



UvA-DARE (Digital Academic Repository)

Inside a White Power echo chamber

Why fringe digital spaces are polarizing politics

Törnberg, P.; Törnberg, A.

DOI

[10.1177/14614448221122915](https://doi.org/10.1177/14614448221122915)

Publication date

2024

Document Version

Final published version

Published in

New Media & Society

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Törnberg, P., & Törnberg, A. (2024). Inside a White Power echo chamber: Why fringe digital spaces are polarizing politics. *New Media & Society*, 26(8), 4511-4533. <https://doi.org/10.1177/14614448221122915>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



Article

Inside a White Power echo chamber: Why fringe digital spaces are polarizing politics

new media & society
2024, Vol. 26(8) 4511–4533
© The Author(s) 2022



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/14614448221122915
journals.sagepub.com/home/nms



Petter Törnberg 
University of Amsterdam, The Netherlands

Anton Törnberg 
University of Gothenburg, Sweden

Abstract

Recent decades have seen a blurring of the line between extremist movements and mainstream politics, driven by rising sectarian polarization. This development has been linked to digital media, with suggestions that so-called *echo chambers* may drive political radicalization. To understand the social processes taking place inside such digital spaces, this article draws on Randall Collins and the Durkheimian tradition to develop a theory of discursive community formation. Empirically, we analyze 20 years of discussion on the White Power forum Stormfront, employing natural language processing to study discursive evolution as members become socialized into the community. Our findings suggest that digital media provide space for conversational rituals that instill in people a sense of social membership and intersubjectivity, contained in the elaboration of a shared discourse, within which certain beliefs become sacred and unquestionable. This provides a potential social mechanism linking echo chambers to the rise of sectarian polarization.

Keywords

Collective identity, natural language processing, polarization, radicalization, White Power

Corresponding author:

Petter Törnberg, Amsterdam Institute for Social Science Research (AISSR), University of Amsterdam, Postbus 15629, 1001 NC Amsterdam, The Netherlands.
Email: p.tornberg@uva.nl

Introduction

On 6 January 2021, a violent far-right mob stormed the US Capitol building. Armed and prepared to take hostages, they came within meters of their intended victims: US Senators and the Vice President. The mob had mobilized on social media, fueled by conspiracy theories and misinformation which brought them together under a collective sense of shared purpose. While in many ways shocking and unprecedented, this event was in other ways typical and predictable: emblematic of a growing pattern of political movements spawned on and through digital spaces, guided by misinformation and emboldened to act. This appears as merely one in a growing list of expressions of the rise of the far right, and the growing legitimacy of violence within an increasingly polarized political mainstream—blurring the boundary between extremist movements and mainstream politics (Mason and Kalmoe, 2022).

For scholars and the public alike, the pattern of digitally mobilized radical politics is raising important questions about the impact of digital media on political life. Within this research, protected digital spaces where like-minded people can meet—often called “echo chambers”—has been at the center of scholarly attention, highlighted as potential drivers of radicalization. Why, precisely, such spaces would be affecting political life has however remained subject to significant debate. The most influential version of the echo chamber hypothesis suggests that isolation from opposing perspectives leads individuals to “self-radicalize” through repeated exposure to one-sided content (Sunstein, 2009). This hypothesis has however been challenged by recent empirical evidence (Guess et al., 2018; Keuchenius et al., 2021; Zuiderveen Borgesius et al., 2016), leading to calls for understanding what is in fact taking place within these digital spaces. What happens to individuals as they interact inside an echo chamber? How and why do echo chambers produce radicalized groups with worldviews that are severed from both reality and society at large? How are members of such groups emotionally charged and primed to commit even violent acts?

We focus on two recent strands of research that point to the important political consequences of such fringe spaces. The first strand focuses on their impact on identity, arguing that fringe spaces contribute to polarization by strengthening political identities, thereby driving a form of political polarization characterized by out-group aversion (Finkel et al., 2020; Klein, 2020; Törnberg et al., 2021). The second strand focuses on the spaces’ impact on discourse, arguing that part of the radicalizing political capacity of fringe digital spaces lies in their remarkable capacity for discursive innovation—creating memes, slang, and conspiracy theories through constant playful experimentation (Nagle, 2017; Phillips, 2015).

In this article, we suggest that these two echo chamber phenomena—strengthening of political identity and intensive discursive innovation—describe two dimensions of one and the same Durkheimian social process of discursive community formation, that is, online communities emerge in and through the construction of a shared culture. We draw on Randall Collins’ work on community formation to develop and apply a theory of the social processes that occur in echo chambers. This theory suggests that as individuals come together under the banner of a shared interest, this interest tends to be transformed into a collective identity, simultaneously articulated into a discourse which serves to

demarcate insiders from outsiders (Collins, 2004; Durkheim, 1912). The community's collectivity and subjectivity are *contained* within these discourses, and some ideas come to function as sacred objects for the community, thereby becoming unquestionable and taken for granted (Benwell and Stokoe, 2006; De Cillia et al., 1999).

To study this social process empirically, we go inside a White Power echo chamber to examine the digital social lives of far-right extremists. We trace the entangled language and identity evolution through which these communities are formed and maintained. Our case is one of the most significant extremist communities on the Internet: the complete corpus of the far-right forum Stormfront.org, comprising 10,172,069 posts and 354,574 users, spanning over 20 years of discussion. Stormfront provides a meeting-place for the White Power movement, and has been linked to many of the far-right hate crimes, terrorist attacks, and mass murders of the last two decades (Kleinberg et al., 2020). It has been referred to as “the murder capital of the internet” for its connection to many recent far-right terrorist attacks (Beirich, 2014). Stormfront constitutes a unique case in having remained a significant far-right forum since its founding, during the early days of the Internet (Scrivens et al., 2020). The Stormfront dataset provides a powerful view into how echo chambers shape the political life of individuals—as they come to see themselves as part of the community and begin to construct a shared worldview, with common enemies and grievances.

Methodologically, the discursive perspective suggests using language as a window into the process of community formation, revealing the formation and maintenance of social identity. We use natural language processing (NLP) to study language evolution in large text material, enabling the examination of identity as a form of discourse (Danescu-Niculescu-Mizil et al., 2013; Kleinberg et al., 2020). Focusing on individual radicalization and member recruitment, we trace the evolution of the language of new members as they become part of the community, and we use the contrast between new and old users to draw out the linguistic expressions of community belonging and *intersubjectivity*, that is, of members' sense of “we-ness.”

Echo chambers and political radicalization

Digital media has been implicated in political radicalization, expressed in both the strengthening of extremist movements and the polarization of mainstream politics (Belew, 2018; Mason and Kalmoe, 2022). Yet, the exact nature of the link between digital media and political radicalization has been subject to significant debate. Many scholars point to the central role of fringe digital spaces—often referred to as “echo chambers” or “filter bubbles” (Bruns, 2019; Pariser, 2011; Sunstein, 2009)—in driving polarization and political extremism. In its most influential formulation, the echo chamber hypothesis suggests that these spaces drive divergence in issue position by isolating individuals from opposing perspectives (Sunstein, 2009). Individuals in echo chambers become “self-radicalized” through repeated exposure to one-sided extremist content, resulting in political views that potentially prime them to engage in isolated, “lone wolf,” violent actions. However, this hypothesis for the radicalizing mechanism of echo chambers has been subjected to significant empirical criticism, as most digital media platforms in fact prove relatively interconnected across the ideological divide—thus challenging the

notion that users will become fully isolated from opposing views (e.g. Guess et al., 2018; Keuchenius et al., 2021; Morales et al., 2021; Zuiderveen Borgesius et al., 2016). This has prompted many in the field to question the conventional echo chamber hypothesis and to call for alternative understanding of the link between fringe digital spaces and political radicalization.

For our purposes, we can identify two recent strands that provide the seeds for such an alternative understanding of how echo chambers are driving political radicalization.

