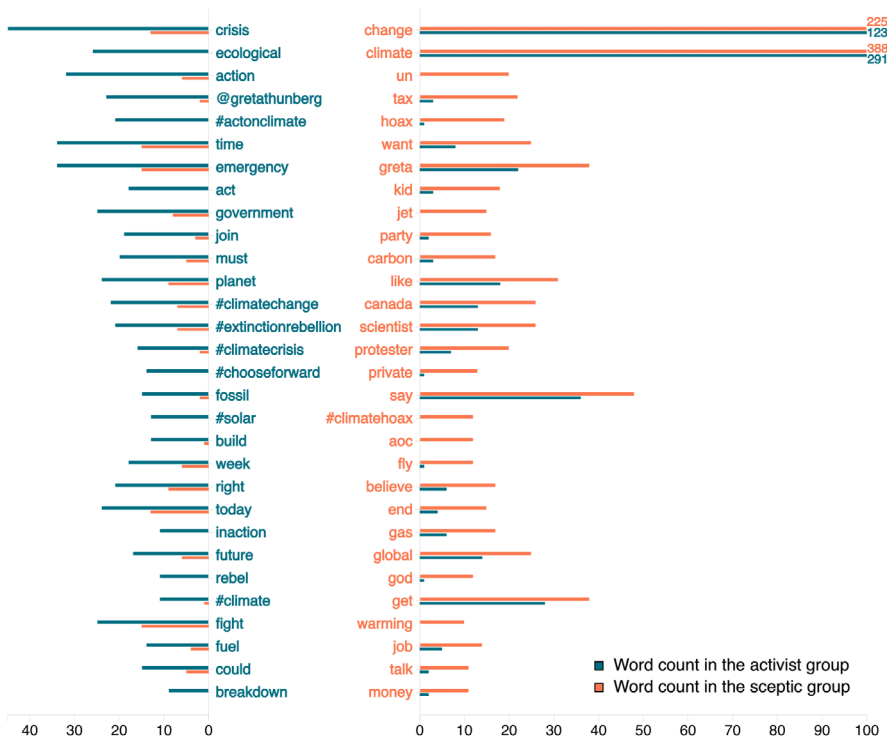
Comments: Most characteristic words among the top-500 retweeted tweets in the activist group (green) and the sceptic group (orange), ranked by the difference between word counts in the two groups. The bars of "change" and "climate", with exceptionally large word counts, are truncated for visualisation purposes, but the actual word counts are attached in number to the right of the bars.

Next, we conducted an exploratory reading of the tweet texts to identify commonly occurring themes and other characteristics using approximately half the

tweets in each group. This procedure was informed by our literature review, and as expected, we encountered many of the same themes as those outlined in prior research. However, our reading did not always comport with the existing literature. For example, whereas Cody and colleagues (2015) showed indirect evidence for discrepancies between activists and sceptics in references to "science", we found that both groups invoked the concept of science at a similar rate. This is surprising, as we initially expected to see more activists invoking scientific evidence that the climate crisis actually exists. Instead, the sceptic group also made frequent references to scientific evidence they claimed to disprove anthropogenic climate change. This observed discrepancy to prior work offers evidence of temporal shifts in the climate discussion, justifying our initially exploratory approach.

In general, the most significant outcome of our reading is the clear difference between the discussion styles and content in the two groups. The dominating sentiment in the sceptic group is negative and even aggressive, while the sentiment in the activist group is more positive. In terms of content, the most-popular tweets among the activists share the theme of "action": they either celebrated actions taken to tackle climate change – especially the pro-climate movements – or called for further action to address the climate crisis, potentially proposing concrete solutions. For example, the global climate strike led by Greta Thunberg was a popular topic among the activists. A large number of tweets are images or videos that demonstrate the strength of the movement or praise certain individuals who participated in the movement, such as the protesters (especially celebrities) who were arrested. The pro-climate speeches by Greta Thunberg and Alexandria Ocasio-Cortez, the US senator who spoke out in favour of strong climate policies, were also widely quoted and commended. Additionally, another set of tweets proposed solutions to address climate change and calls for action to implement them.

Among the sceptics, the most-popular tweets are dominated by the theme of "attack". Most of them are directed sharply at climate change supporters, either making fun of them or accusing them of being hypocritical by engaging in environmentally harmful practices. For example, we saw a large amount of attack and mocking toward climate activists including Greta Thunberg, Alexandria Ocasio-Cortez, and Jane Fonda. Other pro-climate celebrities were accused for flying in private jets. Aside from this, there are also recurring patterns of claiming that climate policy is a front for other agendas (e.g., money, control), speaking negatively of international organisations (e.g., the UN), or invoking "science" or "scientist" to back up an argument. With respect to the style of language used, there is a dominating theme of mocking tones, potentially also paired with the use of uncivil wording, emojis, or exclamation marks.

Based on this exploratory reading process, we identified a list of tweet features plausibly relevant to a tweet's virality. Table 1 contains these features along with the coding rule we developed for identifying them.

