Rage Against the Machine: Exploring Violence and Emotion in Conspiracy Narratives on Parler

Darja Wischerath (iD)
School of Management, University of Bath, Bath, UK

Lukasz Piwek
School of Management, University of Bath, Bath, UK

Jonathan F. Roscoe
BT Applied Research, Ipswich, UK

Brittany I. Davidson
School of Management, University of Bath, Bath, UK

Abstract
The mainstreaming of conspiracy narratives has been associated with a rise in violent offline harms, from harassment, vandalism of communications infrastructure, assault, and in its most extreme form, terrorist attacks. Group-level emotions of anger, contempt, and disgust have been proposed as a pathway to legitimizing violence. Here, we examine expressions of anger, contempt, and disgust as well as violence, threat, hate, planning, grievance, and paranoia within various conspiracy narratives on Parler. We found significant differences between conspiracy narratives for all measures and narratives associated with higher levels of offline violence showing greater levels of expression.

Corresponding author:
Darja Wischerath, School of Management, University of Bath, Claverton Down, Bath, BA2 7AY, UK.
Email: dw942@bath.ac.uk

Introduction

Conspiracy theories are narratives that attempt "to explain the ultimate causes of significant social and political events and circumstances with claims of secret plots by two or more powerful actors" (Douglas et al., 2019). Although real conspiracies can happen (such as the Tuskegee syphilis study where black men in Tuskegee, Alabama, were deliberately not treated for syphilis [Heller, 2017]), they are rare. Generally, conspiracy theories lack plausible accounts of events given the available facts and are often unfalsifiable or believed despite available contrary evidence (Bartlett and Miller, 2010).

The mass proliferation and mainstreaming of conspiracy narratives in recent years pose a new threat to both individuals and society (Bruns et al., 2020). Whilst the majority of those who consume conspiracy narratives remain peaceful, there have been an increasing number of documented instances of violence. For example, acts range from attacks on vaccine clinic workers (Seidman, 2022) and 5G infrastructure (Parveen and Waterson, 2020) to homicides of family members believed to be part of the lizard race (Southern District of California, 2021), as well as large-scale events such as the 2021 Capitol Insurrection (*BBC News*, 2021) or terrorist attacks such as the Christchurch (Davey and Ebner, 2019) or the Utoya shooting (van Buuren, 2013).

Here, we explore the prevalence of anger, contempt, and disgust expressions of emotion in posts on the alt-tech social media platform Parler across multiple conspiracy narratives associated with a range of violent and nonviolent offline behaviors.

Conspiracy narratives in online communities

The participatory affordances of online forums form a unique place for collective sensemaking and community formation (Brown et al., 2022), allowing people to collectively question, share, and discuss complex information (Dailey and Starbird, 2015; Kou et al., 2017), with some evidence that social validation from online communities can aid participation in offline collective action (Smith et al., 2023).

Conspiracy communities are fundamentally social. In online forums and social media platforms, users can gather and share evidence for their chosen conspiracy, thereby collectively shaping conspiracy narratives through their discoveries and creating a perceived abundance of information to be consumed (Sunstein and Vermeule, 2009; Xiao et al., 2021; Zeng and Schäfer, 2021). These social networks can further foster a sense of social belonging that extends beyond the confines of conspiracy narratives (Xiao et al., 2021). Participation in these communities feeds users' innate curiosity, pushes them further down the rabbit hole, and, combined with a strong sense of social commitment, can keep users engaged in communities even when they experience frustrations over an answer that is never found (Sutton and Douglas, 2022; Xiao et al., 2021).

Platforms such as Gab, Parler, or Odysee form an alternative social media ecosystem of a multitude of communities from conspiracies to the far-right (Baele et al., 2023).

Alt-tech platforms tend to mimic affordances of mainstream platforms (e.g., YouTube or Reddit) while advertising minimal content moderation policies and free speech absolutism, attracting organizations and users banned from mainstream sites like Facebook and X (de Keulenaar, 2023; Rogers, 2020; Zeng and Schäfer, 2021). The combination of familiar affordances of social media platforms, explicit lack of content moderation, and a grievance-based identity of perceived oppression by "Big Tech censors" create ideal conditions for conspiracy narratives to thrive and turn to more extreme, and even violence-legitimating, ideas(Cinelli et al., 2022; Jasser et al., 2021; Nouri et al., 2021). For example, some work has found differences in online posting behaviors between violent and nonviolent extremists, particularly on alt-tech platforms such as StormFront (e.g., Scrivens et al., 2023).

Conspiracy narratives, emotion, and sensemaking

Conspiracy narratives provide easy, seemingly logical and causal, answers to questions about blameworthiness, intentionality, and factors that explain how a specific threat emerged (van Prooijen, 2011), making them ideal to aid in making sense of complex or distressing societal events (Franks et al., 2017). That is, blaming identifiable groups and institutions is more effective in reducing distress than admitting the role of uncontrollable factors and randomness as actions of agents can be understood and anticipated (Sullivan et al., 2010). For example, people may find the loss of a loved one from COVID-19 easier to process if they can blame the government, pharmaceutical companies, or the Asian community because "the virus originated in China," causing racial discrimination and violence, rather than face the randomness of such a grave event.

Appraisal theory posits that emotion and emotional appraisal of situations is a known route for people to cope with distressing events (Lazarus, 1991). The emotional response of the subjective situation appraisal leads to emotion expression—for example, someone may feel angry that their social life was impacted by COVID-19 lockdowns. This appraisal leads to action tendencies, such as fight-or-flight responses, which are unique to different emotions (Lazarus, 1991). For example, if someone perceives a situation as a threat, the emotional response might be fear, leading to actions such as avoidance or seeking help. Conversely, if the situation is appraised as unjust, it may elicit anger, resulting in actions like confrontation or advocacy.

Conspiracy narratives make heavy use of emotional appraisal of grievance-laden situations and as polysemic narratives, thus leaving room for individuals to decode the conspiracy narrative to fit with their preexisting beliefs and social identity (Harambam and Aupers, 2021). The uncertainty around the rapid spread of COVID-19 led some people to appraise the situation as threatening and, in an attempt to reduce uncertainty, believe conspiracy theories that 5G technology causes the virus, feeling fear and anxiety that drove them to destroy 5G towers near hospitals as a protective action against the perceived danger (Sweney and Waterson, 2020).

