



Veiled conspiracism: Particularities and convergence in the styles and functions of conspiracy- related communication across digital platforms

new media & society

1–23

© The Author(s) 2025



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14614448251315756

journals.sagepub.com/home/nms**Kilian Buehling** 

Weizenbaum Institute for the Networked Society, Germany; Freie Universität Berlin, Germany

Xixuan Zhang 

Weizenbaum Institute for the Networked Society, Germany; Freie Universität Berlin, Germany

Annett Heft 

Weizenbaum Institute for the Networked Society, Germany; Freie Universität Berlin, Germany; Eberhard Karls University of Tübingen, Germany

Abstract

Digital communication venues are essential infrastructures for anti-democratic actors to spread harmful content such as conspiracy theories. Capitalizing on platform affordances, they leverage conspiracy theories to mainstream their political views in broader public discourse. We compared the word choice, language style, and communicative function of conspiracy-related content to understand its platform-dependent differences and convergence. Our cases are the conspiracy theories of the New World Order and Great Replacement, which we analyzed on 4chan/pol/, Twitter, and seven alternative US news media longitudinally from 2011 to 2021. The conspiracy-related texts were comparatively analyzed using a multi-method approach of computational and quantitative text analyses. Our results show that conspiracy narrations are increasingly present

Corresponding author:

Kilian Buehling, Institute for Media and Communication Studies, Freie Universität Berlin, Garystr. 55, Berlin 14195, Germany.

Email: kilian.buehling@fu-berlin.de

in all venues. While language differs vastly between platforms, we observed a style convergence between Twitter and 4chan. The results show how more coded language veils the spread of racist and antisemitic content beyond the so-called dark platforms.

Keywords

Affordances, alternative media, conspiracy theories, far-right, fringe platforms, mainstreaming

Introduction

Social media platforms and digital communication venues provide central infrastructures for hyper-partisan and anti-democratic actors who spread harmful content and extreme speech (Pohjonen and Udupa, 2017) while circumventing traditional gatekeepers (Engesser et al., 2017). This access to large online audiences is used to mainstream extremist ideas to shift the overall discourse (Rothut et al., 2024). An important vector to mainstream xenophobic and antisemitic beliefs are conspiracy theories, whose online prevalence has attracted increasing scholarly (Mahl et al., 2022a), political (European Commission, 2018), and societal attention in recent years.

Beliefs in conspiracy theories are widespread and can have profound consequences. Of the Americans surveyed in 2021, 26% believed George Soros was behind a hidden plot to destabilize the American government (Uscinski et al., 2022). In Germany and Great Britain, more than 40% of survey respondents polled in 2016 believed that the government was deliberately hiding the truth about how many immigrants live in the country (Smallpage et al., 2020). Studies have shown that conspiracy thinking can be related to institutional distrust, discrimination, and accelerated radicalization and violence (Jolley et al., 2020). Conspiracy theories closely tied to far-right ideology can pose threats to democratic institutions, societies, and individuals. In extreme cases, they are used to justify violence. The terrorist who attacked Muslims in Christchurch, New Zealand, explicitly adhered to the “Great Replacement/White Genocide” conspiracy theories (Ekman, 2022). These theories suggest a hidden plan to minimize the White population and Christian culture by fostering the migration of people of color and reducing birth rates.

To date, many case studies on conspiracy theories on single platforms within short periods have been conducted (Mahl et al., 2022a). The diffusion within and partly across social media platforms has been analyzed from an “infodemic” perspective, that is, the diffusion of novel conspiracy narrations to places where they were not present before (for a review, see Heft and Buehling, 2022). Scholars have highlighted that conspiracy theory narrations are shaped by the characteristics of the platform on which they are uttered, captured by concepts such as platformed conspiracism (Mahl et al., 2023) or platformed antisemitism (Riedl et al., 2024).

However, longitudinal cross-platform studies of conspiracy-related communication that take into account the interplay of platform affordances and their distinct appropriation remain underexplored. These research strands are significant in understanding how platform affordances shape the discourse of conspiracy-related communication. The

development of persistent extreme speech over time, either becoming more explicit or veiled, may gradually shift the audience's perception of acceptable content and increase the risk of mainstreaming far-right narratives (Rothut et al., 2024).

While conspiracy theory narration refers to contents that offer a “proposed explanation of some historical event (or events) in terms of the significant causal agency of a relatively small group of persons – the conspirators – acting in secret” (Keeley, 1999: 116), the distribution of and interaction with this type of content invites a full range of public communication on and about (alleged) conspiracy theories. These communicative functions include further contributions to conspiracy theory narration, engagement in counter-narration and debunking, and employing neutral forms of observation. Thus, we investigated the explicitness or implicitness of conspiracy-related content and how it is communicated.

We expect the styles and functions of conspiracy-related communication in online venues to vary due to the intertwining of platform characteristics and values within platform-specific use cultures. While all communication venues might provide room for narration, counter-narration, and neutral observation of conspiracy theories, platform characteristics may correlate with a higher likelihood for one function or the other. Thus, our research questions are as follows:

RQ 1a: How does the language style of conspiracy-related communication differ between digital communication venues?

RQ 1b: What are the over-time dynamics of differences in the language styles of conspiracy-related communication between digital communication venues?

RQ 2: How do the functions of conspiracy-related communication differ between digital communication venues?

In this study, we elaborated our theoretical assumptions and answered the research questions using an empirical analysis of conspiracy-related content in three communication venues. The platforms selected, namely 4chan, Twitter (now called X), and partisan media sites, were unique, with venues considerably varying in governance and norms, user cultures, and characteristics of creators and audiences. We analyzed similarities and differences in conspiracy-related communication functions and styles across the venues, focusing on two closely interlinked, salient, and transnationally persistent conspiracies of the “New World Order (NWO)” and “Great Replacement/White Genocide” from 2011 to 2021. Using computational text analysis, communication styles were measured by examining the variation and distinctness of semantic features across venues. The communicative functions of conspiracy-related content were classified in a manual content analysis. We further examined communication particularities as entry points for the translation (Ekman, 2022) of conspiracy theories from fringe use cultures to venues that address broader audiences and societal discourses at large.

Our paper is structured as follows: First, we outline how platform features and their appropriation in particular contexts can create distinct use cultures and thus different characteristics of conspiracy-related talk. We address platform-internal particularities and how they might be reflected in communication styles and functions. Second, we

provide an overview of the design and methods of our study. Next, we address narration style features and their longitudinal development. Finally, we highlight the communicative functions of this content. After discussing the results, we highlight the contribution of our findings to contemporary discussions about platform cultures, affordances, and platform moderation policies.

Platform characteristics and the shaping of conspiracy-related communication in distinct user cultures

Platform vernacular and communication styles

Each digital platform and communication venue is characterized by distinct technical features (Bossetta, 2019), governance approaches, and norms that fundamentally impact the authors, users, and the forms of communication they attract (Frischlich et al., 2022; Quandt, 2018). Individuals who engage in digital communication base their communication strategies on the available technical possibilities and tailor their communication styles to the norms and expectations prevailing in a particular environment. In the context of conspiracy-related communication, conspiracy entrepreneurs aim to connect conspiratorial narratives to broader populations with strategies such as emotionalization, pop culture, and Manichean references (Leal, 2020; Oliver and Wood, 2014).

A tangible manifestation of the interactions between the platforms' regulation policies and user behavior is the platform-specific communication style, defined as "linguistic structures below the discursive level, such as grammar, phonology, and lexis" (Bucholtz and Hall, 2005: 597). While the individual choice of vocabulary and other paralingual features (Khalid and Srinivasan, 2020) does not necessarily change a message's factual content, it can influence the recipient's interpretation (Blankenship and Craig, 2011) and mark a group identity through the repeated performance of stylistic features (Bucholtz and Hall, 2005).