Yan Xia, Ted Hsuan Yun Chen, & Mikko Kivelä

Table 1   Tweet features for coding

|	Feature	Coding Rule
Universal (Style)	Mocking	Does the tweet make fun of an entity at its expense? (yes/no)
	Incivility	Does the tweet contain uncivil language? (yes/no)
Universal (Content)	Call to Action	Does the tweet call on others to behave in a certain way? (yes/no)
	Ingroup Praise	Does the tweet speak of the ingroup in a positive manner? (yes/no)
	Outgroup Criticism	Does the tweet speak of the outgroup in a negative manner? (yes/no)
	Science	Does the tweet invoke "science" or "facts" as support? (yes/no)
	Hashtags[a]	Machine extracted count of hashtags used in the tweet.
	Mentions[a]	Machine extracted count of users mentioned in the tweet.
Activists	Solutions	Does the tweet present solutions to addressing climate change? (yes/no)
	Movement	Does the tweet emphasise the strength of the pro-climate movement? (yes/no)
Sceptics	Anti-international	Does the tweet speak negatively of international organisations? (yes/no)
	Hypocrisy	Does the tweet claim that supporters are hypocritical or inconsistent? (yes/no)
	Conspiracy	Does the tweet claim that climate policy is a front for other agendas? (yes/no)

[a] Indicates features directly extracted from the data.

To conduct more systematic analysis on characteristics of the echo chamber cores, we prepared our final dataset by further filtering out tweets by users who authored fewer than three tweets. In this way, we define the core of the echo chambers as individuals who are consistently responsible for the popular tweets in either group. This filtering step also aids the later virality analysis because it allows us to compare the importance of features while holding user characteristics constant. For each of these tweets, we did a second reading, labelling them based on whether they contain each of our previously identified features following the coding rules outlined in Table 1. When encountering external links, we first decided whether the tweet endorses or criticises the external content. If it is endorsed, we took the external content as an extension of the tweet when labelling. We decided to adopt a manual coding approach for labelling the features of each tweet, with the observation that most features we selected to code involved a nuanced understanding of the tweet content. As much as lexicon-based and machine learning methods had been employed in previous studies to automatically detect simpler features of text – including positivity and negativity (Berger & Milkman, 2012; Hansen et al., 2011; Stieglitz & Dang-Xuan, 2013) – those methods are much less effective at recognising features that involve richer emotions and contextual

knowledge. Other studies that included similarly nuanced features, such as incivility, also relied entirely or in part on human coding (Berger & Milkman, 2012; Muddiman et al., 2019).

In this second step, tweets were labelled independently by all three authors, which allowed us to adjudicate disagreements using a vote. Overall, we saw a moderate level of agreement prior to voting, with the inter-rater reliability score Krippendorff's alpha being 0.63 and all three authors reaching a consensus on 86.2 per cent of the entries. Finally, in addition to these manually labelled features, we used computer text processing to directly extract the number of hashtags used and the number of other users mentioned in the tweet. Figure 3 shows the results of our labelling process. Our qualitative findings based on our exploratory reading, on the parts that rely on popularity of content, are confirmed by these quantitative results.

Figure 3   Percentage of tweets labelled with each feature, activist and sceptic groups

Percentage of labelled tweets
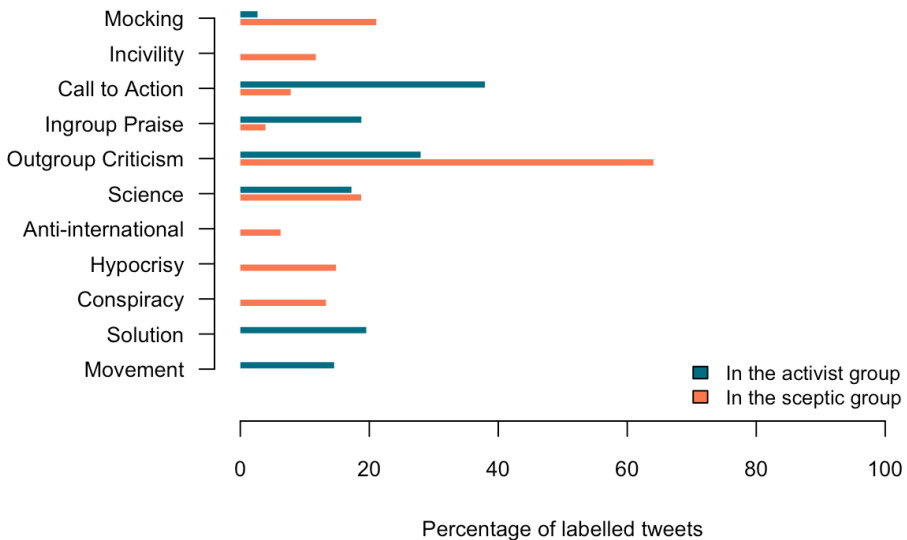
Cross-group spreading tweets

The largely isolated bubble structures apparent in Figure 1 already suggest that cross-group spreading is rare. We examined this more deeply by checking the number of tweets spreading in both groups. The results, shown in Figure 4 – which is a scatter plot of tweets with one axis being the number of retweeters in the activist group and the other axis being the number of retweeters in the sceptic group – indicate that most tweets in our data set mainly spread in only one group, and only nine tweets were retweeted more than ten times in both groups.