Through shared sensemaking, online communities can transform feelings of despair and confusion into a shared identity and purpose and even incite group members to engage in violent political action as a last resort (Törnberg and Törnberg, 2023). For example, protesters of the Anti-Extradition Law Amendment Bill Movement in Hong

Kong were found to be more supportive of violent action when they experienced higher emotions of anger, disgust, or fear, or identified more strongly with the group (Zhu et al., 2022). Here, we expand on this work by examining emotions of anger, contempt, and disgust as a pathway to legitimating discourse around violence in conspiracy narratives on Parler.

Emotions as a pathway to violence and collective action

Group-based emotions, and in particular perceived inequality and injustice, are powerful motivators of collective action (Iyer et al., 2007; van Zomeren et al., 2008). The ANCODI model proposes that historical narratives and reactions to events result in emotions of anger, contempt, and disgust that work together to motivate action, devalue the other group, and legitimize violence against outgroup members (Matsumoto et al., 2015). Anger, contempt, and disgust comprise three components of hate that contribute to aggression and violence (Sternberg, 2003). These emotions become integral to the group's narratives, thus providing guidelines for making appraisals about the outgroup and accelerating action (Matsumoto et al., 2015). For example, initial anger around mask mandates turned to narratives invoking disgust for those wearing masks and encouraging harassment of those wearing masks, even after mask mandates were lifted (Smith, 2022).

Anger, contempt, and disgust are each associated with distinct action tendencies. Anger is typically directed at a situational grievance that the ingroup experiences for which they blame the outgroup (Matsumoto et al., 2013), thus driving action against those deemed responsible (van Zomeren et al., 2004). Contempt and disgust, on the other hand, focus on the disposition of the outgroup members (Matsumoto and Hwang, 2012). Contempt deems the target as inferior and unworthy of esteem and is marked by exclusion tendencies that intend to punish the excluded. Disgust, on the other hand, deems the target to be "contaminated" (e.g., through comparisons to animals or implications of disease (Rozin et al., 2008), instilling a contamination avoidance response rooted in existential threat and excluding the target to the point of elimination (Miceli and Castelfranchi, 2018). Perhaps unsurprisingly, paranoia and anger have previously been identified as predictors of violence (Doyle and Dolan, 2006), and indeed, amongst those high in conspiracy belief, paranoia has been associated with increased justification of violence (Jolley and Paterson, 2020).

Hence, the ANCODI model poses a useful method of analyzing emotions associated with mobilization and has been used in several terrorist and extremist contexts, such as IS propaganda (Ingram, 2016), far-right shooter manifestos (Vanderwee and Droogan, 2023), and others (see Matsumoto et al., 2015) and is thus our lens for this work analyzing increasingly violent narratives.

Violence legitimating narratives

As a non-normative action, violence can only become a viable path of action when it is encouraged and legitimized by narratives in a network (Kruglanski et al., 2022). "Violent talk" (Simi and Windisch, 2020) reinforces the values of violence and its importance as a cultural and political practice, thus legitimizing violence as a viable and even desirable

pathway for addressing grievances. Grievance-based narratives and terrorism-justifying ideologies accomplish this by designating a specific outgroup as the culprit responsible for hardship, making violence against them appear justified (Webber and Kruglanski, 2017). This is also seen in pathways of radicalization models, which pose that legitimation of violence based on grievances is a key intermediary step between beliefs and violent action (McCauley and Moskalenko, 2008; Vegetti and Littvay, 2022).

Narratives can further legitimate violence by delegitimizing the outgroup, effectively excluding them from the group to which norms and values apply (Webber and Kruglanski, 2017). Dehumanization of the outgroup via comparisons to animals and implications of disease is a key strategy for delegitimization, as it degrades the outgroup. This can involve stripping members of their human characteristics or invoking a contamination avoidance response, for example, by labeling a group as "dirty," "diseased," or implying they are animals, thus making violence an acceptable course of action (Ebner et al., 2023; Miceli and Castelfranchi, 2018). When no outgroup is directly responsible for the grievance, uninvolved outgroups can be scapegoated and assigned blame to reduce uncertainty and increase perceptions of control for the ingroup (Glick, 2005; Webber and Kruglanski, 2017). For example, great replacement conspiracy narratives utilize grievances about socioeconomic hardship by blaming immigrants and the UN. Consequently, these narratives propose that eliminating the scapegoated outgroup will relieve the experienced hardship, for example, by freeing up jobs and reducing pressures on the housing market.

Social networks, such as those encountered on social media platforms, can further spread and endorse these narratives, leading to increased identification with extreme communities and reduced inhibitions against the use of violence (Simi and Windisch, 2020). "Violent talk" can help to socialize individuals into extreme communities by communicating values and norms, identifying and dehumanizing targets of violence, and glorifying perpetrators of violence (Simi and Windisch, 2020). Indeed, the 3N model of radicalization posits that social validation of narratives is key to the spread of violence legitimation and reducing individual inhibitions (Webber & Kruglanski, 2017).

However, "violent talk" is inherently indeterminate and may even serve as a substitute for violent behavior for some while increasing the likelihood of violence for others. Whilst it is, therefore, difficult to see the violent talk on online social media platforms as a definitive indicator of users' willingness to engage in violence, it highlights the group's attitudes and norms toward the use of violence by some of its members.

The present research: Emotion expression in violent and nonviolent conspiracy narratives

This paper aims to explore violence legitimation across multiple conspiracy narratives on Parler. Online communities can form around grievances or existing communities can identify a common grievance that can elicit anger. A narrative is then communally constructed to make sense of the situation, often identifying an outgroup to blame for the grievance. In some cases, this outgroup can then be delegitimized and dehumanized, evoking contempt and disgust, which can further be used to legitimize violence against
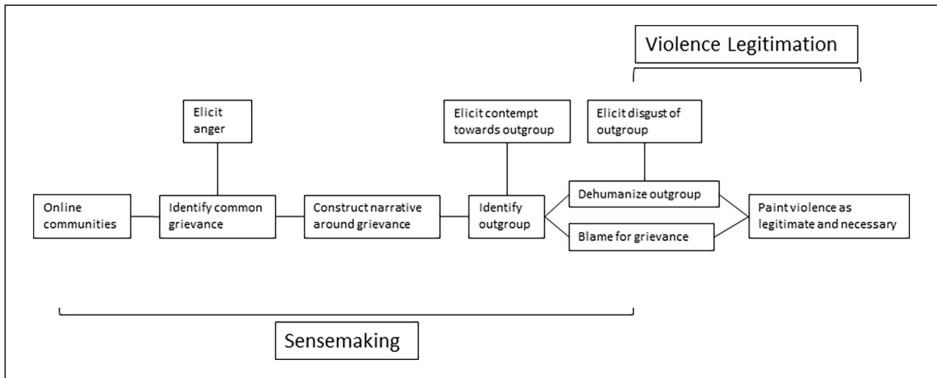
Figure 1. Framework showing narrative evolutions toward violence legitimation and associated emotions.

the outgroup. Whilst this may be an overly linear overview, Figure 1 shows a simplified, diagrammatic overview of our interests, highlighting the role of emotions in sensemaking and legitimation of violence in narratives.