The first strand focuses on impact of echo chambers on identity. Recent research has suggested that such spaces act on politics by *strengthening political identities* (Klein, 2020; Törnberg et al., 2021). This constitutes an important reformulation of the traditional echo chamber hypothesis. First, the new formulation does not make the contentious claim that individuals are fully sheltered from opposing views (Guess et al., 2018; Keuchenius et al., 2021); instead, it merely makes the near-indisputable observation that there exist digital spaces where like-minded people meet. Second, the suggestion is not that these spaces make issue positions more radical, but rather that they strengthen shared identities. This speaks to a view of contemporary politics in which out-group aversion, rather than political ideals and beliefs, is at the core of political conflict.

Such an identity-oriented perspective on politics has become influential within the literature on both mainstream political polarization and extremist movements, characterizing political polarization and radicalization not as divergence of opinion, but as the strengthening of political identities (Finkel et al., 2020). This literature emerged from the observation of a substantial rise in partisan aversion in the US in recent decades, while issue-position polarization has remained fixed and relatively low (Fiorina, 2017), thus suggesting that a novel form of polarization is at play. This contemporary polarization has been referred to as *affective* (Iyengar et al., 2012), *sectarian* (Finkel et al., 2020), or *social* polarization (Mason, 2018), meaning that it is driven by deeply rooted mechanisms of group affiliation in human psychology (Iyengar et al., 2019). By shaping a politics founded on out-group aversion rather than political ideals, such identity-oriented polarization transforms elections from contests over policy disagreements into struggles between warring tribes, separated by a fundamental sense of difference (Sides et al., 2018). Sectarian polarization transforms issue positions into symbols of group belonging, implying that while shifts in issue position may occur as a result of radicalization, they are secondary to a process of socialization (Törnberg et al., 2021). While social identity has always played a role in politics, this literature suggests that we have entered a situation in which partisan identity is coming to dominate or even engulf other identities (Iyengar et al., 2012, 2019; Klein, 2020).

The literature on extremist radicalization has a similar emphasis on identity, seeing radicalization as a form of socialization in which individuals come to view themselves as part of a larger collective, often defined through its opposition to an external group (Borum, 2011; Della Porta, 2013; Vasquez et al., 2015). Identity is seen as a key component in the “increased preparation for and commitment to intergroup conflict” (McCauley and Moskaleiko, 2008: 416). Through this lens, to be radicalized means, primarily, that one’s political identity grows dominant in one’s self-understanding.

The emphasis on political identities in the polarization and radicalization literature speaks to the notion of “collective identity” (Diani, 1992; Snow, 2001) as developed in

social movement theory. This notion suggests that collective identities are constructed as activists interact and share ideas with other members of their in-group (Futrell and Simi, 2004; Polletta and Jasper, 2001). Recent studies have argued that social movements can also make use of digital media for such identity work, using digital messages to develop a common sense of “we” (Gaudette et al., 2020). The focus on social identity resonates with research in digital media studies, which has gathered significant evidence to suggest that social media have become a site for the formation of identities (Clark, 2014; van Haperen et al., 2020) and are capable of fostering communities with a strong sense of group solidarity (Beyer, 2014; Crossley, 2015; Papacharissi, 2011; Turkle, 2011).

The second strand focuses on echo chambers as drivers of discursive and cultural processes, examining how web culture has grown to become a radicalizing political force (Belew, 2018; Nagle, 2017; Reagle, 2015). This literature has found that digital spaces appear to have an innate tendency to produce rich internal subcultures—involving particular vernaculars, slang, memes, and stories (Zannettou et al., 2018). In the early period of the Internet, these mostly consisted of harmless jokes and cultural expressions, such as taking cute pictures of breaded cats or unexpected appearances of the music video for the 1987 hit song “Never Gonna Give You Up.” In more recent years, however, the cultural expressions emerging from these fringe spaces have taken on a political and distinctly reactionary hue (Nagle, 2017; Reagle, 2015; Törnberg and Törnberg, 2016). This period has seen a constant flow of extremist discourses, ideas, and memes from fringe online spaces into the political mainstream, which has contributed to the mainstreaming of White supremacist ideology—ranging from novel hate symbols, such as Pepe the Frog or the OK hand signal, to far-right conspiracies such as the Great Replacement or QAnon (Bail, 2012; Belew, 2018; Reagle, 2015). As this culture has become political, these spaces’ remarkable capacity for discursive innovation, memes, slang, and stories through constant playful experimentation has made them a real political force. As scholars have argued, these memes and stories encode certain political subjectivities constructed around a vague notion of a shared “other,” which can function to drive political conflict (DeCook, 2018; Tuters and Hagen, 2020).

In this article, we suggest that these two observed phenomena of fringe digital spaces—the strengthening of political identity and subcultural discursive innovation—are inextricably interlinked in a Durkheimian process of community formation. We draw on the work of Randall Collins to argue that what takes place in echo chambers can be best understood as the development of collective identities through the elaboration of a cultural and discursive system. As individuals come together in a digital space under the banner of a shared interest, this interest tends to be transformed to a collective identity, simultaneously articulated into a discourse which serves to separate insiders from outsiders and function as linguistic capital within the community (Bourdieu, 1991). Following the literature on discursive identity, their collectivity and subjectivity is *contained* within these discourses (Benwell and Stokoe, 2006; De Cillia et al, 1999).

To cast theoretical light on this social process, we now turn to the work of Randall Collins and the Durkheimian tradition of community formation. We will adapt Collins’ work to elaborate a theoretical framework for understanding the interlinked role of language and community in online media.

Durkheim and Collins on community and rituals

Through his anthropological work on aboriginal communities, Durkheim (1912) sought to identify the core social mechanisms that hold societies together. He found that the central social practice that formed the bonds which held together these communities were the religious rituals in which synchronized movements and chanting brought the tribe to a trance-like state—“collective effervescence”—filling participants with emotional energy and a sense of community. The objects that were the center of attention during these rituals became charged with emotional energy and a sense of intersubjectivity—they became sacred for the community. The totem, Durkheim suggested, was the group’s experience turned into a physical object, an emblem that could bring and sustain the collective energy of the rituals into everyday community life.

Scholars have since broadened and translated Durkheim’s (1912) findings to throw light on the modern secular societies of today—elaborating it into a process of community formation that links social membership, moral beliefs, and cultural production through a single mechanism: ideas are symbols of group membership, and culture is thus generated by the emotional patterns of social interaction within religious rituals. Randall Collins (2004), in particular, has drawn on Erving Goffman to reinterpret Durkheim’s work in the form of a microsociological theory on how groups produce social membership and intersubjectivity—that is, their sense of constituting a *we*. Collins again places rituals at the center—moments of shared attention and emotion—viewing these moments as capable of transforming objects of shared attention into symbols charged with group belonging. These symbols then become used as part of further rituals, creating a chain of interaction that constitutes the foundation of the shared sense of community. Collins uses rituals in a broad sense: dancing together at a club, shouting in unison at a Trump rally—or even something as simple as sharing a cigarette.

While Durkheim and Collins viewed physical co-presence as necessary for interaction rituals, recent work has suggested that these rituals can also take place in mediated environments (DiMaggio et al., 2018; Maloney, 2013; van Haperen et al., 2020). While such online interaction rituals may be lower in intensity, this is compensated by their often sustained and long-term nature. As non-physical meetings limit the possibilities for physical artifacts, barriers for entry, or markers by which one may build confidence that the person with whom one is conversing is indeed an insider (Maloney, 2013), the interaction rituals and their effects instead take place in the discursive realm (Benwell and Stokoe, 2006). This requires adapting Collins’ framework to the digital realm, to describe the process that we refer to as *discursive community formation* (cf. Colombo and Senatore, 2005).