Figure 4   Scatter plot of tweets with respect to number of retweeters in each group
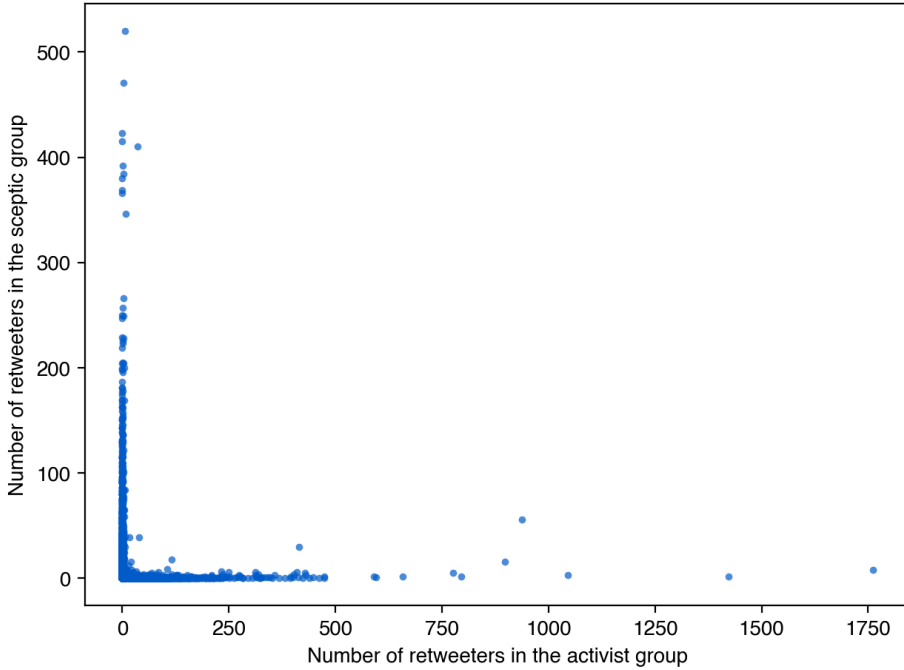
Then, we inspected the tweets that mainly spread in one group but also got a significant number of retweets from the other group to see if any meaningful pattern existed. Among these cross-group spreading tweets, the only recurring pattern we recognised is that they documented some less successful or more extreme approaches to protests taken by the climate activists (e.g., protesters being arrested, protesters burying their heads in the sand). Further, the direction of spread of these tweets ran from the activist group to the sceptic group. This finding is tentative, given the small sample of data, but we speculate that these tweets were "popular" with both groups because even without contextualisation via quote-retweeting, they allowed the respective groups to interpret it as they wanted. However, this cross-group information flow presumably further intensifies the conflict between activists and sceptics, instead of bridging the gap by fostering communication and mutual understanding.

Viral spreading on climate Twitter

Next, we study what kinds of content predict a tweet's viral spreading within the two groups. In doing so, we distinguish virality from the popularity measure in the preceding section. This is because the popularity of a tweet, as measured by the number of retweets, could be an inappropriate metric of how viral the tweet content actually is: there are a number of factors that influence how popular a tweet will become aside from its content. The tweet can be made by an authoritative

person or organisation, or it can start from a favourable location in the network, such as close to actively retweeting users or from an account with a large number of followers (Jenders et al., 2013; Suh et al., 2010). Further, the design choices Twitter has made on how to show tweets to users can have an effect on the popularity. In addition to these confounding factors, random fluctuations in tweeting behaviours are amplified in spreading processes, because the more the tweet is retweeted, the more people will see it and have possibility to retweet it further. This type of rich-get-richer phenomenon is well-known in spreading processes and can lead to situations where small changes in virality can lead to large changes in popularity or large fluctuations in the popularity of similar content (Barrat et al., 2008).

Here, we instead measure the *virality* of a tweet, as proportional to the probability that an average user retweets the tweet after seeing it. With this information, we can find out which type of content, as categorised by the labelling in Table 1, is more likely to spread within the two polarised groups.

Inferring tweet virality

Our goal is to infer the extent to which a tweet is retweeted for its content, apart from confounding contextual factors. These factors include user characteristics such as how authoritative the original tweeting account is and the retweeting account's retweeting tendencies. We also model how the tweet is shown to users. In the end, the virality score of a tweet is computed by finding the score value that best explains the whole process of the tweet spreading in the network.

The core of our model uses two pieces of information. The first – how often a tweet is *retweeted* and by whom – is directly observable from our data. The second – how often a tweet is *seen* and by whom – requires that we define an exposure pathway network based on who follows who and how tweets are shown to followers. In Twitter, a tweet is shown to users that follow the original tweeter or anyone retweeting the original tweet. To model this, we created a follower network with directed links from users to their followers. Such a follower network is not enough, however, because Twitter limits the times a user sees an original tweet, regardless of how many of the user's followees retweeted it. To gain the necessary understanding of the relationship between followee tweeting behaviours and what is displayed on the follower's timeline, we tested Twitter's algorithm for displaying tweets and retweets. Figure 5 contains our findings, which can be summarised in two rules:

1. If an account makes an original tweet, followers can only see the original tweet (i.e., no retweet notifications), regardless of whether others have retweeted it. No information about retweets of the original tweet will be available on the user's timeline.

2. If an original tweet is retweeted by one or more followees, the follower's timeline can only show the retweet notification from the first retweeting followee (and, by the first rule, only if the original tweeter is not a followee).