Gibbs et al. (2015) used the concept of platform vernacular to describe how platform features shape the use of a unique, shared but dynamic set of grammars, styles, and logics that emerge from the ongoing interactions between users and platforms. In their conceptualization of platformed conspiracism, Mahl et al. (2023) built on the concept of Gibbs et al. to explain the influence of incentivization through monetization and the adaptation of users to platform moderation on the spread of conspiracy-related content. The possibilities of how explicitly producers of conspiracy-related communication can spread their message on a given platform – disinhibited by content moderation mechanisms or inhibited by anonymity and ephemerality – reflect "affordances of extreme speech" (de Keulenaar, 2023). As a consequence, extreme speech (Pohjonen and Udupa, 2017), which includes conspiracy theories (Deem, 2019), tends to be fragmented along the borders of digital communication outlets, each governed by diverging understandings of tolerable content. The stylistic adaptation of conspiracy-related communication, either by strategic injection or unintentional reproduction by platform users, has been described in the context of mainstreaming far-right narratives (Rothut et al., 2024).

Communication style is manifested in the use of particular terms such as loaded, slang, or euphemistic terms (Klein, 2023) and linguistic features such as linguistic

complexity (Visentin et al., 2021), emotionality (Zollo et al., 2015), or toxicity (Hoseini et al., 2023). Being adaptive to content moderation, users who post on contentious topics on mainstream platforms and are at risk of being deamplified or banned use neologisms, lexical variation, or dog whistles to circumvent moderation efforts and deliver their content to their audience (Moran et al., 2022). Unique neologisms might be specific to distinct platforms and thus easier to detect. Meanwhile, dog whistles, that is, the encoding of a message into words that have a diverging meaning to the general audience (Abidin, 2021), require a more profound knowledge of specific user cultures. Actors who deliberately veil their extreme speech are themselves navigating the contradiction of remaining visible to their audience (as opposed to violating a platform's terms of service) and using language that obfuscates their content to the degree that is decipherable by most users in the know (Abidin, 2021).

User cultures on 4chan and Twitter and the hybrid characteristics of alternative news media

De Zeeuw et al. (2020) highlighted that anonymous and ephemeral imageboards such as 4chan allow for highly vernacular humor and white supremacist rhetoric. In a discussion forum, users can start threads or contribute to existing ones in a room-based architecture. By abstaining from algorithmically curated feeds and relying on user's self-sorting into more homogeneous communication spaces, room-based platforms can provide a sheltered breeding ground for subcultural milieus in which members signal their belonging through insider abbreviations and slang that is difficult to decode for outsiders (Nissenbaum and Shifman, 2017). 4chan attracts users whose demographic is described as predominantly white and male users and those concerned with far-right politics (de Zeeuw and Tuters, 2020). De Zeeuw et al. (2020) argued that anonymity in online spaces disinhibits behavior that is otherwise deemed socially undesirable (de Keulenaar, 2023; Suler, 2004). In same-topic discussions on 4chan and so-called mainstream platforms such as Twitter, different words are used, or the same words are used with different meanings. De Zeeuw et al. (2020) attributed this difference to mainstream social media affording more authentic user behavior via non-anonymous posts, consistent content moderation, and news media content driven by their own logic of newsworthiness and broadcasting.

The micro-blogging platform Twitter requires users to register an account and a user name to which content is attributed (Frischlich et al., 2022). Its networked architecture is generally more open and directed toward interconnected and distributed forms of communication with greater outreach. Twitter had a reach of 25% among the US online population in 2023, according to the Reuters Institute online survey (Newman et al., 2023). It is attractive among political and media actors who use the platform to disseminate and multiply their content (Jungherr, 2016). Larger platforms such as Twitter (or X) and Facebook are required to invest in content moderation to prevent harmful, hateful, or otherwise illegal content (Nunziato, 2023). While smaller platforms face similar obligations, very large online platforms are subject to greater scrutiny.

All the discussed platforms enable user-driven content production, but the broader reach and distribution of conspiracy-related content also depends on uptake in the

broader digital media ecology, including legacy and alternative news media outlets. Their reporting on and contextualization of conspiracy theories are fundamental to their connectivity to broader publics (Benkler et al., 2018; Bruns et al., 2022; Forberg, 2022). Alternative media generally “represent a proclaimed and/or (self-)perceived corrective, opposing the overall tendency of public discourse emanating from what is perceived as the dominant mainstream media in a given system” (Holt et al., 2019: 862). This broad definition includes alternatives with different aims, functions, and levels of partisanship from left to right (Schwaiger, 2022), ranging from progressive counter-hegemonic media, which long constituted the conceptual core, to more recent alternatives that, in the hyper-partisan extremes, react to perceived hegemonic politics and media with stark opposition and practices that (over-)stretch the boundaries of journalism (Nordheim and Kleinen-von Königslöw, 2021). Yang (2020) conceptualized such radical media on the political far-right as hybrid info-political organizations that strategically switch their repertoires; they navigate between a journalistic logic of professional reporting and a role as a political movement actor (Heft et al., 2021; Mayerhöffer and Heft, 2022). Adhering to both logics in their style and visual appearance enables such alternative media to gain a perception of professional legitimacy while retaining their partisan credibility (Heft et al., 2020; Mayerhöffer and Heft, 2022). In this respect, alternative media bridge partisan communities and broader societal discourse, as they rhetorically embed, reframe, and comment on current events (Ekman, 2019). Right-wing alternative media have been shown to distribute conspiracy-related content (Boberg et al., 2020; Schwaiger, 2022).

Conceptualization and expectations

We summarize our observations in the following expectations:

Regarding *RQ 1a*, we expect conspiracy-related communication styles to vary across digital communication venues, given their differences in features, users, the users’ adaptation to platform features, and the different vernaculars that emerged from these inter-related context conditions.

We conceptualized communication styles by analyzing 1) *lexical items* (i.e. terms used in conspiracy-related communication). Given the differences in which user groups inhabit particular platform niches devoted to conspiracy-related talk and overall platform characteristics, we expect the terms to differ across platforms. Quantitatively, the words used on the different platforms are expected to vary overall, as different user cultures bear different lexical registers. Qualitatively, we expect a higher proportion of explicit, core “conspiracy” terms on 4chan than in the other communication venues such as Twitter and alternative media, on which we would expect less explicit terms.

The terms used when discussing conspiracy theories are closely related to *how* they are communicated. We thus 2) conceptualized the “how” as style elements of conspiracy-related content via the following dimensions:

- a) Certainty of language, denoting a user culture in which the statements lack nuance owing to dichotomous phrasing or the assertion of high assurance of a respective

position. Owing to heightened radicalness and homogeneous user composition, we expect the highest certainty on 4chan and the lowest certainty and thus the most nuance in alternative media articles.

b) Linguistic complexity comprises the use of complicatedly constructed sentences and words. We expect complexity to be lowest on Twitter owing to the character limit of each tweet. Complexity should be highest in alternative media reporting because of its professionalized language.

c) Emotionality refers to the prevalence of words expressing emotions. Assuming that the phrasings of positive and negative emotions are more likely in expressing personal opinions and less in descriptive texts, we expect a lower emotionality in alternative media articles. Owing to the heterogeneous user composition on Twitter compared with the relatively homogeneous, interest-based 4chan community, the highest emotionality is to be expected on Twitter.

Even though the user groups on platforms cultivate distinct behaviors and vernacular, which, to a certain degree, isolate them from the outside (Abidin, 2021), social media users can observe and be active on various platforms. Forberg (2022) found that 4chan users actively employed Twitter's affordances to astroturf the QAnon conspiracy theory and inject their vernacular into the mainstream discourse. Furthermore, user cultures and platform vernaculars are deeply intertwined so that when the vernacular diffuses to other platforms or user groups, it can be misunderstood or change its meaning (Hagen and de Zeeuw, 2023; Tuters et al., 2018). The complex cross-platform information diffusion patterns do not necessarily result in unidirectional transmission from fringe to more mainstream platforms. The process of fringe ideas gaining popularity after being mentioned by media actors (Bail, 2012) is further amplified when media outlets report about them (Benkler et al., 2018; Forberg, 2022). Even neutral or critical reporting runs the risk of unintendedly drawing public attention and thus amplifying conspiracy theories (Bruns et al., 2022; Phillips, 2018), as alternative and legacy news sources are a common reference point for conspiracist claim-making in far-right information ecosystems (Buehling and Heft, 2023; Schatto-Eckrodt et al., 2024).