Discursive community formation

Building on Durkheim, Collins (2004) proposed that interaction rituals have four ingredients: *group assembly*, *barriers to outsiders*, *mutual focus of attention*, and a *shared mood*. These ingredients partially overlap and feed off each other. As these ingredients come together, they in turn have four outcomes: *group solidarity*, *common standards of morality*, *sacred objects*, and *emotional energy in individuals*. The ingredients and

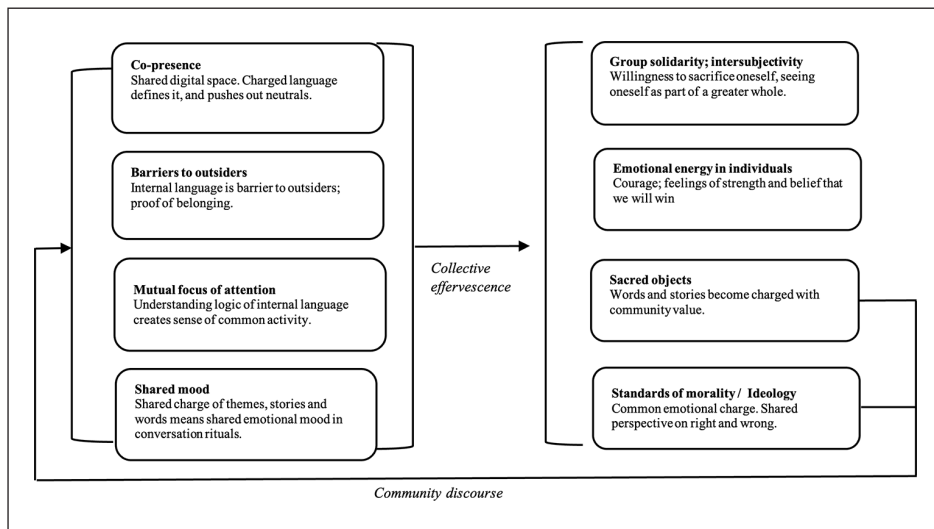


Figure 1. Summarizing figure, adapting Collins (2004: 47) to describe online communities, showing how discourse and language symbols fulfill the symbolic functions that Collins outlines.

outcomes are thus connected, forming a feedback loop: a chain of interaction that is the foundation of discursive community formation (see Figure 1).

On digital media, the interaction ritual takes the form of the exchange of messages on a shared topic. In other words, the rituals are conversations. From the point of view of social membership, conversations are significant not so much for their content, but rather as a moment of shared focus on a common activity: like any ritual, they can be constituted by a shared focus on a shared set of symbols combined with a shared emotion. The difference, of course, is that the four ingredients and the four outcomes are fulfilled by the realm of ideas, discourse, and language. The objects of the shared attention are words, stories, and images, rather than a physical object, and it is these words, stories, and images that convey the shared experience.

The sense of *group assembly* is provided by the common banner under which individuals have gathered. While digital spaces cannot provide a sense of shared physical space, they do feature designs and descriptions that demark the purpose and shared focus of the community: the logo, name, description, and graphic designs provide the foundation on which a common cognitive reality is gradually constructed. Digital spaces are so designed as to raise a banner declaring the shared attributes around which the community gathers.

As part of this, the community over time develops certain *barriers to outsiders*, taking the form of an internal culture and language. These are ways of determining who is part of a community, and who is not. Some such markers can be technically encoded in the digital space, for instance, through information such as number of user posts or community status, or the possibility to choose a recognizable username or picture. The most important means of determining insiders from outsiders, however, lies in the discursive

realm; certain words, themes, stories, ideas, or images come to serve as emblems and evidence of group membership. (This will be discussed in more detail below.)

As communities meet around shared interests, these give rise to certain shared topics and themes which comprise the *mutual focus of attention*. Certain topics become typical for the community, and the conversations will tend to center around these topics. Online meeting places are furthermore technologically structured so as to allow the conversations to share topics—forums, for instance, have subforums and discussion threads, which organize the conversations to make sure participants have a mutual focus of attention.

The stories, languages, and local knowledge that become characteristic of a community not only function as membership emblems and cultural capital for the community, but also carry a certain emotional charge, creating the experience or feeling of a *shared mood*. As members learn the discourse of the community, they also learn what to feel about different topics and stories. The specialized language of the community has a symbolic value and is charged with a special excitement, tension, or enthusiasm through conversation rituals. This is part of what makes them powerful tools for invoking a common cognitive reality in conversation rituals, functioning as conversational or cultural resources that invoke “a shared reality” (Collins, 1981: 1001).

Effects of successful digital rituals. The effects of rituals can be categorized into two classes: first, intersubjectivity, collective identity, and emotional energy; second, a community discourse that contains within it the community and its ideology. These link to the two strands of research on the effects of digital spaces on political polarization.

First, a central effect of successful rituals is that they create a sense of *group solidarity*, strengthening their collective identity. As individual participants develop a stronger sense of solidarity and intersubjectivity, they come to also assume the thoughts, morals, and behaviors internal to their group, viewing themselves less as individuals, and more as part of the community. The ritual, in short, transforms a group of individuals into a community; a shared sense of “we.”

Participants experience rituals as a pleasant experience, filling them with what Collins calls *emotional energy*: a positive feeling that makes participants want to stay in the community, often manifested as confidence, warmth, and enthusiasm. This is, in other words, the dopamine boost that many media scholars have described as being a central driver of social media use, and which can even lead to addiction. This emotional energy is what drives members to act on behalf of the community in other settings, such as participating in a demonstration—or perhaps storming the US Capitol.

Second, the rituals are based on the elaboration of a discursive system: the objects of shared attention are the topics, concepts, beliefs, and interests around which the community is gathered. Some of these can come to function as symbols of the community. Just like the *sacred objects* in the physical rituals described by Collins and Durkheim, these symbols are used as part of further rituals, becoming the cultural artifacts that create a chain of interaction rituals that constitute the foundation of the shared sense of community. These cultural items are charged with membership significance through repeated ritualistic interactions, making these symbols not just indicative of the group, but the very stuff through which intersubjectivity is constructed and

maintained. Words and ideas function as symbols for the community; they become a discursive form of totems.

These discursive symbols come to form the linguistic capital through which rituals are enacted, both as emblems of group membership and as indicative of a shared moral foundation (Collins, 2004). The internal language provides barriers to outsiders ensuring that only those who are “in the know” can participate. This function can be seen in how meme culture tends to exhibit complex layers of intertextual references, abstract and ironic styles, constantly in flux and innovation, requiring both literacy and dedication to decode and stay up-to-date with the latest trends (Knobel and Lankshear, 2007; Shifman, 2013). This challenge is precisely the point (Phillips and Milner, 2017), as this language functions to create a subcultural definition of cultural capital, in opposition to the mainstream culture. These subcultures thus define forms of distinction through a linguistic market, conferring cultural capital and authority on those who master the language (Bourdieu, 1991). This separates outsiders from insiders, through the demarcation of those who are unaware of the subcultural logic and values of the community.

However, the community discourse is not merely an arbitrary collection of language games. The community’s discourse expresses a political subjectivity, and embodies the community ideology: what is seen as good and what is evil (Durkheim, 1912). The formation of a collective identity is necessarily also the creation of difference; as Benhabib (1996: 33) puts it, “every search for identity includes differentiating oneself from what one is not”—identifying the in-group with good, and casting as evil what lies outside the group’s boundaries. Since the identification of shared similarities necessarily implies the creation of a difference—a sense of what the community values and what it opposes—this discursive system comes to define *common standards of morality*. For Durkheim, identity was formed through opposition to the devil and a striving for similarity with a God, ascribed through the rituals of regular participation in religious services. In applying this to the broader context of group formation, the bond of similarity can be made in opposition to an outside force—which, in turn, becomes the personification of evil.

Methodology and data

The Stormfront.org data were collected by custom-written web scrapers, parallelized, and running through proxies to avoid tracing. The scraper was developed in Python by the first author and was executed from a dedicated server. The dataset consists of 10,172,069 posts, and 354,574 users, out of which 99,988 had written one or more posts. The forum data were collected in September 2020. The anonymized data have been made publicly available (Törnberg, 2022). The data were analyzed using Python and standard data analysis packages such as Pandas and nltk.