Figure 5   Twitter's timeline algorithm
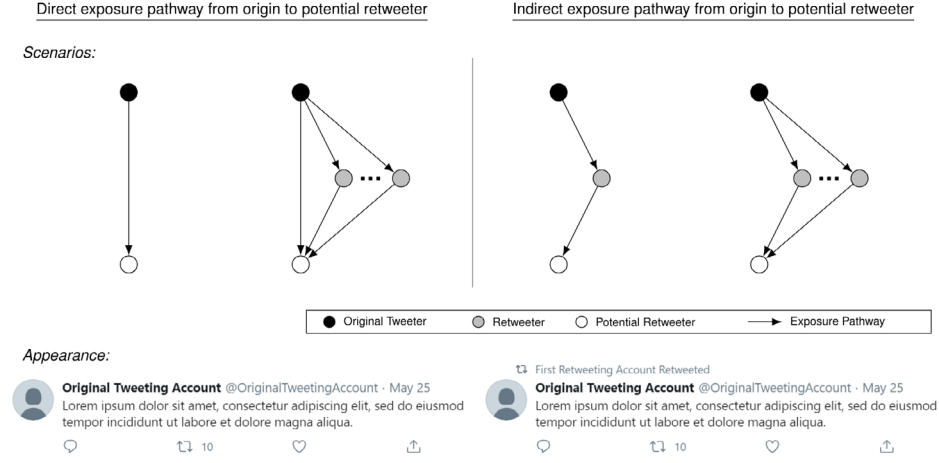
Comments: Different tweeting and retweeting scenarios and what appears on the follower's timeline. In both cases, additional retweeting beyond the source closest to the original tweet is not visible on the timeline.

A cascade – that is, the whole process of a single tweet spreading on the follower network – is modelled similar to the independent cascade model (Kempe et al., 2003) and its extension (Barbieri et al., 2013). In our model, a user retweets a tweet they have seen with a probability that depends on the virality score of the tweet and their own general activity level. More specifically, we first defined the activity (α_u) of a user (u) as the total number of times they have tweeted or retweeted during our data collection period. Then, a tweet (T) is retweeted with probability ($\alpha_u r_T$), a product of the activity of the exposed user (α_u) and the virality score of the tweet (r_T). The likelihood of observing a cascade in our data is computed as the joint probability of exposed users retweeting or failing to retweet independently of each other. For example, if in our data we observe tweet T to successfully activate users u and v, and fail to activate user w in the cascade, then the likelihood of the cascade is $L(r_T) = \alpha_u r_T \, \alpha_v r_T \, (1 - \alpha_w r_T)$.

Under this model, we can infer the virality of any tweet in our dataset by observing the number of successful and failed activations of the tweet and the activity levels of the corresponding users. Intuitively, the more successful activations over failed activations of a tweet we observe, and the less active users the tweet manages to activate, the higher virality the tweet should have. Technically, we achieved such inference of tweet virality using a maximum likelihood method (Saito et al., 2008). Specifically, for every tweet (T) in our dataset, with r_T as an unknown parameter, we first find the likelihood function [$L(r_T)$] of its cascade similar to the above example, and then find the value of r_T that maximises the value of this function. This value of r_T is the inferred virality score of the tweet.

The virality score of a tweet is likely not the same within and between different groups, because some content might resonate within a specific group but

not outside the group. Yet, since the vast majority of popular tweets do not travel between the activist and sceptic groups in our data, in this study we only focus on ingroup virality: that is, for each tweet, we only look at how virally it spreads in the group that it mainly spreads in. This is achieved by simply disregarding the follower network and retweets outside of the main-spreading group when inferring the virality of a tweet, in effect discarding both the successful and failed cascade events for users outside the group.

To conduct the virality inference on our data, we built the follower network of the classified subset of users. Using the Twitter API, we first collected the Twitter followees of each user, and then we constructed the follower network with each user as a single node and a link from each user to each of their followees. Following such a process, we obtained a user follower network of 20,628 nodes (i.e., users) and 2,398,028 links, with which we were then able to infer the virality of every tweet in our dataset.

Viral themes among activists and sceptics

The virality scores and the tweets we manually labelled can be used to investigate which characteristics of a tweet make it viral. We explored this question by finding out which characteristics of a tweet best predict its virality in the activist and sceptic groups, respectively. More precisely, we fit a group lasso model (Yuan & Lin, 2006) for each group, with the log-transformed virality score of a tweet as the response variable, and the tweet characteristics as explanatory variables. To better observe and control for author's effect on tweet virality, both in terms of the characteristics of the author and the structure of their follower network, we first selected a subset of the top-500 tweet set for this study, where each unique author has at least 3 tweets in the subset; the selected subset contains 261 tweets from the activist group and 128 tweets from the sceptic group. Then, we included among the explanatory variables a binary author indicator variable for each unique author in the selected subset. We then grouped all the explanatory variables so that all author indicator variables were in the same group, and each of the remaining variables was in a separate group.

A lasso model fits the data in a way similar to a linear regression model, yet additionally performs regularisation by removing redundant explanatory variables. This decreases the risk of overfitting an excessively complex model to noises in the sampled data, and consequently overestimating the significance of effects and reporting findings that will not replicate in the population (Babyak, 2004; Hawkins, 2004; McNeish, 2015). Consequently, this automatic variable selection process improves the parsimony and generalisability of the model and the validity of its interpretations (Fariss & Jones, 2018; McNeish, 2015). Further, a group lasso model performs variable selection in a grouped manner, so that each group of variables is included or excluded as a whole.