To answer *RQ 1b*, we analyzed how and when conspiracy-related communication styles converge between platforms. We expect communication styles to mutually influence each other, with alternative media being the least influenced owing to their authors' constant navigation between their journalistic and movement-oriented partisan roles.

To answer *RQ 2*, we conceptualized differences between communication venues concerning three functions of conspiracy-related content: (1) conspiracy theory narration (e.g. texts that develop, uncritically reproduce, or propagate a conspiracy); (2) counter-narration (i.e. texts that argue against or refute the conspiracy theory); and (3) neutral and distanced report about a conspiracy theory without suggesting a stance in favor of or against it. These functions represent "programs," "anti-programmes," and "efforts at neutrality" (Rogers, 2017: 83).

Given the different platform characteristics, we expect different proportions of these functions on the three communication venues: 4chan should be most fertile to afford conspiracy narration, and alternative media should primarily enact neutral observation in

their reporting if we take their mimicry of professional journalistic norms for granted, accompanied by commentary framing current events in terms of right-wing conspiracy theories. With its mixed user base, the networked platform Twitter can be expected to provide a space for all functions of conspiracy-related communication.

Study design and methods

Case selection

To facilitate a longitudinal analysis of the style and functions of conspiracy-related communication, the narratives of the *NWO* or *Illuminati* and the *Great Replacement* or *White Genocide* were selected. Both theories can be considered “systemic conspiracies” (Barkun, 2013: 6), including the narrative of a secretive group plotting to steer global events to reach or maintain world domination.

The *NWO* conspiracy theory is one of the most prevalent systemic post-Holocaust conspiracy theories, drawing from antisemitic tropes of a Jewish world conspiracy, while partly replacing the overt accusation of Jewish people by the *Illuminati* or *Zionists* (Simonsen, 2020)¹. The theory and its use for explaining political events and incorporating conspiratorial sub-narratives was popularized worldwide (Girard, 2020). The *Great Replacement* and *White Genocide* conspiracy theories are closely related but still distinct from *NWO*. *Great Replacement* was popularized and adopted by the far-right in the 21st century as a conspiracy theory, suggesting that secretive “globalists” are planning to shift the ethnic and religious proportions in Europe by allowing the immigration of non-Christian refugees from Asia and Africa (Ekman, 2022). This narrative mirrors *White Genocide*, which is more prevalent in the US far-right, relating to an alleged purposeful facilitation of non-white migration to North America and, consequently, the inhibition of a desired white supremacy (Cosentino, 2020).

While both conspiracy theories are distinct, their narratives overlap in the attribution of the conspirators (i.e. *Illuminati* or *globalists*) and their alleged means of secretly puppeteering democratic governments. Both were included because they share the characteristics of long-term prevalence, an apparent appropriation by the far-right and an anti-institutional element that suggests that the perpetrators are already in power (Butter, 2018; Ekman, 2022). Furthermore, they are prevalent in online discourse (Cosentino, 2020; Mahl et al., 2021) and are used to justify the atrocities committed by contemporary far-right terrorists (Ekman, 2022).

Platforms and sample

To collect data from different platforms over time, we developed a comprehensive and precise dictionary to provide a good recall concerning conspiracy theories (Gründl, 2022; Puschmann et al., 2022). The keywords in the dictionary must be applicable across time and platforms to be robust and ensure equivalent data collection along these dimensions (Heft et al., 2023). Data were collected in a broader comparative project on the diffusion dynamics of conspiracy theories, in which we collected English- and German-language data on Twitter, 4chan/pol/, and German/US-American alternative

news media. This study focused on English-language samples of tweets, 4chan/pol/ posts, and alternative news media articles. The dictionary was constructed using the following multistep procedure to accommodate the requirements of equivalent data collection: *theory-driven keyword selection*, resulting in 215 individual terms; *validation* of the sample data collected using the seed dictionary; *computational dictionary expansion* based on the fine-tuned word embedding model GloVe and Local Graph-based Dictionary Expansion (Schindler et al., 2024); manual *validation* of the expanded keywords; and addition of 22 terms to the dictionary (see Appendix A for more details on the dictionary construction).

Data collection and pre-processing

To ensure an equivalent data collection across time and platforms, we randomly selected one calendar week from each half-year period between 2011 and 2021. Randomly collected, consecutive 7 day intervals were chosen to ensure the representativeness of random sampling (Kim et al., 2018) regarding the evolution and variation of language styles while mitigating possible biases related to seasonality (weekly and yearly) or event-based posting spikes (Naab and Küchler, 2022; Riffe et al., 1993). We collected all relevant text data using the dictionary-based sampling procedure, resulting in 4,022,616 documents. We sampled English-language data from tweets² using the Twitter API. 4chan submissions and comments posted to the board /pol/ (“Politically Incorrect”) were accessed via the 4plebs.org archive using the R package fouRplebsAPI (Buehling, 2022). The archiving of 4chan’s/pol/ board only started in 2014, resulting in a truncated dataset on this platform. Articles from alternative media (N=1,406) were collected from the websites Breitbart, Daily Caller, The Blaze, The Gateway Pundit, Infowars, Talking Points Memo, and Occupy Democrats via web scraping. This selection represents a spectrum of alternative news media sites from far-right to the left. For further analysis, all documents were transformed into lowercase. Stop words, Twitter user handles, and URLs were removed.

Even though the comprehensive key-term dictionary is theory-based, computationally augmented, and validated in several steps, the results of keyword-based data collection are prone to type I error and noise (Mahl et al., 2022b). Therefore, an additional classifier was trained to ensure the most accurate identification of conspiracy-related documents in the corpus. On the basis of a training dataset of 2,167 manually annotated documents, we fine-tuned (see Appendix B) the pretrained language model DistilBERT over 100 trials (mean F1: 0.85; accuracy: 0.83). We applied the final model with the optimal set of hyper-parameters (F1: 0.84; accuracy: 0.83) to our dataset. The resulting dataset consisted of 105,507 4chan posts, 771 alternative media articles, and 1,706,057 Twitter tweets, which can confidently be assumed to be conspiracy-related communication.

Operationalization

To operationalize *communication styles* and answer RQ 1, we measured the varying use of lexical and stylistic features between and within the three platform corpora.

To assess the use, distribution, and longitudinal evolution of the lexical items, we analyzed the occurrence of distinct n -grams in our dataset and compared them between 4chan, Twitter, and alternative news media. To identify the most relevant and informative terms used in conspiracy-related communication, each term was weighed via its tf-idf score and ranked accordingly to extract the most relevant terms of conspiracy-related communication for each platform (see Appendix C for details of the calculation). In this procedure, the informativeness of a term is operationalized by its frequency, weighed according to the number of appearances in the corpus documents. Thus, terms used in each document are assigned a low weight despite their high frequency because they are not as informative given the overall corpus (Manning et al., 2009). To calculate tf-idf weights for each platform and random week sample, tf-idf scores are calculated on the document level of the respective sub-corpus. The overall tf-idf score for each term is the sum of all document-level tf-idf scores in the corpus. Theoretically, the most relevant terms can be the same on each platform. The 100 highest-ranking terms were qualitatively evaluated and compared to analyze their explicitness.

We analyzed the similarity of high-relevance terms between platforms as a quantitative measure. To do so, we calculated the pairwise Jaccard similarity between the most relevant word sets, measuring the lexical overlap between the two vocabularies.

For style elements, we measured emotionality, linguistic complexity, and certainty using LIWC22, adopting the operationalization of Visentin et al. (2021)³. LIWC is a dictionary-based tool that classifies over 12,000 words into psycholinguistic categories. The category score is calculated as the proportion of words in a given text that are captured in each category within its dictionary (Boyd et al., 2022). Furthermore, the toxicity of each document was classified using Perspective API (Jigsaw, 2022).