We follow the ethical guidelines for Internet research provided by the Association of Internet Researchers (Franzke et al., 2020) and by the British Sociological Association (BSA, 2017). To ensure anonymity for the members of the community, usernames were deleted when collecting the corpus. Since the analytical focus is on broader discursive patterns, it is not possible to identify individual users from the results here presented. The project has been assessed and approved by the Regional Ethical Review Board.

While operationalizing Collins' theoretical framework is a challenging task, NLP techniques can provide a first step toward examining community formation through a discursive lens. For pre-processing, we identified language of the posts using the Python package `langdetect` and focused on English-language posts ($N=8,806,105$) that are over 120 characters long ($N=6,158,005$, from 81,039 users). We truncated messages at the last word before the 5000th character, to prevent a small number of extremely long messages from distorting the results (these messages are often copies of reports or lists of data, such as a 314,743-character post listing the purported prices of body parts on the Egyptian organ market).

We studied shifts in user language as expressive of shifting identity, focusing on members' linguistic evolution as they over time become socialized into the community. To do so, we numbered each user post by its order in its author's post history. This allowed us to compare users' early posts with their later ones, in order to examine how the language of posts evolves as users engage with the community. We primarily focus on users who remain on the community long-term, which we here define as users who sent at least 50 posts¹ (however, to verify that this parameter did not affect the findings, the analysis was also carried out for the thresholds of 20 and 100, which did not impact the findings). On average, it takes users 289 days to send their 50 first posts. We looked at two parts, matching the two parts of the result of discursive rituals: the convergence of user language with the community, and the content of the shifting language.

Language convergence

We began with the question of the discursive similarity between newly joined members and the forum's overall language, to examine how user language shifts over time as members engage with the community. To measure this, we constructed a language model that represents the forum's overall language by selecting a random sample consisting of 20,000 messages. To measure the distance between two corpora, we used bag-of-words representations; that is, we disregarded grammar and word order and represented texts as sparse vectors of occurrence counts of words. The bag-of-words were represented using Vector Space Model (VSM) with term frequencies as components. To measure the distance between two such bags-of-words, we used the cosine similarity, defined as the normalized dot-product of the two vectors. More formally, if $v = (v_1, v_2 \dots v_n)$ and $u = (u_1, u_2 \dots u_n)$ are two term frequency vectors, their cosine distance is:

$$D(u, v) = 1 - \cos(u, v) = 1 - \frac{\sum_{i=1}^n u_i v_i}{|u_i| |v_i|}$$

Content of lexical change

To examine which words make up the difference between early and late posts, we used Log-Likelihood to identify the most overrepresented words in the early posts compared to the later posts. This allowed us to find words and bigrams that are statistically

over- and underrepresented in the community language: words that characterize the particular discourse of the community.

We created two subcorpora based on posts from members who had written at least 50 posts, focusing on the first and last 20 of these posts (i.e. 1–20 vs 30–50). This selection allowed us to see the language before and after the members have converged on the forum language, and was selected such that the corpora have a similar number of words. This enabled us to inductively identify words and bigrams that are statistically over- and underrepresented in the community language, using Log-Likelihood comparison between the two VSMs (Boukes and Trilling, 2017; van Vliet et al., 2020).

One limitation of this approach is that it does not distinguish between jargon changes and thematic changes. That is, some of the shifts will reflect changing ways of referring to the same issue, while others will reflect a change in the issues discussed. We therefore developed a custom-made technique using word embeddings to identify forum vernacular replacing mainstream terms. We created a word embedding model (word2vec) of the combined materials of the early and late posts of the members (Goldberg and Levy, 2014). We then looked at the most overrepresented words in one corpus and compared them with the most similar and most underrepresented words in the other corpus. We focused on the top 300 most overrepresented words of the 5000 most used words. For each of these, we looked at the top 20 words that are closest in embedding space, and multiplied their similarity score in the embedding (i.e. how close they were in the vector space, from 0 to 1) with the Log-Likelihood score that shows how much the word has increased in use in later messages. This approach does have its limitations, as embeddings at times include opposite words, but it is an effective approach to identify community-specific terms that have replaced more commonly used words.

Result

Language convergence

Central to discursive community formation is the idea that communities tend to develop an internal culture, built through sustained, low-intensity exchange of messages, in which particular ideas, words, stories, and beliefs come to symbolize membership and connection to the community. The words signal belonging, and to learn this internal language is thus central to becoming part of the community and successfully participating in its discursive rituals. This suggests that users who enter the community need to learn this community language, and that learning this culture is necessary to attain the emotional energy that drives users to continue their participation. This, in turn, can be seen as users *converging* on the community language: coming to take up the discourse, themes, and stories of the community.

To examine whether users indeed converge with the language of the community, we looked at the cosine distance between new members and the community, as a function of how many posts they have contributed. Figure 2 shows the convergence of users by their post number. The results are quite striking. The members begin far from the forum language, but relatively quickly converge as they engage with the community. After about

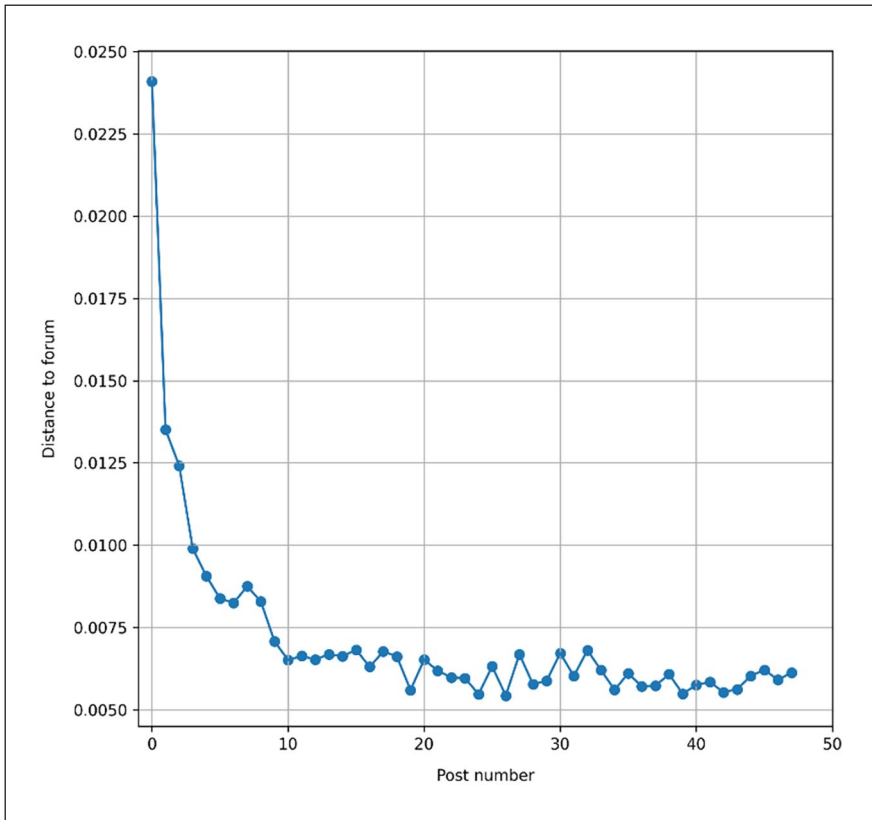


Figure 2. This figure shows the cosine distance between members' posts, in posting order, with the overall community language, over time. As the figure shows, new members quickly converge on the forum discourse. (As cosine distance is not a statistical measure, a confidence interval cannot be provided.)

20 posts, almost complete convergence has taken place. This suggests that members quickly absorb the defining discourse of the community.

Members who fail to take up the language should furthermore be expected to be more likely to leave the community, as successful participation in conversation rituals requires adopting the community discourse. To see this, we compared the posts of the members who stayed with those who did not, to see the extent to which their posts differ. We looked at members who posted only a few messages and compared their trajectories with members who became long-term contributors.

As Figure 3 illustrates, members who only remain for a few posts are significantly further away from the forum discourse in the initial posts than those who become long-term members. It takes about five posts for the short-term members to reach the place where the long-term members begin in their first posts. After this, these members appear to cease their convergence with the forum language, instead moving further away, before leaving the forum.