We selected with cross validation the most appropriate level of regularisation with which the model had the best predictive performance on the validation set that was unseen during model training. Estimated coefficients were then interpreted in the same way as those from linear regression models with log-transformed outcomes. More specifically, a coefficient value of x indicates that for every one-unit increase in the explanatory variable, the virality of the tweet is predicted to change by $(e^x - 1) \times 100\%$. Here, we present in Table 2 the transformed coefficients, which can be directly interpreted as the predicted percentage change in tweet virality for every unit increase in the explanatory variables.

Table 2   Effects of the explanatory variables on tweet virality, activist and sceptic groups

	Explanatory variable	Predicted change in virality (%)
	Mocking	0.0
	Call to Action	0.0
	Ingroup Praise	0.9
	Outgroup Criticism	1.5
Activist group (n = 261)	**Science**	-2.7
	Movement	12.3
	Solution	0.0
	Hashtags	0.5
	Mentions	1.7
	Authors	-32.3~46.4
	Mocking	0.0
	Incivility	16.2
	Call to Action	14.5
	Ingroup Praise	0.0
	Outgroup Criticism	-7.9
Sceptic group (n = 128)	Science	0.0
	Anti-international	0.0
	Hypocrisy	10.1
	Conspiracy	0.0
	Hashtags	1.6
	Mentions	1.1
	Authors	-27.3~26.5

Comments: The predicted percentage change in tweet virality for every unit increase in the explanatory variables. *n* = number of samples used to fit the model in each group. The variables highlighted in grey indicate those not excluded by the lasso model.

We first look at what characteristics of a tweet are related to its virality in the activist group. As shown in Table 2, the best model selected with the group lasso method excludes Mocking and Solution, indicating that whether a tweet in the activist group uses a mocking tone or includes a solution to climate change is a

poor predictor of its virality. Among the retained features, the virality of a tweet seems to be mostly predicted by who the author of the tweet is, seen in the large effect estimates of the author indicator variables. The number of hashtags and the number of mentions in a tweet positively contribute to its virality, which resonates with previous work (Suh et al., 2010).

Among the tweet characteristics we labelled, Movement seems to be the most important one in predicting the virality of a tweet in the activist group. The model indicates that if the tweet discusses the strength of a pro-climate movement, then its virality increases by 12.3 per cent in the activist group. Meanwhile, Outgroup Criticism and Ingroup Praise also positively predict the virality of a tweet, but with a much weaker relationship. Interestingly, the occurrence of Science seems to have a small negative effect on the virality of a tweet in the activist group. Controlling for other features, if the tweet invokes science as support, then its virality decreases by 2.7 per cent.

Table 3   Most- and least-viral tweet pairs in the activist group

User	Most-viral tweet	Least-viral tweet
A1	**Virality score: 0.155** **[Movement]** LOOK AT ALL THOSE PEOPLE marching in Alberta to demand #climateaction! Even here people want action on the #ClimateEmergency. Denying our situation doesn't help. #ActOnClimate #ClimateStrike #FridaysforFutures #cdn-poli #climate #energy #elxn43 @GretaThunberg. Via @vineshpratap [video]	**Virality score: 0.035** **[Science, Call to Action]** Air pollution is now more deadly than war, smoking and TB. It kills 7 million people every year. We have solutions to keep our communities safe and deal with the #climate crisis. Let's implement them. #GreenNewDeal #AirPollution #ClimateChange #energy #tech #PanelsNotPipelines [video]
A2[a]	**Virality score: 0.125** **[Ingroup Praise, Outgroup Criticism]** Attacks on Thunberg is motivated by one thing. She is intelligent, eloquent, compassionate, and young. She has scared some hateful and reactionary so-called 'grown ups', #ActOnClimate #ClimateCrisis [link]	**Virality score: 0.035** **[Science]** Climate change has rapidly and dramatically affected the arctic region. Even just ten years ago it was impossible for container ships to go through the Northern Sea Route. [link]

Comments: Pairs of the most-viral and least-viral tweets posted by the same user within the selected tweets in the activist group. The virality score and labels of each tweet are shown along with the text.

[a] Tweets from non-public figures, paraphrased while retaining the same features to protect users' privacy.

To more intuitively illustrate these effects, we present in Table 3 several pairs of the most-viral and least-viral tweets posted by the same user from our dataset.

As shown, the most-viral tweet by user A1 describes the strength of a pro-climate movement (thus labelled Movement) and has a virality score of 0.155. On the other hand, A1's least-viral tweet discusses the seriousness of the current air pollution situation with scientific evidence (mostly in the video attached) and calls for action to deal with it (thus labelled Science and Call to Action). It has a virality score of only 0.035. Meanwhile, the most-viral tweet by user A2, which praises Greta Thunberg and criticises her opponents (thus labelled Ingroup Praise and Outgroup Criticism), obtains a virality score of 0.125. A2's least-viral tweet, similar to A1's, also uses scientific evidence to show the seriousness of climate change (thus labelled Science), and only has a virality score of 0.035.

With respect to the sceptic group, the best model selected with the group lasso excludes Ingroup Praise, Science, Anti-international, and Conspiracy, suggesting that such themes are not useful in predicting the virality of a tweet in the sceptic group. Similarly – as in the activist group – the identity of the tweet author is the strongest predictor of the virality of a tweet, and the number of mentions and hashtags have positive effects.