We relied on a manual content analysis based on a standardized codebook to capture the functions of conspiracy-related content and answer RQ 2. In coding, conspiracy-related texts were differentiated by their stance and type of narration. Neutral content describes or mentions a conspiracy theory without a positive or negative evaluation. Furthermore, posts uncritically narrating, reproducing, or propagating a conspiracy theory are differentiated from posts rejecting or debunking it. Two trained student assistants coded a random sample of posts from all sample weeks. In cases of intercoder disagreement, the document was discussed by both coders to arrive at a unanimous decision (see Appendix D for the complete codebook).

Results

Owing to the diverging user activity, the sample size was distributed unequally across platforms, with the largest collected from Twitter, followed by 4chan and alternative media sources using any of the keywords. While we cannot draw inferences from the absolute number of texts per platform, the longitudinal distribution of conspiracy-related communication can be comparatively assessed across platforms (Figure 1). The sample size increased over time on all platforms. The increase in conspiracy-related discourse peaked in 2020 on 4chan and Twitter, and reached its maximum in alternative media in 2018.

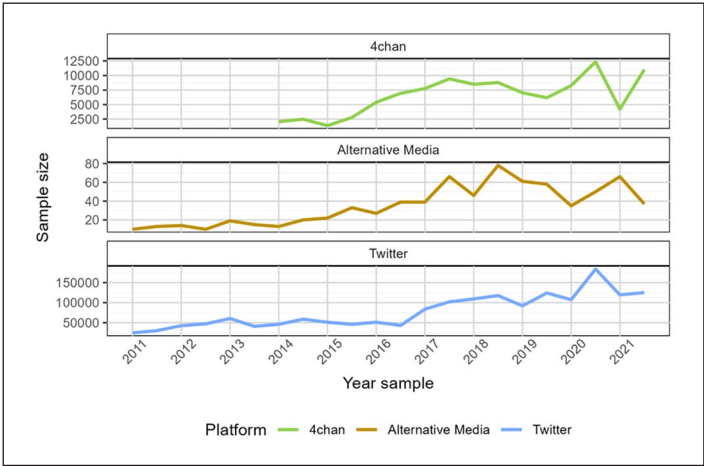


Figure 1. Sample size per platform after BERT classification.

Lexical analysis

Weighting all terms in the platform-specific corpora via tf-idf results in a ranking of the most used and informative terms. Apart from the overall explicit language, especially on 4chan, the most distinctive words on 4chan were racist and antisemitic slurs (see Appendix E for the top 20 terms per corpus). References to contemporary politics (“Trump” or “country”) were rare. The conspiracy-related communication on Twitter also referred to antisemitic conspiracy theories but did so by putting higher emphasis on more elaborate conspiracy theories (“puppet” and “NWO”) and the use of dog-whistles to veil and motivate their narrations (“globalists,” “Soros,” and “Gates”). Alternative news media articles containing keywords from the search-term dictionary mainly referenced US politicians (“Trump”) and institutions (“President,” “government,” and “FBI”). Thus, they are more closely aligned with current political events rather than abstract conspiracy narrations.

The results of the quantitative lexical analysis of the most informative terms per platform are shown in Figure 2, which depicts various dimensions of the pairwise platform comparisons. Overall, the cross-sectional term similarity was highest between Twitter and 4chan, with a Jaccard similarity of 0.28 for the 100 highest-ranked terms. Including only the top-most informative words might result in measurements that are more sensitive to minor changes in word use. Therefore, all similarity measures in the lexical analyses were carried out using the 500 most informative terms in each platform. This revealed a higher similarity of 0.48. The overall lexical similarity between Twitter and alternative media (top 100 terms: $J=0.183$, top 500 terms: $J=0.30$) was expectedly higher than that between 4chan and alternative media (top 100 terms: $J=0.11$, top 500 terms: $J=0.25$).

Figure 2 (lower panel) depicts a longitudinal analysis of the similarities between the most informative terms for each observation period between the platforms. The word choice on 4chan and Twitter became more similar over time concerning both the 100 and

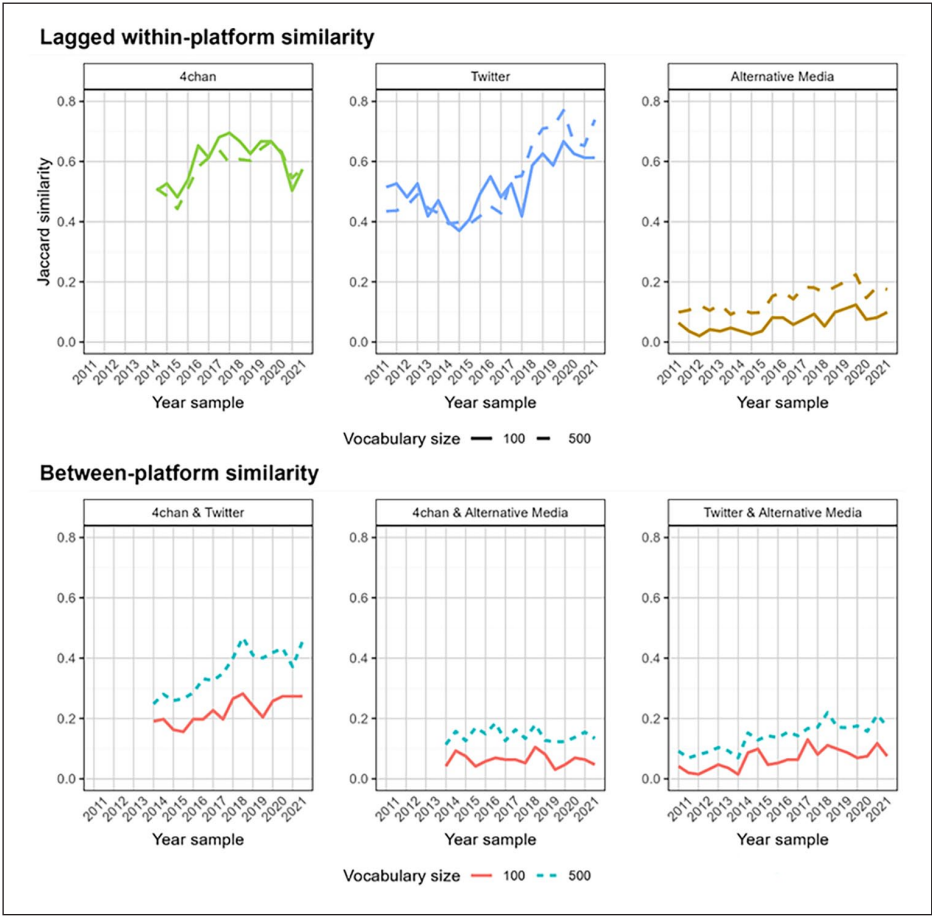


Figure 2. Lexical similarity between platforms. **Upper panel:** Lagged within-platform similarity of top terms. **Lower panel:** Longitudinal development of top term similarity between platforms.

500 most important terms. Starting from a Jaccard similarity of 0.25 in 2014, the overlap of the top 500 terms nearly doubled to 0.46 in 2021. The similarity between 4chan and alternative media remained constantly low, approximately 0.13 (0.06 for the 100 most important terms). The similarity between alternative media and Twitter showed equally low levels at the beginning of the observation period, but the similarity in word choice in conspiracy-related communication doubled until the end of the observation period.

Within each platform over time, the similarity of the most important words in each period was compared with the most important words in the previous period, indicating a time lag of around six months. This lagged within-platform term similarity increased on 4chan between 2015 and 2019 and decreased afterward. On Twitter, it is roughly on the

same level in the early observation periods but then decreased, only to increase again from 2015. Both platforms showed a moderately high inertia of change in term use and a higher similarity within platforms than with other platforms. The within-platform Jaccard similarity of alternative media over time was lower than that on other platforms, signifying a choice of words more oriented toward current events than on the continuous reproduction of stable narratives.