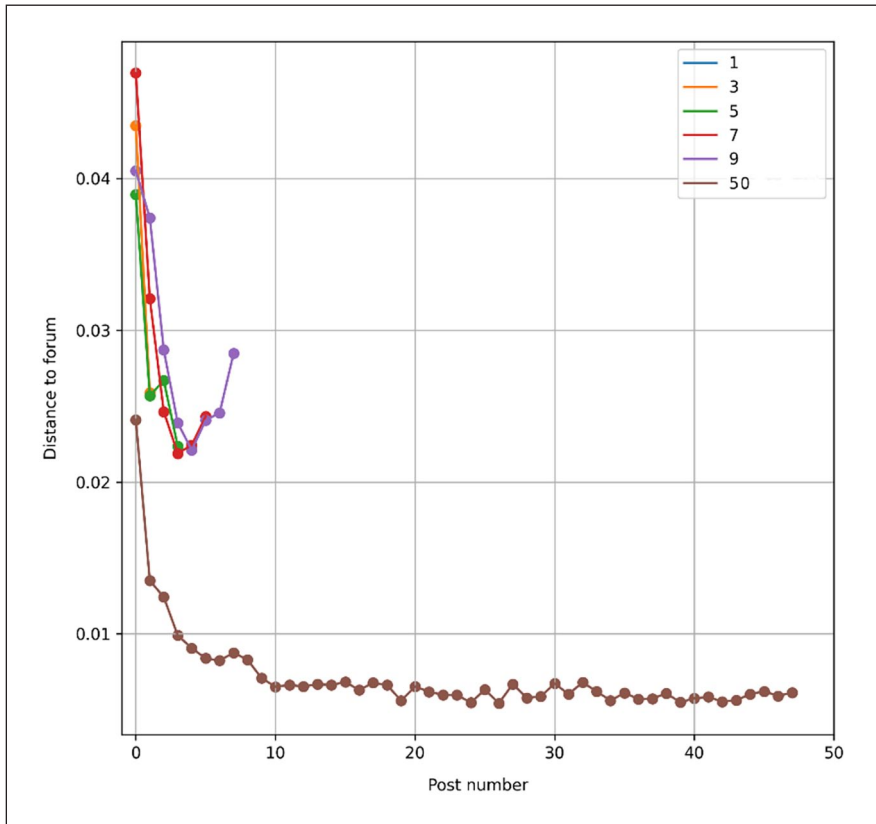


Figure 3. This figure shows the cosine distance between members' posts, in posting order, with the overall community language, over time. As the figure shows, new members quickly converge on the forum discourse. Members who leave the forum after a smaller number of posts start further from the forum language, and tend to diverge further in the last posts before leaving the forum. (As cosine distance is not a statistical measure, a confidence interval cannot be provided.)

Content of lexical change

A second central idea of discursive community formation is that the language of the community contains within it a worldview. Studying the discourse of a community can therefore throw light on the ideology of the community—their definitions of good and evil. To do so, we must examine what is characteristic of the community's language—what changes in the member's language as they converge with the community. To examine this, we turn to inductive methods to examine the content of the lexical change between early and later users. Figure 4 shows the most log-likelihood overrepresented words when comparing early and later posts. Table 1 shows the community-specific terms that have replaced more commonly used words. As these methods inductively capture aspects of a high-dimensional shift, some of the results are easily interpreted, while others

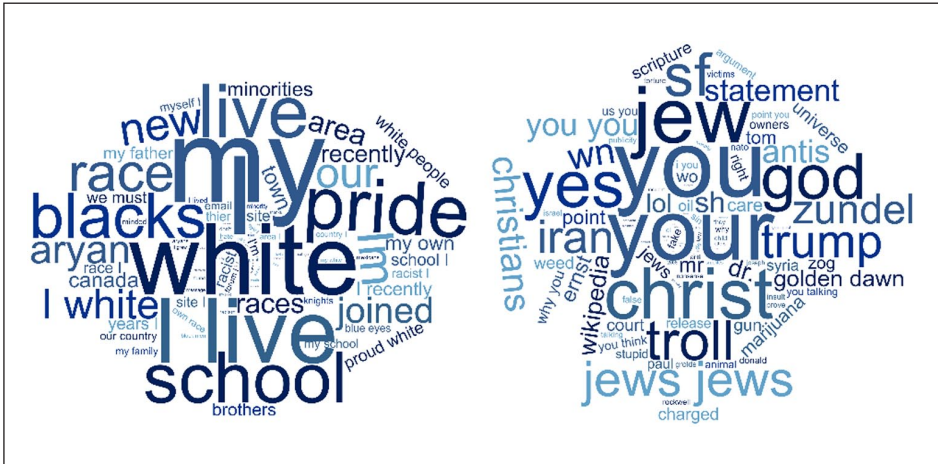


Figure 4. The left shows the most Log-Likelihood overrepresented words among new members, the right among the established members. These are not traditional word clouds, in which the word size is proportional to word frequency, since that the size of word is proportional to its log-likelihood overrepresentation of the words. The included words are in the top-5000 list of the most common words from both corpuses. Includes bigrams, which means that some words are repeated. The corpuses contain 5,937,177 tokens.

require closer examination. We will now look at some of these shifts to seek to understand their changes that they represent.

The first thing to note here is the shift in the use of pronouns and “indexical” statements—such as “you,” “me,” “here,” and “this”—which both scholars of discourse analysis and ethnomethodologists point to as important sites through which identity and interpersonal relations are expressed (Fairclough, 1989; Fairclough and Wodak, 1997). The use of these words can reveal how the authors view themselves and their relationship with their audience: the word “I,” for instance, is suggestive of a sense of individuality, whereas the use of “we” suggests that the author views themselves as representing something larger than themselves. Indexical terms are also indicative of a sense of being situated or co-located in a particular setting in which meaning and value are embedded.

Figure 4 and Table 1 show a clear fall in the use of first-person singular (“I,” “my,” “Im”) as members become established in the community. Table 1 shows that these words are replaced by second-person plural (“you,” “your”), which is used to refer to the Stormfront community, as well as the word “wn,” for “White Nationalist.”

To draw this out in more detail, we examined how the proportion of posts containing the chosen words changes as the user makes more contributions to the forum (see Figure 5). These graphs show a shift in pronouns which suggests that the members refer to themselves less as individuals, and more as situated within a community—with “you,” “your,” and “sf” (short for “Stormfront”) being used to mark their situatedness within the community. While the literature (e.g. DiMaggio et al., 2018) suggests that we should expect an increase in the use of “we” or “us” as individuals come to see themselves as part of a community, this does not appear to be the case. Similarly, there is only a slight increase

Table 1. This shows the change of words that are used with a similar meaning as members become more established in the community. The replacement words are selected by their log-likelihood overrepresentation multiplied by their word embedding similarity, with a threshold of 5 being used for inclusion. Words that have no new words replacing them are left out.

	Replacement words
my	your
i	you, lol, they
white	nonwhite
myself	you, wn
blacks	groids, they, negroes, nigs
black	groid, negro, nonwhite
im	you, lol
our	your
aryan	aryamsm
minorities	they
site	sf, blog
slavic	uralic
hey	lol, yeah, haha, yea, kidding, nah
dislike	think, disagree
feel	think
ive	sf
government	zog
edl	uaf, efp
media	msm, zog
information	evidence
turks	armenians, azeris, tatars
wp	wn
movement	party, wn, wnism
guys	guy, lol, you, haha
people	they, wns
politically	morally
replies	comments, comment
caucasian	australoid, arabids
enemy	zog

in the use of “they”/“them.” Closer examination of the use of these terms suggests that members do not write as if representing the community—through an inclusive “we”—but rather as if *addressing* the community—an inclusive “you” or “sf.” This unexpected finding speaks to how users perceive themselves in relation to the online community—implying that users view themselves as *situated within* the community, rather than as its official representatives.