Among the labelled characteristics, Incivility and Call to Action seem to be the most important positive predictors of tweet virality in the sceptic group, while Hypocrisy also has a positive effect. Specifically, if the tweet contains uncivil language, then its virality increases by 16.2 per cent in the sceptic group. If it contains calls to action, its virality increases by 14.5 per cent. If it contains hypocrisy claims, its virality increases by 10.1 per cent. Intriguingly, despite the overwhelming number of these tweets, Outgroup Criticism predicts decreased tweet virality in the sceptic group. If a tweet makes an outgroup criticism, then its predicted virality decreases by 7.9 per cent.

In Table 4, we show the most-viral and least-viral tweet pairs within the selected tweets in the sceptic group. Specifically, the most-viral tweets by user S1, S2, and S3 – all with virality scores over 0.11 – respectively involve uncivil language, calls to action, and hypocrisy claims, which are the three most significant characteristics we find to be positively predicting virality in the sceptic group. Meanwhile, the least-viral tweets from the three users involve characteristics shown to have negative or no effect on tweet virality, such as Mocking, Conspiracy, and Outgroup Criticism. We can also observe that Outgroup Criticism co-occurs with viral themes (e.g., Hypocrisy) in viral tweets, but occurs by itself in less-viral ones – which in part explains the predicted negative effect of Outgroup Criticism on virality.

Table 4   Most- and least-viral tweet pairs in the sceptic group

User	Most-viral tweet	Least-viral tweet
S1[a]	**Virality score: 0.190** **[Incivility]** I learned basic skills like sewing, cooking, woodwork, automobiles, and metalwork in high school home economics and shop classes :) Kids are now taught BS Socialism and Fake Climate Change... :/ KAG	**Virality score: 0.037** **[Mocking]** Thanks to California banning Plastic Straws... The Climate is becoming much better... It's only 45 today when it would have been 46... It's working... yay...
S2[a]	**Virality score: 0.118** **[Call to Action]** Just one year ago, no political party would say: - Climate crisis is a hoax - Immigration is too high - Corporate welfare and supply management must be eliminated - The budget can be balanced within two years PPC is changing the conversation and bringing change along with it. #VotePPC	**Virality score: 0.044** **[Conspiracy, Outgroup Criticism]** The network of globalist elites are all supporting each other. They promote fake climate change to scare people into submission. They are liars.
S3[a]	**Virality score: 0.113** **[Hypocrisy, Outgroup Criticism]** Greens Sarah Hanson-Young: The government needs to declare a climate emergency! Also Sarah Hanson-Young: I have taken 58 flights this year with taxpayer money. Do as I say, not do as I do. [image]	**Virality score: 0.072** **[Outgroup Criticism]** When will the Victoria police bill all these climate change protestors? [link]

Comments: Pairs of the most-viral and least-viral tweets posted by the same user within the selected tweets in the sceptic group. The virality score and labels of each tweet are shown along with the text.

[a] Tweets from non-public figures, paraphrased while retaining the same features to protect users' privacy.

Discussion

Our findings on virality-related characteristics first complement our empirical observations of common climate discussion themes in confirming a potential shifting trend of focus in the Twitter climate debate. Compared with Pearce and colleagues' (2014) analysis of Twitter discussions around the 2013 IPCC report, which found "science" to be one of the most prominent themes, our results show that the discussion of climate change science, although still an active topic, is not among the most-popular themes (see Figure 2), and is even among the least-viral themes (see Table 2) in both groups. The viral themes we found, including Move-

ment among activists, and Incivility and Hypocrisy among sceptics, further suggest that the core of Twitter climate discourse might have switched from the existence of climate change to climate activism – either emphasising the climate movements from the activist side, or attacking climate activism from the sceptic side.

On the other hand, our results also show the difference between popular themes and viral themes. Specifically, the use of uncivil language – although not a dominating theme in the sceptic group (see Figure 3) – is the most important characteristic that predicts tweet virality among the sceptics (see Table 2). In the meantime, outgroup criticism themes increase tweet virality in the activist group where they are less prevalent, and decrease tweet virality in the sceptic group where they are more prevalent. Apart from the theme co-occurrence issue that we discussed earlier, such discrepancy might be related to the different types of entities under criticism in the two groups. In the activist group, criticism is mostly directed at governments, which is democratically understood to be a socially acceptable target of dissatisfaction. In the sceptic group, however, criticism is usually directed at certain individuals and accompanied by mocking, incivility, or hypocrisy claims, thus potentially only attracting support from a limited subgroup of people. Such a phenomenon also aligns with the sensation-seeking literature (Bench & Lench, 2013; Zuckerman, 2010) in suggesting that people tend to lose interest in patterns under repeated exposure and are more easily elicited by novel stimuli.

Finally, we consider together the most important predictor of tweet virality in each group: Movement among activists and Incivility among sceptics. Clearly, they first resonate with our common theme analysis in showing the heterogeneity of discourse in the two groups and potentially hinting at different types of links that tie the community together. More specifically, Twitter climate activists nowadays might be most effectively united through the discussion of climate movements, while Twitter climate sceptics most likely bond over shared hostility toward climate activism.