While the more abstract measure of term similarity across platforms showed word-choice convergence, zooming in on the specific terms generating this pattern revealed the nature of this similarity. The overlapping terms between 4chan and Twitter consisted of words used to embellish the conspiracy narrations studied, including “globalism,” “Soros,” and “false flag,” which are markers of the actual conspiracy theories and anti-semitic dog whistles. More overt instances of conspiracist reasoning, such as the direct reference to Jewish or Black people, are exclusively expressed on 4chan. The increasing convergence in word choice between 4chan and Twitter was caused by the increased use of terms relating to current US domestic political events, such as “Trump,” “Biden,” and “vaccines,” in addition to the conspiracist terms. The overlap between Twitter and alternative media was also mainly driven by terms relating to real politics. At the beginning of the observation period, overlapping words referred partly to the conspiracy narration (“Soros” and “war on Whites”), general domestic US politics (“Romney” and “government”), and race-related discussion. They were later dominated by the COVID-19 pandemic, foreign politics (“Ukraine” and “China”), and domestic political polarization in the United States (“Mueller,” “Democrats,” and “president”). A complete list of the 100 most important terms per platform and week sample, and the pairwise overlap with each other platform can be found in the supplementary material provided via the Open Science Framework⁴.

To gain further insights into the use of specific terms, we calculated the most similar terms for each most distinctive term per platform. Similar to Peeters et al. (2021), we trained word2vec models for each yearly sub-corpus to detect similar terms on the basis of word vector representations (see Appendix F for details). We found that 4chan posts were centered around more explicit antisemitism, racism, and conspiracy theories (i.e. Jewish people were discussed in terms similar to those used for “Bolsheviks,” “Zionists,” and “Freemasons”; Illuminati similar to “NWO,” “Satanism,” and “globalism”; and immigration similar to “mass immigration” and Great Replacement vocabulary). On Twitter, the use of Illuminati was similar to that of other conspiracy-related terms such as “Freemasonry” and “secret societies,” but included pop culture references to music albums. Jewish people were referenced more in political and historical terms (“Holocaust,” “Palestinians,” and “Arabs”). Conspiracy-related tweets mainly referenced immigration similar to right-wing populist rhetoric (“mass immigration,” “uncontrolled,” and “multiculturalism”). While alternative news media partly engaged in overt conspiracism (in 2017 and 2018, Jewish people were discussed similarly to George Soros and “puppet”), conspiracy-related posts generally relate to terms based in foreign (Israel-related) or domestic (illegal,” “immigrants,” and “border”) politics. The term “Illuminati” was not mentioned in any conspiracy-related alternative news media article.

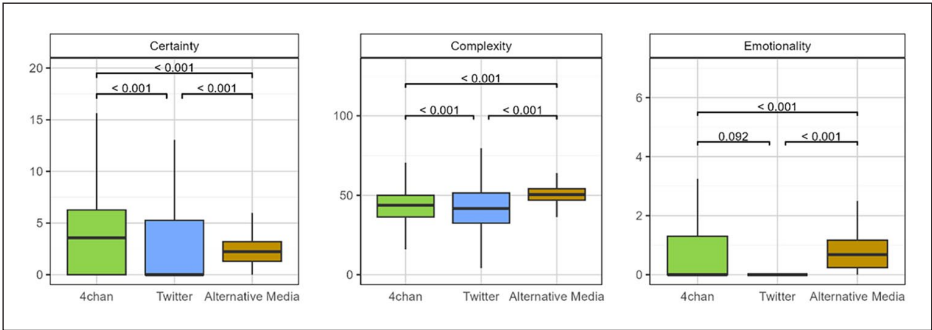


Figure 3. Linguistic styles between platforms; cross-sectional distribution of LIWC scores. Labels denote the p-value of t-test distribution test results.

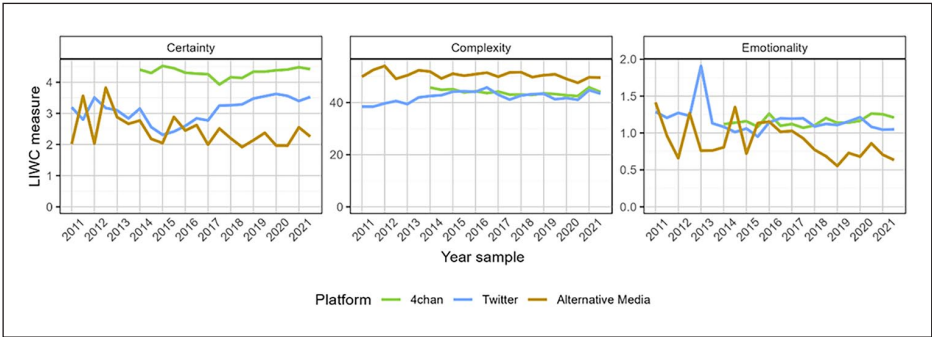


Figure 4. Longitudinal development of the LIWC result means.

Style Elements

LIWC classification was performed on the individual texts in the corpus to analyze the prevalence of language styles. As described in the methods section, we distinguished between posts containing certain complex and emotional languages, and their toxicity. Figure 3 shows boxplots of the cross-sectional LIWC classification distributions of the styles across the platforms, while Figure 4 shows the longitudinal development of the means of the style markers.

The analysis of language style on 4chan revealed that posts on this platform were the most certain overall, consistently holding in every sampled period, supporting the expectation stated in Section 2. The complexity of conspiracy-related communication on 4chan is comparable with that of other platforms, but the mean complexity score in each period of measurement was significantly lower than that reported in alternative media news articles but higher than that on Twitter. 4chan’s emotionality was low in the median, but its cross-sectional distribution was heavily skewed.

The median tweet certainty was the lowest among the three digital communication venues. However, it is highly skewed, and the analysis of the over-time development of

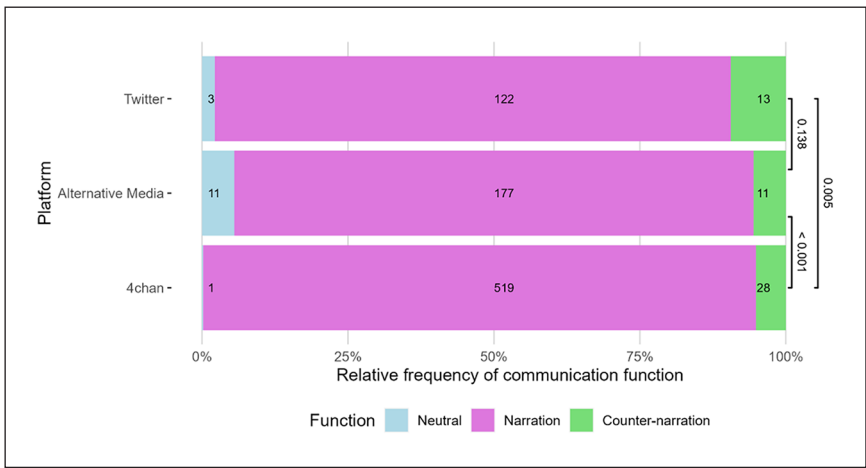


Figure 5. Relative frequency of conspiracy-related communication functions in the random sample coded. The absolute number of documents in each category is shown inside the bars. Labels denote the p-value from Fisher’s exact test of distribution independence.

mean certainty scores revealed that while decreasing until 2015, the certainty expressed in conspiracy-related tweets started to increase again afterward. As expected, Twitter’s complexity was lowest in the cross-sectional analysis and over time. The median emotionality on Twitter was slightly higher than that on 4chan, but not as heavily skewed. The distribution test revealed no significant difference between Twitter’s and 4chan’s distribution of emotionality, contradicting the expectations of a higher emotionality caused by the collision of different political leanings on the platform.

The certainty score of alternative media articles was lowest on average and over time while having a higher median value than tweets. The toxicity scores of all online venues are presented in Appendix G. The scores met the prior expectations, as the most toxic language can be found on 4chan/pol/, followed by Twitter. Alternative media toxicity was the lowest of all platforms observed.