Drawing on the proposed pathway of the ANCODI theory (Matsumoto et al., 2015), we explore the prevalence of anger, contempt, and disgust emotions in conspiracy narratives. Although this work is exploratory, we expect conspiracy narratives that are associated with more offline violence (e.g., antivaccine or great replacement narratives) to evoke more contempt and disgust emotions and legitimize violence more than narratives that are not associated with offline violence (e.g., flat earth narratives). Hence, our research questions are:

RQ1: Do expressions of anger, contempt, and disgust differ between conspiracy narratives?

RQ2: Do expressions of hate, grievance, threat, planning, paranoia, and violence differ between conspiracy narratives?

RQ3: Is the expression of anger, contempt, and disgust emotions correlated with expressions of violence, grievance, threat, planning, paranoia, and hate within conspiracy narratives?

Ethical approval has been granted by the University of Bath's ethics committee (ref: 3417-4230).

Materials and methods

We compare five conspiracy narratives prevalent in social media discourse: flat earth, anti-5G, false flag, anti-vaccine, and great replacement narratives. We chose these five narratives as they were popular in 2020, often in response to current events of that period,

Table 1. Selected conspiracy narratives, main beliefs, and associated outcomes.

Conspiracy narrative	Main belief	Outgroup	Nonviolent outcomes	Violent outcomes
Flat earth	The earth is not a sphere. Evidence is hidden by NASA and the global elites.	NASA, the global elite	Frequent offline conferences	None
Anti-5G	Microwaves radiated through 5G cause bodily harm and are used as population control.	Bill Gates, the government	Use of pseudoscientific, and harmful, gadgets to reduce radiation	Arson of 5G masts; physical harassment of engineers
False flag	Mass shootings are staged events by the government to increase gun control measures.	The government	Verbal harassment of bereaved families	Association with QAnon and US capitol insurrection
Antivaccine	The pharma industry intentionally hides the harms of vaccines and uses vaccines as population control.	Pharma industry, the government	Reduced uptake of vaccines	Harassment and attacks on vaccine clinic staff
The great replacement	Alleges that white populations are being replaced by non-white immigrants.	Immigrants, Jews, the government, the UN	Increased antisemitism and xenophobia	Mass shootings and terrorist attacks (e.g., Christchurch attack, Buffalo shooting)

such as the COVID-19 pandemic and social movements like the Black Lives Matter protests, and thus were likely to be encountered by an average user of Parler. Further, these narratives exhibit a range of violent and nonviolent outcomes associated with their respective communities and showcase a variety of beliefs and outgroups, which range from a general distrust of "the global elite" to identifying governmental bodies and ethnic minorities, whereas flat earth narratives have not been associated with any violent outcomes (see Table 1). To preserve ecological validity, we intentionally focused on broad narratives that may contain several subnarratives. For example, some antivaccine narratives address justifiable hesitation and governmental trust, whilst others spread ideas about microchips and extreme bodily harm that are less easily understood. Due to Parler's timeline affordances, we believe that users would encounter a mix of narratives for each topic, which is reflected in our data collection methodology.

The Parler social network

Parler is an alt-tech platform mimicking the affordances of X (previously Twitter). Launched in 2018, the social network had minimal rules and content guidelines, banning only "unlawful acts" and spambots (Parler, 2021). The lax content moderation rules made it a haven for users banned from and those who thought their speech to

be censored by mainstream platforms. Registered users were able to write short posts up to 1000 characters long, share images and links, and engage with other users' posts through comments and votes. Posts were displayed in individual feeds from followed users and were not algorithmically curated. We selected Parler as it was a popular alt-tech platform, with over 13.25 million users in January 2021. Parler was particularly popular among users banned from mainstream platforms and played a key role in mobilization efforts for the January 6, 2021, storming of the US Capitol. The platform was shut down by Amazon Web Services after the riots but has now returned online after a period of instability (Ingram and Brooks, 2023). Parler has been studied extensively in various contexts, such as the Capitol storming (Jakubik et al., 2023), the 2020 US presidential election (Chen et al., 2023), and the QAnon Movement (Sipka et al., 2022). Due to our focus on only conspiratorial narratives (rather than comparisons to nonconspiratorial narratives), Parler was an ideal source of data to use.

Data and data collection

We used a subset of the open Parler dataset (Aliapoulios et al., 2021). The dataset comprises 183 million Parler posts made between August 2018 and January 2021 as well as metadata from 13.25 million user profiles. Our dataset consists of the post content, the date created, an anonymous creator ID, and a post ID. For our dataset, we pulled posts from January 2020 until January 2021 that contained specific hashtags associated with each of the five conspiracy narratives. In doing this, we created separate datasets for each conspiracy narrative, thus the same posts could occur in more than one dataset. We retained these duplicates as it is common for posts to be applicable to more than one conspiracy narrative, and thus we wanted to ensure each dataset of each conspiracy narrative was "complete."

   During data cleaning and processing, we discovered that the dataset contained posts that were solely used for self-promotion purposes, as well as posts that contained only hashtags, emojis, or text that made no sense on its own (e.g., "This is so good!") because they contained links or images that were not included in the dataset. Further, some users used a large number of hashtags, emojis, and short phrases after every post as a signature-esque part of their content to signal their beliefs regardless of the post's content (see Figure 2). This resulted in a lot of posts being falsely pulled or mistagged due to the hashtags used.