Despite their evident difference, the Movement and Incivility themes both seem to serve the function of enhancing emotional connections within the group while rejecting potential involvement from outside the group. Specifically, promoting pro-climate movements likely increases enthusiasm from climate activists, while escalating the resistance of climate sceptics toward the issue. Meanwhile, incivility in comments against climate activism likely amplifies resonance among sceptics, yet elicits contempt and animosity from activists. Following Dahlgren's (2021) claim that people are constantly exposed to information from the outgroup, which originally served as evidence against polarisation, our results instead show that this exposure probably exacerbates polarisation, echoing Bail and colleagues' (2018) similar finding from a field experiment. In this sense, our findings reveal – from an information-spreading perspective – how the echo chambers of climate discussions are potentially further consolidated and separated.

Conclusion

In this article, we set out to study climate change discussions on Twitter through the lens of information spreading. Our work first presents an up-to-date picture of popular climate communication themes on Twitter both within and across the activist and sceptic groups. We show that climate activists and climate sceptics generally communicate within their own groups in disparate styles, and we additionally find a virtual absence of information sharing across the groups. These results corroborate prior findings that show evidence of echo chambers in climate communication on Twitter (Williams et al., 2015).

More importantly, we make a distinct contribution by examining the tweet characteristics that predict viral spreading within the two groups. First, we find that the virality-predicting themes showcase interesting matches and mismatches with the popular themes. Further interpreting the strongest predictors of viral spreading – Movement among activists and Incivility among sceptics – we argue that while these themes reflect different types of bonds that tie the community together, they both tend to enhance ingroup connections while repulsing outgroup engagement. This finding has implications in the broader context of climate change politics and communication, in that it reveals the potential for viral spreading to exacerbate polarisation in the climate debate on Twitter.

Funding and acknowledgements

References

Adam, K. (2019, October 22). Why didn't Greta Thunberg win the Nobel Peace Prize? *The Washington Post*. https://www.washingtonpost.com/world/europe/why-didnt-greta-thunberg-win-the-nobel-peace-prize/2019/10/11/e84e6efc-eba4-11e9-9306-47cb0324fd44_story.html.

Asikainen, A., Iñiguez, G., Ureña-Carrión, J., Kaski, K., & Kivelä, M. (2020). Cumulative effects of triadic closure and homophily in social networks. *Science Advances*, 6(19). https://doi.org/10.1126/sciadv.aax7310

Babyak, M. A. (2004). What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic medicine*, 66(3), 411–421. https://doi.org/10.1097/01.psy.0000127692.23278.a9

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. 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

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26(10), 1531–1542. https://doi.org/10.1177/0956797615594620

Barbieri, N., Bonchi, F., & Manco, G. (2013). Topic-aware social influence propagation models. *Knowledge and Information Systems*, 37(3), 555–584. https://doi.org/10.1007/s10115-013-0646-6

Barrat, A., Barthelemy, M., & Vespignani, A. (2008). *Dynamical processes on complex networks*. Cambridge University Press.

Baumann, F., Lorenz-Spreen, P., Sokolov, I. M., & Starnini, M. (2020). Modeling echo chambers and polarization dynamics in social networks. *Physical Review Letters*, *124*(4), 048301. https://doi.org/10.1103/PhysRevLett.124.048301

Bench, S. W., & Lench, H. C. (2013). On the function of boredom. *Behavioral Sciences*, *3*(3), 459-472. https://doi.org/10.3390/bs3030459

Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research*, *49*(2), 192–205. https://doi.org/10.1509/jmr.10.0353

Birch, S. (2020). Political polarization and environmental attitudes: A cross-national analysis. *Environmental Politics*, *29*(4), 697–718. https://doi.org/10.1080/09644016.2019.1673997

Carrington, D. (2019, March 14). Greta Thunberg nominated for Nobel Peace Prize. *The Guardian.* https://www.theguardian.com/world/2019/mar/14/greta-thunberg-nominated-nobel-peace-prize

Chen, T. H. Y., Salloum, A., Gronow, A., Ylä-Anttila, T., & Kivelä, M. (2020, July 6). Polarization of climate politics results from partisan sorting: Evidence from Finnish Twittersphere. *arXiv preprint arXiv:2007.02706*. https://doi.org/10.31235/osf.io/49qgu

Cody, E. M., Reagan, A. J., Mitchell, L., Dodds, P. S., & Danforth, C. M. (2015). Climate change sentiment on twitter: An unsolicited public opinion poll. *PloS one*, *10*(8), 1–18. https://doi.org/10.1371/journal.pone.0136092

Dahal, B., Kumar, S. A., & Li, Z. (2019). Topic modeling and sentiment analysis of global climate change tweets. *Social Network Analysis and Mining*, *9*(24). https://doi.org/10.1007/s13278-019-0568-8

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

Fariss, C. J., & Jones, Z. M. (2018). Enhancing validity in observational settings when replication is not possible. *Political Science Research and Methods*, *6*(2), 365–380. https://doi.org/10.1017/psrm.2017.5

Garimella, K., Morales, G. D. F., Gionis, A., & Mathioudakis, M. (2018). Quantifying controversy on social media. *ACM Transactions on Social Computing*, *1*(1), 1–27. https://doi.org/10.1145/3140565

Guerini, M., Strapparava, C., & Ozbal, G. (2011). Exploring text virality in social networks. *Proceedings of the International AAAI Conference on Web and Social Media*, *5*(1).