Communication functions

Figure 5 shows the distribution of the functions of conspiracy-related communication across platforms. By comparing the functions, 4chan’s conspiracy-related posts nearly exclusively narrated or propagated conspiracy theories. Only 5% of the posts refuted or counter-narrated conspiracy theories, and one post neutrally described them. Contrary to our expectations, most documents on Twitter and alternative media did not neutrally report or debunk conspiracy theories. They rather function as a conspiracy theory narration in most instances. As the sample of alternative news media articles was drawn from a diverse range of media outlets, the communication functions specific to each outlet are reported in Appendix H. The results show that 80% of the left-leaning sample were coded as neutral or counter-narration, whereas only 7% of the right-leaning

alternative media articles did not engage in conspiracy theory narration. To determine if the distribution of communication functions is statistically different across platforms, we used Fisher's exact test (Fisher, 1970) to test whether the communication function distributions were artifacts of random sampling or if they are indeed drawn from underlying distributions. The test results show that the documents from Twitter and alternative media were significantly different from those from 4chan regarding their function of conspiracy-related communication. Twitter and alternative media do not differ significantly.

Discussion

This study's quantitative and qualitative analyses aimed at characterizing how conspiracy-related communication is performed in online communication venues that vastly differ regarding their technical infrastructure, platform policies, content producers, and overall user culture. The different style dimensions must be analyzed individually and comparatively to determine the overall development of communication styles in each online venue.

The mostly unregulated, ephemeral, and fringe social media platform 4chan hosted the most toxic and certain conspiracy-related content. Posts on 4chan were more emotional than those on alternative media but insignificantly different from those on Twitter. Conspiracy-related posts are mainly used in a narrative function, with only a few discussions employing a rebuttal to the propagation of conspiracy theory. Analyzing the word choice on the platform revealed the prominence of racist slurs and more coded conspiracy theory elements. Furthermore, a convergence of word choices between 4chan and Twitter users became clear.

Supporting our assumptions, tweets that contained conspiracy-related content were located between the two other online venues analyzed regarding toxicity, certainty, and emotionality. Furthermore, counter-narration frequency does not differ from alternative media statistically significantly, even though the Twitter user base includes more fact-checkers, disagreeing political activists, and journalists. A likely explanation is that the user community that discussed conspiracy theories on Twitter did not overlap with potentially counter-narrating users as much as expected. Therefore, arguments against conspiracy theories are uncommon. Tweets had the lowest linguistic complexity across all platforms, with minimal difference from that of 4chan posts. This permanent similarity indicates that Twitter's character limit of 140 (280 from 2017 onward) for tweets does not seem to be the driver of this subpar complexity. Twitter's increased toxicity and certainty after 2015 were synchronous with the shift in word choice after Donald Trump's inauguration as US president and COVID-19 pandemic outbreak. Furthermore, while texts published on 4chan or alternative media were only regulated by content moderation policies to a minor extent, Twitter experienced several waves of content moderation efforts, introducing dedicated larger-scale, algorithmically augmented moderation staff after public criticism resulting from suspected manipulation campaigns in the wake of Donald Trump's election (Fielitz and Jaspert, 2023). Thus, the increasing toxicity on Twitter in this study is likely the lower bound of the actual user culture prevalent on the platform.

The overarching results show the relative stability of user cultures within and the fundamental differences between online platforms. This insight supports the claims of Avalor et al. (2024) that toxicity patterns in online conversations remain relatively stable within social media. Furthermore, on a larger scale, our results support de Keulenaar's (2023) argument based on a single-example analysis that the minimal, eclectic content moderation, ephemeral structure, and positioning of 4chan on the fringes of online social media provide "affordances of extreme speech." With the increasing popularity of a platform (Twitter) or the claim to journalistic professionalism (alternative media) and the subsequent aim to appeal to larger "mainstream" audiences, the boundaries of what can be openly expressed as conspiracy theories become increasingly narrow. Furthermore, the observed cross-platform convergence of word choice in conspiracy-related communication is driven by discussion and reporting about politicians, institutions, and events. While cross-platform causal effects of style changes are not tested in this study, the intermedia agenda-setting literature points to an enduring influence of news media on social media communication, whereas social media sets the agenda during pivotal events (Su and Borah, 2019). This tendency can also be linked to an increasing interpretation of politics through a conspiratorial lens. Politicians hinting toward conspiracy theories to attack opponents and public rebuttals of the attacked trigger general public discussions and reporting about conspiracy theories and, consequently, the prevalence of conspiracy-related communication.

When quantitatively measuring language styles, researchers should acknowledge clear methodological limitations. The style classification software used in this analysis, LIWC and Perspective API, were developed for different application domains and did not consider text context. Highly toxic comments on 4chan will, at most, appall non-in-group platform users. By contrast, regular 4chan users are unlikely to disengage from the discussion due to toxic language. In the context of this study, the results indicate an "outsider's" perspective on 4chan's content. All LIWC results were based on the developers' proprietary dictionaries, which are generated and updated regularly. While transparent and reproducible, dictionary methods lack the capacity to incorporate context and are inflexible regarding internet slang. However, the costs of training reliable, high-performance classification models for all psychometric measures of interest in this study must be balanced with the possible performance gains. Hence, the authors consciously decided in favor of LIWC.

Conclusion

The aims of this study and the results presented are consistent with the recent scholarly contributions to the debate about extreme speech, platform affordances, and mainstreaming of far-right ideology. Far-right conspiracy theories are based on anti-semitic and racist premises and can be harmful to public discourse and minoritized individuals who object to these narrations. However, they do not need to be propagated via hate speech. Consequently, conspiracy-related communication can be found on all platforms, including unmoderated and anonymous (4chan), moderated and popular (Twitter), and alternative news media seeking to appear as professional journalistic venues. The platforms mainly differ in *how* conspiracy theory is presented,

likely linked to the varying “affordances of extreme speech” (de Keulenaar, 2023) that the respective platforms provide for their content producers.

The implications of our findings are two-fold. First, platform affordances, in all likelihood, affect the possible selection or self-selection of active platform users. Consequently, our results show that users are adaptive in how they communicate and that these varying communication styles are structured along the borders of the platforms. Second, individual platform policies do not appear to reduce the absolute number of far-right conspiracy-related communications under study. Instead, they result in a nuanced topology of the overall digital communication landscape. While tackling this development might be out of scope for the individual platform operator, the emerging platform ecosystem provides niches for more and less overt anti-democratic content to which followers can be funneled (Zimdars, 2024) or that can be used as a source of mobilization and information (Buehling and Heft, 2023). Across all platforms, conspiracy theories were increasingly used as an explanatory context of current political events, underlining their potential for mainstreaming far-right and anti-democratic ideology in public discourse. This dynamic has the potential to subtly normalize extremist narratives by reducing public resistance or increasing their acceptability (Rothut et al., 2024), thereby expanding the influence and integration of extremist actors into broader society.

Acknowledgements

This research could only be realized by the exceptional and persevering work of our coders, Joana Becker, Angelika Juhász, Katharina Sawade, and Dominik Hokamp. They provided high-quality work while reading extreme and toxic social media posts, for which we are very thankful. The authors would also like to thank the HPC Service of FUB-IT, Freie Universität Berlin, for computing time. Additional thanks go to the anonymous reviewers, whose insightful comments improved our manuscript.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by grants from the German Federal Ministry of Education and Research (grant nos. 13N16049 [in the context of the call for proposals Civil Security – Societies in Transition] and 16DII135 [in the context of the Weizenbaum Institute]).

ORCID iDs

Kilian Buehling  <https://orcid.org/0000-0002-5244-7547>

Xixuan Zhang  <https://orcid.org/0000-0002-4636-3881>

Annett Heft  <https://orcid.org/0000-0001-6637-795X>

Supplemental material

Supplemental material for this article is available online.