   To improve accuracy of narrative taggings, we temporarily removed hashtags from posts with more than seven hashtags and then reran a keyword search to retag these posts. Posts were then assigned to their new narrative based on the updated keyword search. If no keywords matched, posts retained their original narrative tag. Further, posts were filtered for the English language. In total, the final dataset comprised $n = 30,478$ text-only posts (flat earth: 1,097 posts; 5G: 1,546 posts; antivaccine: 21,116 posts; false flag: 1,852 posts; the great replacement: 4,867 posts). Furthermore, we found that not every post is explicitly conspiratorial; however, we retained these posts as they contribute to the overall narrative of distrust and alternative reasoning and our data thus mirrors the experience of the user most closely.
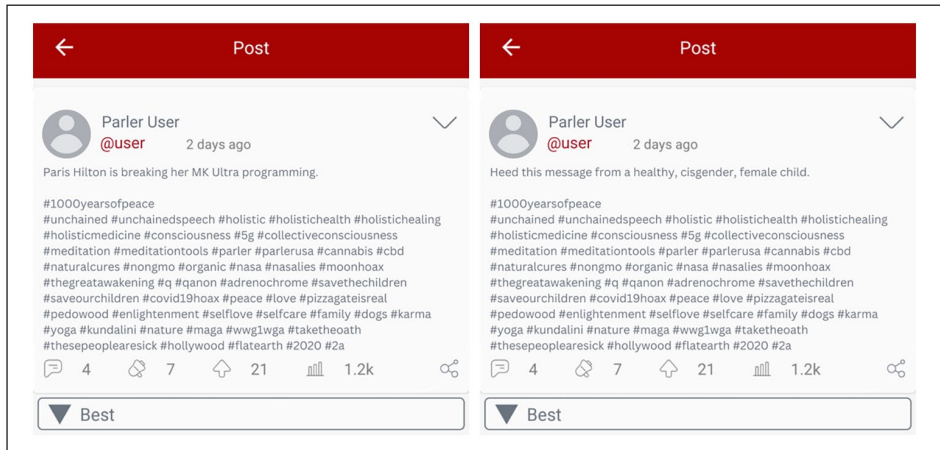
Figure 2. Two example posts of not explicitly conspiratorial texts with signature-esque uses of hashtags. Text is paraphrased from the data.

Dictionary analysis

The present research examines the proportion of words (%) in a post that corresponds to a given dictionary category. To calculate this, we first tokenize the text at the word-level and remove stop words, resulting in $N = 1,155,068$ word tokens. We then compare the words in each post to the words in the dictionary category, divide the count of matched words against the number of total words in the post, resulting in the percentage of words in a post that correspond to a dictionary category. The analysis was conducted in R, and full documentation and code can be found on the OSF (https://osf.io/gjcns/).

We measured emotions of anger, contempt, and disgust using the moral justification dictionary (Wheeler and Laham, 2016). The dictionary comprises 12 words for contempt (e.g., ridicul*, detest*), 20 words for anger (e.g., infuriat*, mad*), and 27 words for disgust (e.g., appall*, gross*). We selected the *Grievance Dictionary* (van der Vegt et al., 2021) as a measure for overt expressions of violence (269 words, e.g., shoot, stab), threat (151 words, e.g., attack, punish), hate (175 words, e.g., sicken, hate), planning (183 words, e.g., employ, implement), grievance (64 words, e.g., hurt, frustrat*), and paranoia (133 words, e.g., afraid, worry). The grievance dictionary dimensions were chosen to measure (1) overt expressions and discussions of violence via dimensions of violence, threat, hate, and planning; (2) underlying general grievances; and (3) paranoid thinking that has been associated with conspiratorial belief and violence (Hafez and Mullins, 2015; McCauley and Moskalenko, 2008; Oliver and Wood, 2014). Raw frequencies of dictionary categories can be found in Appendix A.

For further dictionary accuracy, we conducted a manual check of the words in each dictionary. Language use is nuanced, and certain words, such as "sick," can be used in multiple ways with a variety of meanings and implications. A manual check is, therefore, vital in the context of online data, as simple metrics like word counting can be skewed when taken out of context. We selected the top 10 most used words of each dictionary
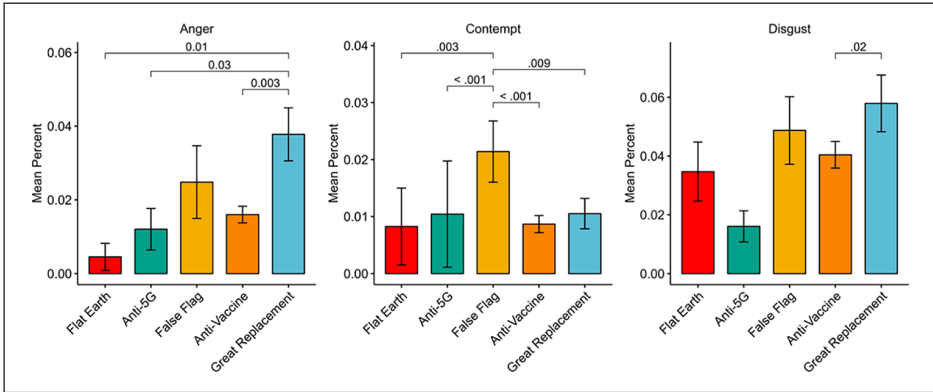
Figure 3. Mean percent of anger, contempt, and disgust words per post and significant differences between narratives. Bars left to right are in order of increasing associated violence. Error bars show standard error of the mean.

category, as well as 5 random words, which two raters separately and blindly labeled TRUE or FALSE against the dictionary category. We removed three words in total, one from disgust ("sick"), hate ("fight"), and planning ("agenda") dictionaries to reduce the skew in the dataset driven by false positives. While there may be instances of the words being used in the intended context of the dictionary, from scanning large portions of the data, we found this not to be the case and therefore took the decision to remove these words entirely. Further details can be found on the OSF.

Results

In the interest of space, we only report significant interactions. Full results, including nonsignificant findings, can be found on the OSF.

RQ1: How do expressions of anger, contempt, and disgust differ between conspiracy narratives?

We examined the proportion of post texts that expressed anger, contempt, or disgust across different conspiracy narratives using Kruskal-Wallis tests and Dunn post-hoc comparisons with Benjamini-Hochberg corrections. We chose BH corrections due to the exploratory nature of the analyses. We expected that proportions of anger, contempt, and disgust would be higher for narratives associated with more offline violence (Figure 3).

Anger: There were significant differences between conspiracy narratives for anger words, $H(4)$ 7.8, $p = .001$. Great replacement narrative posts contained a higher proportion of anger words than anti-5G narratives ($Z = 2.61$, $p = .03$), antivaccine narratives ($Z = -3.66$, $p = .003$), and flat earth narratives ($Z = 3.1$, $p = .01$).

Contempt: There were significant differences in contempt word usage between conspiracy narratives, $H(4) = 23.9$, $p < .001$. Post-hoc Dunn tests with Benjamini-Hochberg

adjustments revealed that false flag narratives contained a higher proportion of contempt words than anti-5G narratives ($Z = 3.91, p < .001$), flat earth narratives ($Z = 3.29, p = .003$), great replacement narratives ($Z = -2.92, p = .009$), and antivaccine narratives ($Z = -4.4, p < .001$).