We also see a move from mainstream terms to community-specific vernacular and themes, which function as markers of community belonging, and contain within them the White supremacist ideology of the forum. Figure 4 and Table 1 show a number of such community-specific terms, including “anti-white” and “white genocide,” which

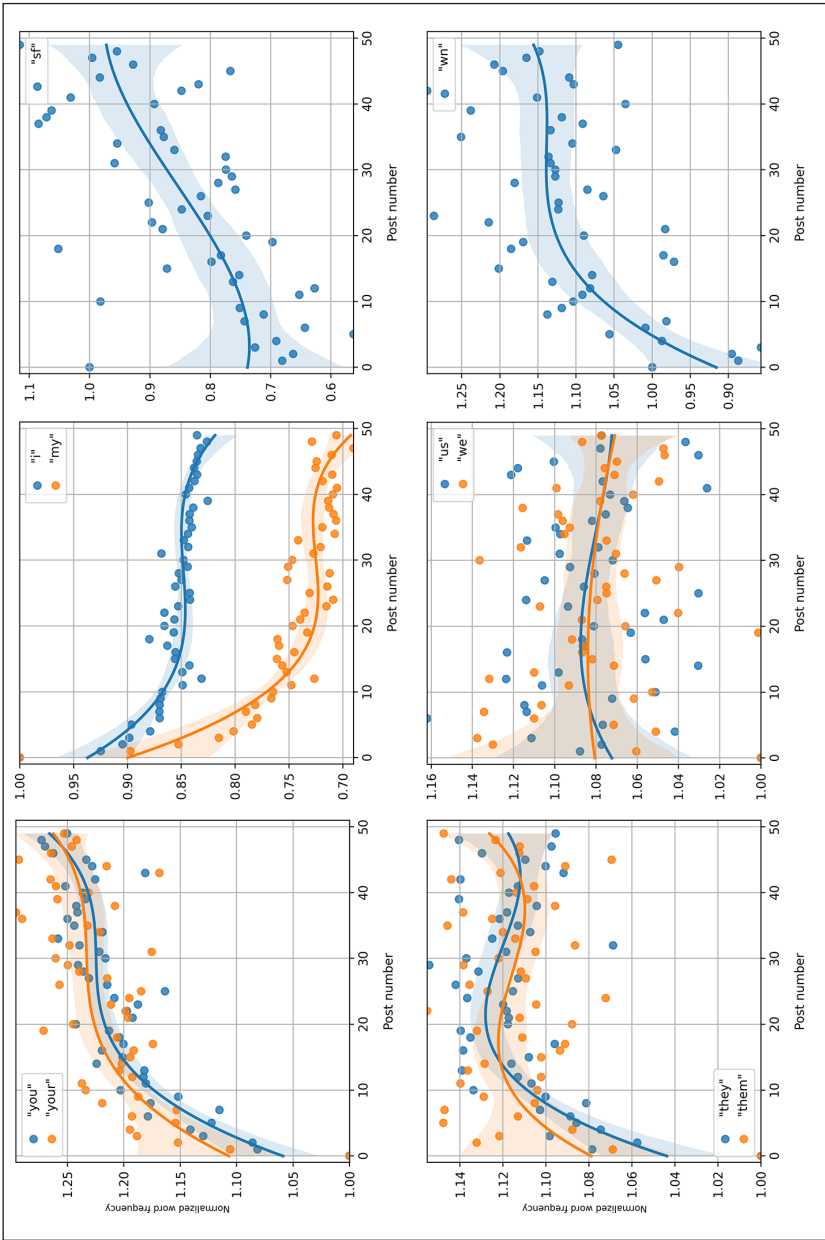


Figure 5. The fraction of the words in the first, second, third, and so on, posts of members who write at least 50 posts, normalized by the fraction of posts in the first message, to show the relative increase or decrease. The data are estimated by an order-3 polynomial linear regression model with 0.95 confidence interval.

emphasize the community's underlying ideological premise that White people are the true victims of racism, and which point to the Great Replacement conspiracy theory. Another example is the change in use from "government" to "zog"—short for "Zionist Occupational Government," referring to an anti-Semitic conspiracy theory claiming that Jews secretly control Western governments. "Media" is similarly replaced with "msm" ("mainstream media") or, again, "zog."

Such shift to a community-specific vernacular is also visible in words that are used to refer to Black people (see Figure 6). The commonly accepted mainstream terms "black"/"blacks" go through a precipitous fall in frequency as members engage with the forum. Some of this fall comes from a thematic move in focus toward Jews as the main out-group, reflecting the forum's strong anti-Semitic tendencies. But as Table 1 shows, these terms are also replaced with community-specific terms, such as "negroes," "groids," or "nigs" (it should be noted that the forum does not allow use of the more common racist N-word).

The use of "negroid" references historical race theory, suggesting that humankind can be divided into different races. A similar shift can be seen in references used for a number of national and ethnic groups, where over time Stormfront members learn to emphasize ethnicity over nationality, and race over ethnicity. For instance, "turks" are replaced by "armenians," "azeris," "tatars," that is, ethnic groups living in Turkey. Similarly, "white" and "caucasian" both fall in use, being replaced by "australoids" and "arabids." "Australoid" was a race-theoretical term for people of Australia, Melanesia, and parts of Southeast Asia, whereas "Arabids" was used to capture a racial division between peoples of Semitic ethnicities and peoples of other ethnicities. These terms thus serve to separate Jewish people from White people, suggesting that they constitute a separate racial group. This inductive analysis thus illustrates how the discourse of the community contains within it a racist ideology and political subjectivity.

Discussion

Echo chambers—here used in the broad sense of digital spaces where like-minded people can meet—have been a central focus of academic attention and seen as a potential driver in the recent rise of "sectarian" political polarization (Finkel et al., 2020; Mason and Kalmoe, 2022). In this article, we have drawn on Durkheim and Collins to propose a theory of what is in fact taking place inside these digital spaces.

We adapted Durkheim and Collins' work to suggest that echo chambers drive *discursive community formation*: they provide space for social rituals that instill in participants a sense of social membership and intersubjectivity through the elaboration of a distinct internal culture founded on the construction of a difference. As social media allow us to meet with like-minded individuals under the banner of a shared interest, these meeting spaces over time produce collective identities. The much-observed playful linguistic innovation that characterizes these communities fulfills the function of demarcating community boundaries and defining a linguistic capital within the community.

To be a part of a community is thus to develop a shared reality. We acquire the language of a community, absorb its symbols, and employ these to perform storytelling about ourselves, thus creating links between our personal identity and our community

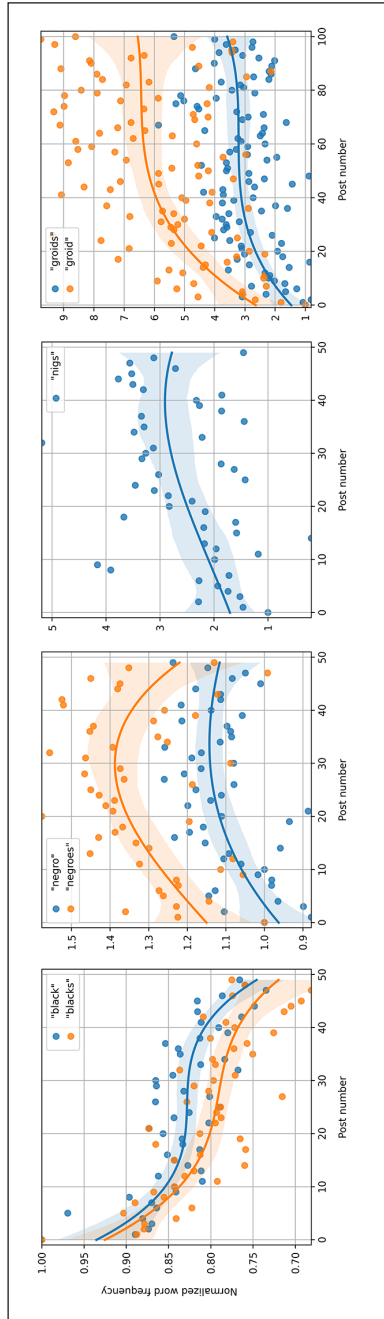


Figure 6. Focusing on the shift from “black”/“blacks” to community-specific jargon.