Notes

1. While not every reference to the Illuminati/NWO conspiracy theory is necessarily intended to cover explicit antisemitism, the conspiracy theory itself is antisemitic. Its roots are

- tracible from medieval anti-Judaism to the French revolution and 20th century antisemitism (Simonsen, 2020), up to the 21st century conspiracy narrations (Berlet, 2011). Thus, when individuals propagate systemic NWO conspiracy theories—“a task that requires them to recognize the relevance of past conspiracies and past conspiracy theories – they invariably come into contact with the antisemitic legacy of the conspiracy culture” (Byford, 2011: 102).
2. The Twitter data were collected in two rounds according to the seed and expanded dictionaries. We collected all tweets in the first round. Owing to the changes in Twitter’s API policy, we sampled only the tweets through two API calls for the second round of data collection.
 3. The exact LIWC22 categories used for each dimension are detailed in Appendix G.
 4. Online supplement via OSF: <https://osf.io/a79bx/>

References

- Abidin C (2021) From “Networked publics” to “Refracted publics”: a companion framework for researching “Below the radar” studies. *Social Media + Society* 7(1): 2056305120984458.
- Avalle M, Di Marco N, Etta G, et al. (2024) Persistent interaction patterns across social media platforms and over time. *Nature* 628: 582–589.
- 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.
- Barkun M (2013) A Culture of conspiracy: apocalyptic visions in contemporary America. In: Juergensmeyer M, Embree AT, Tibi B, et al. (eds) *Comparative Studies in Religion and Society*. 15th ed. Berkeley, CA: University of California Press, pp. 255.
- Benkler Y, Faris R and Roberts H (2018) *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford: Oxford University Press.
- Berlet C (2011) Protocols to the left, protocols to the right: conspiracism in American political discourse at the turn of the second millennium. In: Landes R (ed.) *The Paranoid Apocalypse: A Hundred-year Retrospective on the Protocols of the Elders of Zion*. New York: New York University Press, pp. 186–216.
- Blankenship KL and Craig TY (2011) Language use and persuasion: multiple roles for linguistic styles. *Social and Personality Psychology Compass* 5(4): 194–205.
- Boberg S, Quandt T, Schatto-Eckrodt T, et al. (2020) Pandemic populism: Facebook pages of alternative news media and the corona crisis—a computational content analysis. arXiv:2004.02566.
- Bossetta M (2019) *The Digital Architectures of Social Media: Platforms and Participation in Contemporary Politics*. København: University of Copenhagen.
- Boyd RL, Ashokkumar A, Seraj S, et al. (2022) *The Development and Psychometric Properties of LIWC-22*. Austin, TX: University of Texas at Austin.
- Bruns A, Hurcombe E and Harrington S (2022) Covering conspiracy: approaches to reporting the COVID/5G conspiracy theory. *Digital Journalism* 10(6): 930–951.
- Bucholtz M and Hall K (2005) Identity and interaction: a sociocultural linguistic approach. *Discourse Studies* 7(4–5): 585–614.
- Buehling K (2022) Fourplebsapi: R package for accessing 4Chan posts via the 4Plebs.org API (Version 0.9.0)[R]. Available at: <https://doi.org/10.5281/zenodo.6637440>
- Buehling K and Heft A (2023) Pandemic protesters on Telegram: how platform affordances and information ecosystems shape digital counterpublics. *Social Media + Society* 9(3): 20563051231199430.
- Butter M (2018) *Nichts Ist, Wie Es Scheint «: Über Verschwörungstheorien*. Frankfurt: Suhrkamp Verlag.
- Byford J (2011) *Conspiracy Theories: A Critical Introduction*. New York: Springer.

- Cosentino G (2020) From pizzagate to the great replacement: the globalization of conspiracy theories. In: Cosentino G (ed.) *Social Media and the Post-truth World Order: The Global Dynamics of Disinformation*. New York: Springer, pp. 59–86.
- de Keulenaar E (2023) The affordances of extreme speech. *Big Data & Society* 10(2): 20539–517231206810.
- de Zeeuw D and Tuters M (2020) Teh internet is serious business: on the deep vernacular web and its discontents. *Cultural Politics* 16(2): 214–232.
- de Zeeuw D, Hagen S, Peeters S, et al. (2020) Tracing normification: a cross-platform analysis of the QAnon conspiracy theory. *First Monday* 25: 10643.
- Deem A (2019) The digital traces of #whitegenocide and alt-right affective economies of transgression. *International Journal of Communication* 13: 3183–3202.
- Ekman M (2019) Anti-immigrant sentiments and mobilization on the Internet. In: Smets K, Leurs K, Georgiou M, et al. (eds) *Sage Handbook of Media and Migration*. New York: Sage, pp. 551–562.
- Ekman M (2022) The great replacement: strategic mainstreaming of far-right conspiracy claims. *Convergence* 28(4): 1127–1143.
- Engesser S, Ernst N, Esser F, et al. (2017) Populism and social media: how politicians spread a fragmented ideology. *Information, Communication & Society* 20(8): 1109–1126.
- European Commission (2018) *Action Plan against Disinformation, JOIN/2018/36 Final*. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018JC0036&qid=1715599259096>
- Fielitz M and Jaspert M (2023) *Paradigmen Der Plattformpolitik Im Umgang Mit Hass Und Extremismus: Eine Chronologische Einordnung*. Amadeu Antonio Stiftung. Available at: <https://doi.org/10.19222/202314/08>
- Fisher RA (1970) Statistical methods for research workers. In: Kotz A and Johnson NL (eds) *Breakthroughs in Statistics: Methodology and Distribution*. New York: Springer, pp. 66–70.
- Forberg PL (2022) From the fringe to the fore: an algorithmic ethnography of the far-right conspiracy theory group QAnon. *Journal of Contemporary Ethnography* 51(3): 291–317.
- Frischlich L, Schatto-Eckrodt T and Völker J (2022) *Rückzug in Die Schatten? Die Verlagerung Digitaler Foren Zwischen Fringe Communities Und "Dark Social" Und Ihre Implikationen Für Die Extremismusprävention*. CoRE – Connecting Research on Extremism in North Rhine-Westphalia.
- Gibbs M, Meese J, Arnold M, et al. (2015) #Funeral and Instagram: death, social media, and platform vernacular. *Information, Communication & Society* 18(3): 255–268.
- Girard P (2020) Conspiracy theories in Europe during the twentieth century. In: Butter M and Knight P (eds) *Routledge Handbook of Conspiracy Theories*, pp. 569–581. London: Routledge.
- Gründl J (2022) Populist ideas on social media: a dictionary-based measurement of populist communication. *new media & society* 24(6): 1481–1499.
- Hagen S and de Zeeuw D (2023) Based and confused: tracing the political connotations of a memetic phrase across the web. *Big Data & Society* 10(1): 20539517231163175.
- Heft A and Buehling K (2022) Measuring the diffusion of conspiracy theories in digital information ecologies. *Convergence: The International Journal of Research into New Media Technologies* 28(4): 940–961.
- Heft A, Buehling K, Zhang X, et al. (2023) Challenges of and approaches to data collection across platforms and time: conspiracy-related digital traces as examples of political contention. *Journal of Information Technology & Politics* 21: 323–339.
- Heft A, Knüpfer CB, Reinhardt S, et al. (2021) Toward a transnational information ecology on the right? Hyperlink networking among right-wing news sites in Europe and the United States. *The International Journal of Press/politics* 26(2): 484–504.