Disgust: There were significant differences in disgust between conspiracy narratives, $H(4) = 16.2, p = .003$. Dunn post-hoc tests with Benjamini-Hochberg corrections revealed significant differences between great replacement narratives and antivaccine narratives ($Z = -3.05, p = .02$).

RQ2: Do expressions of violence, threat, hate, planning, grievance, and paranoia differ between conspiracy narratives?

We examined the proportion of post texts that expressed violence, threat, hate, grievance, or paranoia across different conspiracy narratives using Kruskal-Wallis tests and Dunn post-hoc comparisons with Benjamini-Hochberg corrections. We expect that the proportions of violence, threat, hate, grievance, and paranoia will be higher for narratives associated with more offline violence (Figure 4).

Violence: There were significant differences in violence expressions between conspiracy narratives, $H(4) = 521, p < .001$. Expressions of violence were higher for false flag narratives compared to anti-5G narratives ($Z = 12.6, p < .001$), flat earth narratives ($Z = 14.3, p < .001$), great replacement narratives ($Z = -12.8, p < .001$), and antivaccine narratives ($Z = -21.3, p < .001$). Great replacement narratives contained a higher proportion of violence words than anti-5G narratives ($Z = 2.90, p = .005$), flat earth narratives ($Z = 5.79, p < .001$), and antivaccine narratives ($Z = -10.3, p < .001$). Lastly, anti-5G narratives contained a higher proportion of violence words compared to flat earth narratives ($Z = -2.76, p = 0.007$) and antivaccine narratives ($Z = -3.02, p = .004$).

Threat: There were significant differences in threat-word use between conspiracy narratives, $H(4) = 668, p < .001$. False flag narratives contained a higher proportion of threat words than anti-5G narratives ($Z = 16.3, p < .001$), flat earth narratives ($Z = 18.3, p < .001$), great replacement narratives ($Z = -16.1, p < .001$), and antivaccine narratives ($Z = -24.4, p < .001$). Great replacement narratives contained a higher proportion of threat words than anti-5G narratives ($Z = 4.17, p < .001$), flat earth narratives ($Z = 7.66, p < .001$), and antivaccine narratives ($Z = -9.42, p < .001$). Vaccine narratives exhibited a higher proportion of threat words than flat earth narratives ($Z = 3.34, p = .001$). Anti-5G narratives contained a higher proportion of threat words than flat earth narratives ($Z = -3.39, p = .001$).

Hate: There were significant differences between conspiracy narratives in the proportion of hate words, $H(4) = 485, p < .001$. False flag narratives contained a significantly higher proportion of hate words than anti-5G narratives ($Z = 15.4, p < .001$), flat earth narratives ($Z = 15.7, p < .001$), great replacement narratives ($Z = -11.4, p < .001$), and antivaccine narratives ($Z = -19.3, p < .001$). Great replacement narratives had a higher proportion of hate words than anti-5G narratives ($Z = 7.49, p < .001$), flat earth narratives ($Z = 8.51, p < .001$), and antivaccine narratives ($Z = -9.77, p < .001$). Antivaccine
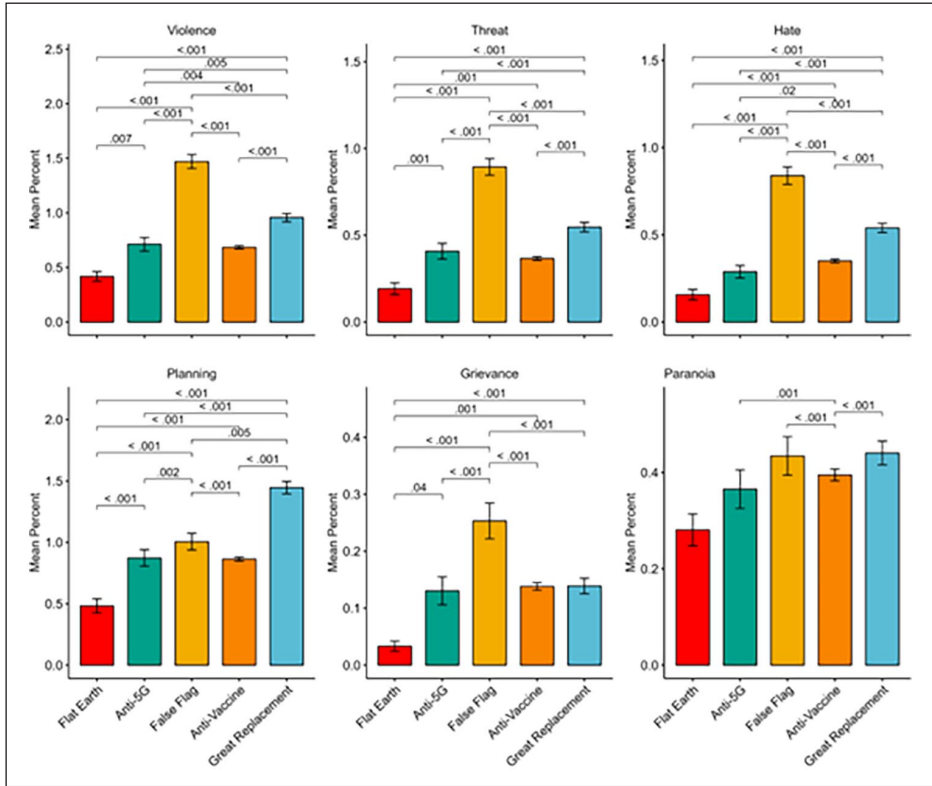
Figure 4. Mean percent of violence, threat, hate, grievance, and paranoia words per post and significant differences between narratives. Bars left to right are in order of increasing associated violence. Error bars show standard error of the mean.

narratives had a higher proportion of hate words than anti-5G narratives ($Z = 2.40, p = .018$) and flat earth narratives ($Z = 4.17, p < .001$).

Planning: There were significant differences between conspiracy narratives in the proportion of planning words, H(4) $= 207, p < .001$. False flag narratives used a higher percentage of planning words than anti-5G narratives ($Z = 3.19, p = .002$), flat earth narratives ($Z = 7.67, p < .001$), and antivaccine narratives ($Z = −4.77, p < .001$). Further, great replacement narratives had a significantly higher proportion of planning words than anti-5G narratives ($Z = 6.46, p < .001$), flat earth narratives ($Z = 10.8, p < .001$), and antivaccine narratives ($Z = −12.2, p < .001$). Anti-vaccine narratives used more planning words than flat earth narratives ($Z = 5.39, p < .001$). Anti-5G narratives used more planning words than flat earth narratives ($Z = −4.37, p < .001$).