(Danescu-Niculescu-Mizil et al., 2013; Kleinberg et al., 2020). These stories about ourselves, our role in the world, and our link to our community are simultaneously stories about the world which guide our action.

Discursive community formation thus provides a potential causal mechanism that links digital media to the rise of sectarian political polarization. First, it leads to the strengthening of political identities that is associated with sectarian and identity-based forms of political polarization (Finkel et al., 2020; Klein, 2020; Mason, 2018). The notion of “online self-radicalization” is thus misleading: *au contraire*, these online spaces are profoundly social, and the process of online radicalization should be understood as a process of socialization. Furthermore, this suggests that echo chambers and broader media fragmentation enable the observed strengthening of partisan identities that has been linked to the rise of affective or sectarian polarization.

Second, discursive community formation links beliefs and stories that instantiate political subjectivities to our social identity; such beliefs and stories thereby become sacred, functioning as links between us and our community, and are, therefore, part of that which is taken for granted and beyond question. Such sacred beliefs have in recent years received significant scholarly attention within social psychology, which has shown how various deep-rooted psychological mechanisms protect these beliefs from being challenged (Kahan, 2017). Our identities shape our cognition through mechanisms such as “confirmation bias,” “deductive,” and “motivated reasoning”—in which our objective judgment and our rationality are affected by our identities and interests (Johnson-Laird, 2006; Nickerson, 1998; Wood and Porter, 2016). This type of “Identity-Protective Cognition” (Kahan, 2017) is understood as a way of avoiding dissonance and estrangement from one’s social group by subconsciously resisting any information that threatens the group’s defining values. In short, such sacred ideas are not fully subject to rational interrogation, as they operate in the realm of social identity rather than that of rational deliberation. Discursive community formation suggests that echo chambers naturally lead to the elaboration of such sacred ideas, thus suggesting a potential link between fringe digital spaces and the rise of misinformation.

Limitations and future research

The current study suggests a potentially important perspective on the effects of fringe digital spaces. However, it has some limitations that may be noted to inform future research.

First, our operationalization of Collins’ theory should be read as a first tentative step. For instance, Collins’ definition of an interaction ritual is inclusive of what can be construed as a symbol—vernacular, stories, images, discourses, beliefs, values, and attitudes can all come to shape a community. The bag-of-words approach adopted in this article, however, only captures changes in the frequency distribution of words, while discarding relationships between words and grammatical structures. Such an approach can be powerful and informative in respect of discourses, but in focusing solely on the distribution of words, it cannot capture stories or more complex language games. Future research may thus expand on this work by including other phenomena such as images, emotional energy, or complex discourses, and employ more sophisticated NLP methods.

Second, while examining what is present in conversations in digital communities allows us to study the construction of a shared reality and discursive playground, identifying ideas that come to be taken for granted or seen as being beyond reproach often involves examining what is *not* in the text. Ideas that are not up for debate, naturally, do not tend to come up in debate—at least not directly. For Stormfront, this may concern, for instance, the overarching belief in White supremacy, and the value of preserving White identity. These are constantly hinted at, but rarely raised explicitly. Future research may thus seek to develop methods to identify which ideas have acquired a sacred role for a community.

Finally, in this article, we have studied community formation in a White Power community. As the processes described are likely to play out in a broad range of digital spaces, future research may seek to examine the extent to which these processes also take place in more mainstream digital spaces.

Authors' Note

Petter Törnberg is also affiliated to University of Neuchâtel, Switzerland.

Anton Törnberg is also affiliated to University of Lund, Sweden.

Data Availability

The data used in this study have been made permanently publicly available for download at <https://doi.org/10.34740/kaggle/dsv/4176720>

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: PT would like to acknowledge funding from the Dutch Research Council (NWO) VENI (grant number VI.Veni.201S.006).

ORCID iDs

Petter Törnberg  <https://orcid.org/0000-0001-8722-8646>

Anton Törnberg  <https://orcid.org/0000-0002-5847-3033>

Note

1. It should be noted that more complex models that better capture user engagement with the forum and that would therefore show larger differences between long- and short-term users are likely to exist. For instance, it is likely that users who send their 50 posts over a prolonged period of time will have engaged more with the forum than users who send the same number of posts over a brief duration. However, we have opted for as simple a model as possible, and leave it to future research to examine what factors determine the pace of a user's socialization.

References

- Bail CA (2012) The fringe effect: civil society organizations and the evolution of media discourse about Islam since the September 11th attacks. *American Sociological Review* 77(6): 855–879.
- Beirich H (2014) White homicide worldwide. Southern Poverty Law Center Intelligence report. Available at: <https://www.splcenter.org/20140331/white-homicide-worldwide>

- Belew K (2018) *Bring the War Home: The White Power Movement and Paramilitary America*. New York: Harvard University Press.
- Benhabib S (1996) *Democracy and Difference: Contesting the Boundaries of the Political*. New York: Princeton University Press.
- Benwell B and Stokoe E (2006) *Discourse and Identity*. London: Edinburgh University Press.
- Beyer JL (2014) *Expect Us: Online Communities and Political Mobilization*. London: Oxford University Press.
- Borum R (2011) Radicalization into violent extremism I: a review of social science theories. *Journal of Strategic Security* 4(4): 7–36.
- Boukes M and Trilling D (2017) Political relevance in the eye of the beholder: determining the substantiveness of TV shows and political debates with Twitter data. *First Monday* 22(4): 1–22.
- Bourdieu P (1991) *Language and Symbolic Power*. New York: Harvard University Press.
- British Sociological Association (BSA) (2017) Statement of ethical practice. Available at: https://www.britisoc.co.uk/media/24310/bsa_statement_of_ethical_practice.pdf
- Bruns A (2019) *Are Filter Bubbles Real?* London: John Wiley & Sons.
- Clark M (2014) *To Tweet Our Own Cause: A Mixed-Methods Study of the Online Phenomenon "Black Twitter."* Chapel Hill, NC: University of North Carolina at Chapel Hill Graduate School.
- Collins R (1981) On the microfoundations of macrosociology. *American Journal of Sociology* 86(5): 984–1014.
- Collins R (2004) *Interaction Ritual Chains*. New York: Princeton university press.
- Colombo M and Senatore A (2005) The discursive construction of community identity. *Journal of Community & Applied Social Psychology* 15(1): 48–62.
- Crossley AD (2015) Facebook feminism: social media, blogs, and new technologies of contemporary US feminism. *Mobilization: An International Quarterly* 20(2): 253–268.
- Danescu-Niculescu-Mizil C, West R, Jurafsky D, et al. (2013) No country for old members: user lifecycle and linguistic change in online communities. In: *Proceedings of the 22nd international conference on world wide web*, Rio de Janeiro, Brazil, 13–17 May, pp. 307–318. New York: ACM.
- De Cillia R, Reisigl M and Wodak R (1999) The discursive construction of national identities. *Discourse & Society* 10(2): 149–173.
- DeCook JR (2018) Memes and symbolic violence: #proudboys and the use of memes for propaganda and the construction of collective identity. *Learning, Media and Technology* 43(4): 485–504.
- Della Porta D (2013) *Clandestine Political Violence*. London: Cambridge University Press.
- Diani M (1992) The concept of social movement. *The Sociological Review* 40(1): 1–25.
- DiMaggio P, Bernier C, Heckscher C, et al. (2018) Interaction ritual threads: does IRC theory apply online? In: Weininger EB, Annette L and Lizardo O (eds) *Ritual, Emotion, Violence: Studies on the Micro-sociology of Randall Collins*. London: Taylor & Francis, pp. 81–124.
- Durkheim E (1912) *The Elementary Forms of the Religious Life*. New York: Free Press.
- Fairclough N (1989) *Language and Power*. New York: Longman.
- Fairclough N and Wodak R (1997) Critical discourse analysis. In: van Dijk TA (ed.) *Discourse as Social Interaction*. New York: Springer, pp. 258–284.
- Finkel EJ, Bail CA, Cikara M, et al. (2020) Political sectarianism in America. *Science* 370(6516): 533–536.
- Fiorina MP (2017) *Unstable Majorities: Polarization, Party Sorting, and Political Stalemate*. New York: Hoover Press.
- Franzke AS, Bechmann A, Zimmer M, et al. (2020) *Internet Research: Ethical Guidelines 3.0*. Available at: <https://aoir.org/reports/ethics3.pdf>