- Heft A, Mayerhöffer E, Reinhardt S, et al. (2020) Beyond Breitbart: comparing right-wing digital news infrastructures in six western democracies. *Policy & Internet* 12(1): 20–45.
- Holt K, Ustad Figenschou T and Frischlich L (2019) Key dimensions of alternative news media. *Digital Journalism* 7(7): 860–869.
- Hoseini M, Melo P, Benevenuto F, et al. (2023) On the globalization of the QAnon conspiracy theory through telegram. *Proceedings of the 15th ACM web science conference 2023*: 75–85.
- Jigsaw (2022) *Perspective API – Attributes and Languages*. Available at: <https://developers.perspectiveapi.com/s/about-the-api-attributes-and-languages>
- Jolley D, Mari S and Douglas KM (2020) Consequences of conspiracy theories. In: Butter M and Knight P (eds) *Routledge Handbook of Conspiracy Theories*. London: Routledge, pp. 231–241.
- Jungherr A (2016) Four functions of digital tools in election campaigns: the German case. *The International Journal of Press/politics* 21(3): 358–377.
- Keeley BL (1999) Of conspiracy theories. *The Journal of Philosophy* 96(3): 109–126.
- Khalid O and Srinivasan P (2020) Style matters! Investigating linguistic style in online communities. *Proceedings of the International AAAI Conference on Web and Social Media* 14: 360–369.
- Kim H, Jang SM, Kim S-H, et al. (2018) Evaluating sampling methods for content analysis of Twitter data. *Social Media + Society* 4(2): 2056305118772836.
- Klein E (2023) *Loaded Language and Conspiracy Theorizing* [Ph.D.]. Available at: <https://www.proquest.com/docview/2821587281/abstract/E1294929A9F0483DPQ/1>
- Leal H (2020) Networked disinformation and the lifecycle of online conspiracy theories. In: Butter M and Knight P (eds) *Routledge Handbook of Conspiracy Theories*. London: Routledge, pp. 497–511.
- Mahl D, Schäfer MS and Zeng J (2022a) Conspiracy theories in online environments: an interdisciplinary literature review and agenda for future research. *new media & society* 25: 1781–1801.
- Mahl D, von Nordheim G and Guenther L (2022b) Noise pollution: a multi-step approach to assessing the consequences of (Not) validating search terms on automated content analyses. *Digital Journalism* 11: 298–320.
- Mahl D, Zeng J and Schäfer MS (2021) From “Nasa lies” to “Reptilian Eyes”: mapping communication about 10 conspiracy theories, their communities, and main propagators on Twitter. *Social Media + Society* 7(2): 17482.
- Mahl D, Zeng J and Schäfer MS (2023) Conceptualizing platformed conspiracism: analytical framework and empirical case study of BitChute and Gab. *new media & society* 26: 6938–6957.
- Manning C, Raghavan P and Schuetze H (2009) *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.
- Mayerhöffer E and Heft A (2022) Between journalistic and movement logic: disentangling referencing practices of right-wing alternative online news media. *Digital Journalism* 10(8): 1409–1430.
- Moran RE, Grasso I and Koltai K (2022) Folk theories of avoiding content moderation: how vaccine-opposed influencers amplify vaccine opposition on Instagram. *Social Media + Society* 8(4): 20563051221144252.
- Naab TK and Küchler C (2022) How to select from the stream of comments? Sampling user comments from Facebook pages of news outlets for manual content analysis. In: Jünger J, Gochermann U, Peter C, et al. (eds) *Methoden Und Forschungslogik Der Kommunikationswissenschaft*, vol. 17, pp. 196–229. Köln: Herbert von Halem Verlag.
- Newman N, Fletcher R, Schulz A, et al. (2023) *Reuters Institute Digital News Report 2023* [SSRN Scholarly Paper]. Available at: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2023-06/Digital_News_Report_2023.pdf

- Nissenbaum A and Shifman L (2017) Internet memes as contested cultural capital: the case of 4chan's /b/ board. *new media & society* 19(4): 483–501.
- Nordheim GV and Kleinen-von Königslöw K (2021) Uninvited dinner guests: a theoretical perspective on the antagonists of journalism based on serres' parasite. *Media and Communication* 9(1): 88–98.
- Nunziato DC (2023) *The Digital Services Act and the Brussels Effect on Platform Content Moderation* (SSRN Scholarly Paper 4425793). Available at: <https://papers.ssrn.com/abstract=4425793>
- Oliver JE and Wood TJ (2014) Conspiracy theories and the paranoid Style(s) of mass opinion. *American Journal of Political Science* 58(4): 952–966.
- Peeters S, Tuters M, Willaert T, et al. (2021) On the vernacular language games of an antagonistic online subculture. *Frontiers in Big Data* 4. Available at: <https://www.frontiersin.org/articles/10.3389/fdata.2021.718368>
- Phillips W (2018) The oxygen of amplification. *Data & Society* 22: 1–128.
- Pohjonen M and Udupa S (2017) Extreme speech online: an anthropological critique of hate speech debates. *International Journal of Communication* 11: 1173–1191.
- Puschmann C, Karakurt H, Amlinger C, et al. (2022) RPC-Lex: a dictionary to measure German right-wing populist conspiracy discourse online. *Convergence: The International Journal of Research into New Media Technologies* 28(4): 1144–1171.
- Quandt T (2018) Dark participation. *Media and Communication* 6(4): 36–48.
- Riedl MJ, Joseff K, Soorholtz S, et al. (2024) Platformed antisemitism on Twitter: anti-Jewish rhetoric in political discourse surrounding the 2018 US midterm election. *new media & society* 26(4): 2213–2233.
- Riffe D, Aust CF and Lacy SR (1993) The effectiveness of random, consecutive day and constructed week sampling in newspaper content analysis. *Journalism Quarterly* 70(1): 133–139.
- Rogers R (2017) Foundations of digital methods: query design. In Schäfer MT and van Es K (eds) *The Datafied Society. Studying Culture through Data*, pp. 75–94. Amsterdam University Press.
- Rothut S, Schulze H, Rieger D, et al. (2024) Mainstreaming as a meta-process: a systematic review and conceptual model of factors contributing to the mainstreaming of radical and extremist positions. *Communication Theory* 34(2): 49–59.
- Schatto-Eckrodt T, Clever L and Frischlich L (2024) The seed of doubt: examining the role of alternative social and news media for the birth of a conspiracy theory. *Social Science Computer Review* 42(5): 1160–1180.
- Schindler DJ, Jha S, Zhang X, et al. (2024) LGDE: local graph-based dictionary expansion (arXiv:2405.07764). arXiv. DOI: 10.48550/arXiv.2405.07764.
- Schwaiger L (2022) *Gegen Die Öffentlichkeit: Alternative Nachrichtenmedien Im Deutschsprachigen Raum*. transcript.
- Simonsen KB (2020) Antisemitism and conspiracism. In: Butter M and Knight P (eds) *Routledge Handbook of Conspiracy Theories*, pp. 357–370. London: Routledge.
- Smallpage SM, Drochon H, Uscinski JE, et al. (2020) Who are the Conspiracy theorists? Demographics and conspiracy theories. In: Butter M and Knight P (eds) *Routledge Handbook of Conspiracy Theories*, pp. 263–277. London: Routledge.
- Su Y and Borah P (2019) Who is the agenda setter? Examining the intermedia agenda-setting effect between Twitter and newspapers. *Journal of Information Technology & Politics* 16(3): 236–249.
- Suler J (2004) The online disinhibition effect. *Cyberpsychology & Behavior* 7(3): 321–326.
- Tuters M, Jokubauskaitė E and Bach D (2018) Post-truth protest: how 4chan cooked up the piz-zagate bullshit. *M/C Journal* 21(3): Article3.
- Uscinski J, Enders A, Klofstad C, et al. (2022) Have beliefs in conspiracy theories increased over time? *PLOS ONE* 17(7): e0270429.

- Visentin M, Tuan A and Di Domenico G (2021) Words matter: how privacy concerns and conspiracy theories spread on twitter. *Psychology & Marketing* 38(10): 1828–1846.
- Yang Y (2020) *The Political Logic of the Radical Right Media Sphere in the United States*. University of Washington. Available at: <https://www.proquest.com/docview/2437414567>
- Zimdars M (2024) Alt-health influencers and the threat of social media deplatforming. *Journal of the Association for Information Science and Technology* 75: 1216–1229.
- Zollo F, Novak PK, Vicario MD, et al. (2015) Emotional dynamics in the age of misinformation. *PLoS ONE* 10(9): e0138740.

Author biographies

Kilian Buehling is a postdoctoral research associate at the Freie Universität Berlin and the Weizenbaum Institute for the Networked Society. He studied Economics at Technische Universität Dresden, and his research interests are information diffusion, transnational communication processes, and network dynamics of anti-democratic and conspiracy theory groups.

Xixuan Zhang is a research associate and doctoral candidate at the Freie Universität Berlin. She studied media and political communication at Freie Universität Berlin and media informatics at Technische Universität Berlin. Her research interests are in the fields of digital activism, online discourse, and the networked public sphere. She also researches the application of computational methods, ranging from text mining to network analysis and machine learning.

Annett Heft is a Professor of far-right extremism research with a focus on media and public spheres at the University of Tübingen and head of the research group Dynamics of Digital Mobilization at the Weizenbaum Institute and the FU Berlin. Her main fields of research are the comparative study of political communication and mobilization in Europe with a focus on digital public spheres, right-wing communication infrastructures, transnational communication, as well as quantitative methods, and computational social science.