Grievance: There were significant differences in the use of grievance words between conspiracy narratives, H(4) $= 132, p < .001$. False flag narratives contained a higher proportion of grievance words than anti-5G narratives ($Z = 8.09, p < .001$), flat earth narratives ($Z = 9.56, p < .001$), great replacement narratives ($Z = −8.61, p < .001$), and
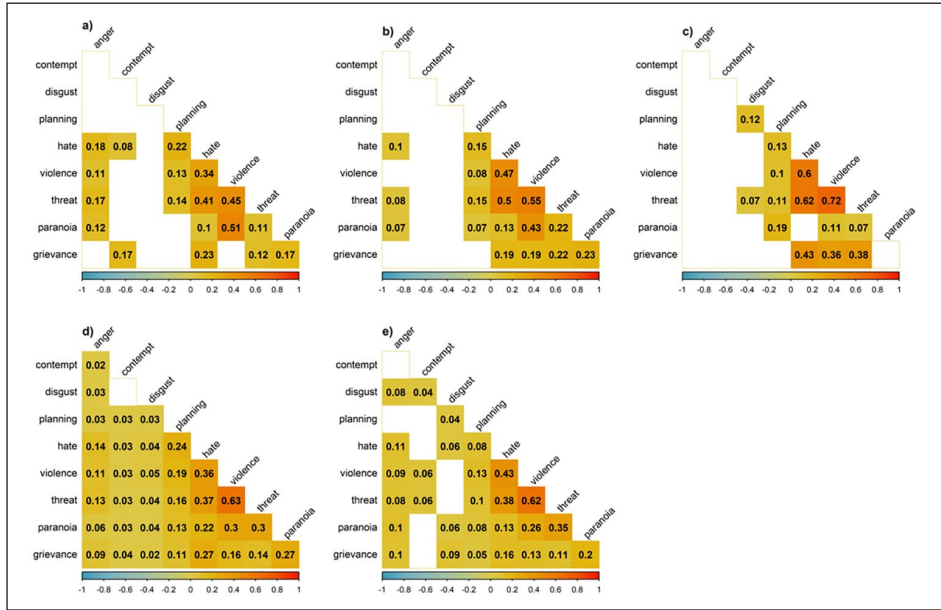
Figure 5. Correlation heatmaps for each conspiracy narrative. (a) flat earth, (b) anti-5G, (c) false flag, (d) anti-vaccine, (e) great replacement. Alpha = .01.

anti-vaccine narratives ($Z = -10.6$, $p < .0001$). Great replacement narratives contained a higher proportion of grievance words than flat earth narratives ($Z = 3.86$, $p < .001$). Antivaccine narratives contained a higher proportion of grievance words than flat earth narratives ($Z = 3.45$, $p < .001$). Anti-5G narratives contained a higher proportion of grievance words than flat earth narratives ($Z = -2.16$, $p = .04$).

Paranoia: There were significant differences in paranoia words between conspiracy narratives, $H(4) = 44.2$, $p < .001$. Anti-5G narratives contained a higher proportion of paranoia words than anti-vaccine narratives ($Z = -3.57$, $p = .001$). False flag narratives contained a higher proportion of paranoia words than antivaccine narratives ($Z = -4.34$, $p < .001$). Great replacement narratives contained a higher proportion of paranoia words than antivaccine narratives ($Z = -4.56$, $p < .001$).

RQ3: Are the expressions of anger, contempt, and disgust emotions correlated with expressions of violence, grievance, and hate within conspiracy narratives?

We examined whether expressions of anger, contempt, and disgust emotions were correlated with expressions of violence, threat, hate, planning, grievance, and paranoia within conspiracy narratives. We expected correlations to be positive and significant regardless of the association of the narrative with violence (Figure 5). As the data was non-parametric, we used Spearman's correlations.

Flat earth: Anger was not significantly correlated with contempt and disgust ($p >$ .01). Anger was weakly correlated with hate ($\rho = 0.18$, $p <$.001), violence ($\rho = 0.11$, $p <$.001), threat ($\rho = 0.17$, $p <$.001), and paranoia ($\rho = 0.12$, $p < $ .001). Contempt was weakly correlated with grievance ($\rho = 0.17$, $p <$.001) and hate ($\rho = 0.08$, $p = $ .005). Hate, violence, and threat were moderately correlated ($\rho = 0.434$–0.51, $p <$.001). Violence was moderately correlated with paranoia ($\rho = 0.51$, $p <$.001).

Anti-5G: Anger, contempt, and disgust were not significantly correlated ($p >$ .01). Anger was positively correlated with hate ($\rho = 0.1$, $p < $ .001), threat ($\rho = 0.08$, $p < $ .001), and paranoia ($\rho = 0.07$, $p = $ .009). Hate, threat, and violence were moderately positively correlated with each other ($\rho = 0.47$- 0.55, $p < $ .001). Violence was moderately correlated with paranoia ($\rho = 0.43$, $p < $ .001) and grievance ($\rho = 0.19$, $p < $ .001).

False flag: Anger, contempt, and disgust were not significantly correlated ($p > $ .01). Disgust showed a significant weak positive correlation to planning ($\rho = 0.12$, $p < $ .001) and threat ($\rho = 0.07$, $p = $ .002). Further, hate, threat, and violence showed moderate to high positive correlations ($\rho = 0.6$–0.72, $p < $ .001). Violence was positively correlated with paranoia ($\rho = 0.11$, $p < $ .001) and grievance ($\rho = 0.36$, $p < $ .001).