- Futrell R and Simi P (2004) Free spaces, collective identity, and the persistence of US white power activism. *Social Problems* 51(1): 16–42.
- Gaudette T, Scrivens R and Venkatesh V (2020) The role of the internet in facilitating violent extremism: insights from former right-wing extremists. *Terrorism and Political Violence*. Epub ahead of print 16 July. DOI: 10.1080/09546553.2020.1784147.
- Goldberg Y and Levy O (2014) Word2vec explained: deriving Mikolov et al.'s negative-sampling word-embedding method. *arXiv*. Available at: <https://arxiv.org/abs/1402.3722>
- Guess A, Nyhan B, Lyons B, et al. (2018) *Avoiding the Echo Chamber about Echo Chambers: Why Selective Exposure to Like-minded Political News Is Less Prevalent Than You*. Miami, FL: Knight Foundation.
- Iyengar S, Lelkes Y, Levendusky M, et al. (2019) The origins and consequences of affective polarization in the United States. *Annual Review of Political Science* 22: 129–146.
- Iyengar S, Sood G and Lelkes Y (2012) Affect, not ideology: a social identity perspective on polarization. *Public Opinion Quarterly* 76(3): 405–431.
- Johnson-Laird PN (2006) Deductive reasoning. In: Nadel L (ed.) *Encyclopedia of Cognitive Science*. New York: John Wiley & Sons, pp. 109–136.
- Kahan DM (2017) *Misconceptions, misinformation, and the logic of identity-protective cognition*. Cultural cognition project working paper series no. 164. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2973067
- Keuchenius A, Törnberg P and Uitermark J (2021) Why it is important to consider negative ties when studying polarized debates: a signed network analysis of a Dutch cultural controversy on Twitter. *PLoS ONE* 16(8): e0256696.
- Klein E (2020) *Why We're Polarized*. New York: Simon & Schuster.
- Kleinberg B, van der Veet I and Gill P (2020) The temporal evolution of a far-right forum. *Journal of Computational Social Science* 4(1): 1–23.
- Knobel M and Lankshear C (2007) Online memes, affinities, and cultural production. In: Knobel M and Lankshear C (eds) *A New Literacies Sampler*. New York: Peter Lang, pp. 199–227.
- McCauley C and Moskaleiko S (2008) Mechanisms of political radicalization: pathways toward terrorism. *Terrorism and Political Violence* 20(3): 415–433.
- Maloney P (2013) Online networks and emotional energy: how pro-anorexic websites use interaction ritual chains to (re)form identity. *Information, Communication & Society* 16(1): 105–124.
- Mason L (2018) *Uncivil Agreement: How Politics Became Our Identity*. Chicago, IL: The University of Chicago Press.
- Mason L and Kalmoe NP (2022) *Radical American Partisanship: Mapping Violent Hostility, Its Causes, and the Consequences for Democracy*. Chicago, IL: The University of Chicago Press.
- Morales GDF, Monti C and Starnini M (2021) No echo in the chambers of political interactions on Reddit. *Scientific Reports* 11(1): 1–12.
- Nagle A (2017) *Kill All Normies: Online Culture Wars from 4Chan and Tumblr to Trump and the Alt-Right*. London: John Hunt.
- Nickerson RS (1998) Confirmation bias: a ubiquitous phenomenon in many guises. *Review of General Psychology* 2(2): 175.
- Papacharissi Z (2011) *A Networked Self: Identity, Community, and Culture on Social Network Sites*. London: Routledge.
- Pariser E (2011) *The Filter Bubble: What the Internet Is Hiding from You*. London: Penguin.
- Phillips W (2015) *This Is Why We Can't Have Nice Things: Mapping the Relationship between Online Trolling and Mainstream Culture*. New York: MIT Press.
- Phillips W and Milner RM (2017) Decoding memes: Barthes' punctum, feminist standpoint theory, and the political significance of #yesallwomen. In: Harrington S (ed.) *Entertainment Values*. London: Palgrave Macmillan, pp. 195–211.

- Polletta F and Jasper JM (2001) Collective identity and social movements. *Annual Review of Sociology* 27(1): 283–305.
- Reagle JM (2015) *Reading the Comments: Likers, Haters, and Manipulators at the Bottom of the Web*. New York: MIT Press.
- Scrivens R, Davies G and Frank R (2020) Measuring the evolution of radical right-wing posting behaviors online. *Deviant Behavior* 41(2): 216–232.
- Shifman L (2013) Memes in a digital world: reconciling with a conceptual troublemaker. *Journal of Computer-Mediated Communication* 18(3): 362–377.
- Sides J, Tesler M and Vavreck L (2018) *Identity Crisis: The 2016 Presidential Campaign and the Battle for the Meaning of America*. New York: Princeton University Press.
- Snow DA (2001) Collective identity and expressive forms. In: Smelser N and Baltes P (eds) *International Encyclopedia of the Social and Behavioral Sciences*. London: Elsevier Science, pp. 2212–2219.
- Sunstein CR (2009) *Going to Extremes: How Like Minds Unite and Divide*. London: Oxford University Press.
- Törnberg A and Törnberg P (2016) Combining CDA and topic modeling: analyzing discursive connections between Islamophobia and anti-feminism on an online forum. *Discourse & Society* 27(4): 401–422.
- Törnberg P, Andersson C, Lindgren K, et al. (2021) Modeling the emergence of affective polarization in the social media society. *PLoS ONE* 16(10): e0258259.
- Törnberg P (2022). Stormfront.org – White Power forum posts 2001-2020 [Data set]. Kaggle. Available at: <https://doi.org/10.34740/KAGGLE/DSV/4176720>
- Turkle S (2011) *Life on the Screen*. New York: Simon & Schuster.
- Tuters M and Hagen S (2020) (((They))) rule: memetic antagonism and nebulous othering on 4chan. *New Media & Society* 22(12): 2218–2237.
- van Haperen S, Uitermark J and van der Zeeuw A (2020) Mediated interaction rituals: a geography of everyday life and contention in Black Lives Matter. *Mobilization: An International Quarterly* 25(3): 295–313.
- van Vliet L, Törnberg P and Uitermark J (2020) The Twitter parliamentarian database: analyzing Twitter politics across 26 countries. *PLoS ONE* 15(9): e0237073.
- Vasquez EA, Wenborne L, Peers M, et al. (2015) Any of them will do: in-group identification, out-group entitativity, and gang membership as predictors of group-based retribution. *Aggressive Behavior* 41(3): 242–252.
- Wood T and Porter E (2016) The elusive backfire effect: mass attitudes' steadfast factual adherence. *Political Behavior* 41(1): 135–163.
- Zannettou S, Caulfield T, Blackburn J, et al. (2018) On the origins of memes by means of fringe web communities. In: *Proceedings of the internet measurement conference*, Boston, MA, 31 October–2 November, pp. 188–202. New York: ACM.
- Zuiderveen Borgesius F, Trilling D, Möller J, et al. (2016) Should we worry about filter bubbles? *Internet Policy Review: Journal on Internet Regulation* 5(1): 1–17.

Author biographies

Petter Törnberg is a postdoctoral researcher in human geography at University of Amsterdam and University of Neuchâtel. His current research focuses on using computational methods and digital data to examine political and cultural conflicts on digital platforms.

Anton Törnberg is Associate professor at the Department of Sociology and Work Science at the University of Gothenburg. His research chiefly focuses on digital data and social movements and combining computational methods with qualitative approaches.