Hansen, L. K., Arvidsson, A., Nielsen, F. Å., Colleoni, E., & Etter, M. (2011, June 28–30). Good friends, bad news-affect and virality in twitter. In J. J. J. H. Park, L. T. Yang, & C. Lee (Eds.), *Future information technology: 6th International Conference on Future Information Technology* (pp. 34–43). Springer. https://doi.org/10.1007/978-3-642-22309-9

Hanteer, O., & Rossi, L. (2019). An innovative way to model twitter topic-driven interactions using multiplex networks. *Frontiers in Big Data*, *2*(9). https://doi.org/10.3389/fdata.2019.00009

Hawkins, D. M. (2004). The problem of overfitting. *Journal of Chemical Information and Computer Sciences*, *44*(1), 1–12. https://doi.org/10.1021/ci0342472

Jang, S. M., & Hart, P. S. (2015). Polarized frames on "climate change" and "global warming" across countries and states: Evidence from twitter Big Data. *Global Environmental Change*, *32*, 11–17. https://doi.org/10.1016/j.gloenvcha.2015.02.010

Jenders, M., Kasneci, G., & Naumann, F. (2013). Analyzing and predicting viral tweets. *Proceedings of the 22nd international conference on world wide web* (pp. 657–664). https://doi.org/10.1145/2487788.2488017

Karypis, G. & Kumar, V. (1998). A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on scientific Computing*, *20*(1), 359–392. https://dl.acm.org/doi/10.5555/305219.305248

Kempe, D., Kleinberg, J., & Tardos, É. (2003, August). Maximizing the spread of influence through a social network. *Proceedings of the ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 137–146).

Kirilenko, A. P. & Stepchenkova, S. O. (2014). Public microblogging on climate change: One year of twitter worldwide. *Global Environmental Change*, *26*, 171–182. https://doi.org/10.1016/j.gloenvcha.2014.02.008

McCright, A. M., & Dunlap, R. E. (2011). The politicization of climate change and polarization in the American public's views of global warming, 2001–2010. *The Sociological Quarterly*, *52*(2), 155–194. https://doi.org/10.1111/j.1533-8525.2011.01198.x

McNeish, D. M. (2015). Using lasso for predictor selection and to assuage overfitting: A method long overlooked in behavioral sciences. *Multivariate Behavioral Research*, *50*(5), 471–484. https://doi.org/10.1080/00273171.2015.1036965

Muddiman, A., McGregor, S. C., & Stroud, N. J. (2019). (Re)claiming our expertise: Parsing large text corpora with manually validated and organic dictionaries. *Political Communication*, *36*(2), 214–226. https://doi.org/10.1080/10584609.2018.1517843

Munger, K. (2019). The limited value of non-replicable field experiments in contexts with low temporal validity. *Social Media + Society*, *5*(3), 1–4. https://doi.org/10.1177/2056305119859294

Musco, C., Musco, C., & Tsourakakis, C. E. (2018). Minimizing polarization and disagreement in social networks. *Proceedings of the 2018 World Wide Web Conference* (pp. 369–378).

Nagarajan, M., Purohit, H., & Sheth, A. (2010). A qualitative examination of topical tweet and retweet practices. *Fourth International AAAI Conference on Weblogs and Social Media*. Citeseer.

Nelimarkka, M., Laaksonen, S.-M., & Semaan, B. (2018). Social media is polarized, social media is polarized: Towards a new design agenda for mitigating polarization. *Proceedings of the 2018 Designing Interactive Systems Conference* (pp. 957–970).

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin UK.

Pearce, W., Holmberg, K., Hellsten, I., & Nerlich, B. (2014). Climate change on Twitter: Topics, communities and conversations about the 2013 IPCC working group 1 report. *PloS one*, *9*(4), e94785. https://doi.org/10.1371/journal.pone.0094785

Saito, K., Nakano, R., & Kimura, M. (2008). Prediction of information diffusion probabilities for independent cascade model. *International conference on knowledge-based and intelligent information and engineering systems* (pp. 67–75). Springer.

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, *29*(4), 217–248. https://doi.org/10.2753/MIS0742-1222290408

Suh, B., Hong, L., Pirolli, P., & Chi, E. H. (2010). Want to be retweeted? Large scale analytics on factors impacting retweet in twitter network. *2010 IEEE Second International Conference on Social Computing* (pp. 177–184). IEEE.

Tyagi, A., Uyheng, J., & Carley, K. M. (2020). Affective polarization in online climate change discourse on twitter. *arXiv preprint arXiv:2008.13051*.

Weng, L., Menczer, F., & Ahn, Y.-Y. (2013). Virality prediction and community structure in social networks. *Scientific reports*, *3*, 2522. https://doi.org/10.1038/srep02522

Williams, H. T., McMurray, J. R., Kurz, T., & Lambert, F. H. (2015). Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global environmental change*, *32*, 126–138. https://doi.org/10.1016/j.gloenvcha.2015.03.006

Yuan, M., & Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *68*(1), 49–67. https://doi.org/10.1111/j.1467-9868.2005.00532.x

Zuckerman, M. (2010). Sensation seeking. *The Corsini encyclopedia of psychology*, 1-4.