Antivaccine: Anger showed a weak positive correlation with contempt ($\rho = 0.02$, $p = $ .004) and disgust ($\rho = 0.03$, $p < $ .001). Contempt was not significantly correlated with disgust ($p > $ .01). Anger showed a positive weak significant correlation with planning ($\rho = 0.03$, $p = $ .001), hate ($\rho = 0.13$, $p < $ .001), violence ($\rho = 0.11$, $p < $ .001), threat ($\rho = 0.13$, $p < $ .001), paranoia ($\rho = 0.06$, $p < $ .001), and grievance ($\rho = 0.09$, $p < $ .001). Contempt was positively correlated with planning ($\rho = 0.03$, $p < $ .001), hate ($\rho = 0.03$, $p < $ .001), violence ($\rho = 0.03$, $p < $ .001), threat ($\rho = 0.03$, $p < $ .001), paranoia ($\rho = 0.03$, $p < $ .001), and grievance ($\rho = 0.04$, $p < $ .001). Disgust showed a weak positive correlation with planning ($\rho = 0.03$, $p < $ .001), hate ($\rho = 0.04$, $p < $ .001), violence ($\rho = 0.05$, $p < $ .001), threat ($\rho = 0.04$, $p < $ .001), paranoia ($\rho = 0.04$, $p < $ .001), and grievance ($\rho = 0.02$, $p < $ .001). Further, hate, violence, and threat were moderately positively correlated ($\rho = 0.36$ - 0.63, $p < $ .001). Violence was positively correlated with paranoia ($\rho = 0.3$, $p < $ .001) and grievance ($\rho = 0.16$, $p < $ .001).

The great replacement: Anger showed a weak positive correlation with disgust ($\rho = 0.08$, $p < $ .001). Contempt showed a weak correlation with disgust ($\rho = 0.04$, $p < $ .001). Anger was not correlated with contempt ($p > $ .01). Anger showed a weak positive correlation with hate ($\rho = 0.11$, $p <$.001), violence ($\rho = 0.09$, $p < $ .001), threat ($\rho = 0.08$, $p < $ .001), paranoia ($\rho = 0.1$, $p < $ .001), and grievance ($\rho = 0.1$, $p < $ .001). Contempt showed a weak positive correlation with violence ($\rho = 0.06$, $p < $ .001) and threat ($\rho = 0.06$, $p < $ .001). Disgust was positively correlated with planning ($\rho = 0.04$, $p < $ .001), hate ($\rho = 0.06$, $p < $ .001), paranoia ($\rho = 0.06$, $p < $ .001), and grievance ($\rho = 0.09$, $p < $ .001). Hate, violence, and threat were moderately correlated ($\rho = 0.38$–0.62, $p < $ .001). Violence was correlated with paranoia ($\rho = 0.26$, $p < $ .001) and grievance ($\rho = 0.13$).

Discussion

The aim of this paper was to investigate whether there were differences in expressions of anger, contempt, and disgust emotions between different conspiracy narratives, as well as overt expressions of grievances, hate, threat, violence, and paranoia. We also explored how these expressions were correlated within each narrative. Our findings support the ANCODI model as a model of violence legitimation in conspiracy narratives (Matsumoto and Hwang, 2012) and show the role of anger, contempt, and disgust as sensemaking emotions that open pathways to legitimating and encouraging violence. Specifically, we found that anger and contempt were more frequently expressed in narratives linked to violent events, showing that both situational outrage and a continued devaluation and dehumanization of the outgroup can legitimize and motivate action against the outgroup. Our findings further highlight the discussion of violence, threat, hate, grievance, and planning in more violent narratives. We found that emotions of anger, contempt, and disgust were correlated with grievance, paranoia, threat, hate, and violence, providing further evidence for the use of emotions in violence-legitimating narratives.

RQ1: Do expressions of ANCODI emotions differ between conspiracy narratives?

We found significant differences across conspiracy narratives in expressions of anger, contempt, and disgust. As expected, the proportions of post texts containing ANCODI emotions were highest for narratives that were associated with more violent events (e.g., false flag or great replacement narratives) and lowest for flat earth narratives. Disgust words were highest in great replacement narratives, which differed significantly only from antivaccine narratives. These results align with the ANCODI model, suggesting that narratives associated with more offline violence may be using anger, contempt, and disgust to address situational grievances, as well as delegitimize an outgroup.

RQ2: Do expressions of hate, grievance, threat, planning, and violence differ between conspiracy narratives?

We found significant differences in expressions of grievance, hate, planning, threat, violence, and paranoia. Grievance, hate, planning, threat, and violence were most prevalent in narratives associated with more violence (false flag and great replacement narratives). Violence, threat, and hate words were most prevalent in false flag narratives. Whilst part of the prevalence of violence words may have been driven by the discussion of violent events themselves, the prevalence of threat and hate words indicates that users were discussing violence-legitimating strategies in relation to the events themselves.

Paranoia words showed similar high-prevalence levels across all narratives, though false flag and great replacement narratives were still showing the highest prevalence. This is likely due to the fact that paranoia is a key component of all conspiracy narratives (Oliver and Wood, 2014).

RQ3: Is the expression of ANCODI emotions correlated with expressions of violence, grievance, threat, planning, and hate within conspiracy narratives?

Within narratives, we found weak correlations of ANCODI emotions with expressions of grievance, paranoia, hate, threat, and violence and moderate correlations between violence, threat, and hate, as well as violence, paranoia, and grievance. These findings further reflect the ANCODI model's emphasis on the interconnection of emotions and their role in violence legitimation. For example, in antivaccine and great replacement narratives, anger and hate are correlated with grievance expressions, reflecting situational sensemaking that can further legitimize hate and violence in narratives. Further, disgust was correlated with planning and hate in the false flag, antivaccine, and great replacement narratives, which indicates that delegitimization of the outgroup can aid in hateful rhetoric and mobilization.

We reveal how anger, contempt, and disgust are expressed across multiple conspiracy narratives associated with a variety of violent and nonviolent outcomes and highlight the role these emotions play in facilitating and legitimating collective action, supporting the ANCODI model (Matsumoto et al., 2013). For effective development of prevention and countermeasures, understanding the motivating forces behind collective and individual action in an ecologically valid setting is key. Hence, our study contributes novel insights by exploring narratives and expressions of emotion in a setting by using real social media data and capturing people's thoughts and emotions outside of a laboratory setting. We found that violence, threat, and hate were discussed primarily by narratives that are associated with more offline violence and were correlated with anger, contempt, and disgust. Whilst we cannot make causal inferences, our results suggest that discussion of violence and related concepts can occur at the same time as emotions associated with violence legitimation.

In real-world data, it can be difficult to fully differentiate conspiracy narratives: superconspiracies such as QAnon tie together conspiracy theories and use singular event conspiracy theories such as those in the false flag category as proof of a wider sinister plot (Harambam and Aupers, 2021). This was evident in our data where hashtags from other conspiracy theories were used to corroborate arguments or simply state one's beliefs, leading to narratives mixing and expanding to other seemingly less conspiratorial themes. Particularly in the age of algorithmic feeds, this highlights the connected nature of all conspiracy narratives (Cinelli et al., 2022). This means that a user will likely encounter a variety of narratives on their feed, even if they are originally only interested in one topic, which may inadvertently expose them to further grievances and violence legitimating narratives and aid their journey "down the rabbit hole." Given the wide adaptation of conspiracy narratives by extremist actors (Bartlett and Miller, 2010), this can also lead to users getting exposed to extreme ideological viewpoints and further aid radicalization journeys.

Methodological contribution

Given the constraints of dictionary analyses as a methodology for detecting complex emotional constructs such as disgust, two raters calibrated the dictionaries to see if they

were fit for the study context. This is an important element often missing from the social sciences, where the idea of calibrating measurement tools is less common than in STEM subjects. For example, in natural language processing, tuning models like BERT for specific applications (e.g., hate speech) is common to improve model performance. While these methods are more complex than dictionary approaches, we believe it is critical to evaluate how we are measuring constructs, especially complex emotions like disgust or contempt. Beyond considering usual dictionary limitations (e.g., spelling errors), we focus here on the contextual use of words in the observed community.

Notably, three words were not contextually relevant ("sick" in disgust, "agenda" in planning, and "fight" in hate) and were removed for accuracy. After refining the dictionary, we observed a large decrease in the occurrence of disgust-related words within all but primarily in antivaccine narratives. This shift is likely due to the use of "sick" as a term for illness rather than an expression of disgust. We further found weak changes in correlations of disgust with other measures in antivaccine narratives. We also found a small decrease in planning word occurrence for great replacement narratives, as the word is likely to be used in connection to Agenda 2030. Correlations of threat, hate, and violence decreased for all narratives as "fight" was a word within all three measures.

Our findings illustrate the difficulties of using dictionary methods for the detection of emotion and complex linguistic rhetoric. Text mining approaches like dictionaries are becoming increasingly popular within fields of political and social sciences (e.g., Ebner et al., 2023; Kennedy et al., 2022; van der Vegt et al., 2021), due to their ability to deal with large datasets, low computational intensity, and accessibility to non-technical researchers. However, dictionaries do not take context information into account; while some context-sensitive dictionaries have been developed (Ebner et al., 2023; Kennedy et al., 2022), it is important to note that these dictionaries still face limitations in distinguishing multiple meanings of words. Thus, including or excluding words from the dictionary can introduce both type 1 and type 2 errors. Hence, we only report the findings from our calibrated dictionaries. Combining qualitative insights with quantitative text mining (e.g., computational grounded theory, Nelson, 2020) can provide deeper insights and address the shortcomings of text mining methods.

Limitations

Social media environments include information beyond post text, such as username, profile photo, added links or media, or emojis that we did not analyze. Emojis and attached media can include crucial information about the text that conveys emotions or context (Kaye et al., 2021). Usernames and profile photos also can convey an account's group identity and sometimes even posting intent. Matsumoto and Hwang (2015) stress that the communication of anger, contempt, and disgust occurs not only through the use of emotion-laden words but also through nonverbal means such as gestures, facial expressions, and images. Here, we analyzed only text and excluded special characters, which means we potentially overlooked additional context that could have provided further insights into the emotions conveyed beyond the text of the post.

Further, as false flag narratives inherently talk about violent topics, measurements of violence may be inflated. A qualitative check of the data suggests that violent talk is still

present in the data, as represented in the other measurement dimensions of threat and hate. We note that our work focused exclusively on conspiratorial narratives on Parler; further research could explore the differences in levels of evoked emotions across various alt-tech and mainstream platforms (see also Jakubik et al., 2023).

Conclusion

Overall, we found significant differences in expressions of anger, contempt, and disgust across conspiracy narratives, with narratives that were associated with more offline violence using more emotive words. Our work highlights the important role of emotion in legitimizing violence and poses implications for early-stage prevention efforts, which can utilize grievance-focused counternarratives to divert those interested in more harmful narratives (see also Reed et al., 2017).

Additionally, we note the importance of validating the methods, and specifically here, dictionaries through qualitative checks. We note that researchers must consider calibrating and evaluating their measurement tools, especially when looking at specific communities, whose use of words may be highly specific and thus cause inaccurate and inflated results when using dictionaries/word counting techniques in particular. Combining computational approaches with qualitative insights is, therefore, key for future work aiming to explore complex psychological and rhetorical mechanisms in large textual datasets. We, therefore, encourage future work to utilize a mixed methods approach to strengthen computational findings.

ORCID iD

Darja Wischerath https://orcid.org/0000-0003-4657-503X

References

Aliapoulios M, Bevensee E, Blackburn J, et al. (2021) A large open dataset from the Parler social network. *Proceedings of the International AAAI Conference on Web and Social Media* 15: 943–951.

Baele SJ, Brace L and Coan TG (2023) Uncovering the far-right online ecosystem: an analytical framework and research agenda. *Studies in Conflict & Terrorism* 46(9). Routledge: 1599–1623.

Bartlett J and Miller C (2010) The power of unreason: conspiracy theories, extremism and counter-terrorism. *Demos*. Epub ahead of print 2010.

BBC News (2021) Capitol riots timeline: What happened on 6 January 2021? 10 February. Available at: https://www.bbc.com/news/world-us-canada-56004916 (accessed 22 November 2023).

Brown O, Smith LGE, Davidson BI, et al. (2022) The problem with the internet: An affordance-based approach for psychological research on networked technologies. *Acta Psychologica* 228: 103650.

Bruns A, Harrington S and Hurcombe E (2020) "Corona? 5G? or both?": the dynamics of COVID-19/5G conspiracy theories on Facebook. *Media International Australia* 177(1): 12–29.

Chen B, Lukito J and Koo GH (2023) Comparing the #StopTheSteal Movement across multiple platforms: differentiating discourse on Facebook, Twitter, and Parler. *Social Media + Society* 9(3).

Cinelli M, Etta G, Avalle M, et al. (2022) Conspiracy theories and social media platforms. *Current Opinion in Psychology* 47: 101407.

Dailey D and Starbird K (2015) "It's Raining Dispersants": Collective sensemaking of complex information in crisis contexts. In: *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing*, Vancouver BC Canada, 28 February 2015, pp. 155–158. ACM. Available at: https://dl.acm.org/doi/10.1145/2685553.2698995 (accessed 8 August 2022).

Davey J and Ebner J (2019) *The "Great Replacement": The Violent Consequences of Mainstreamed Extremism*. London: Institute for Strategic Dialogue.

de Keulenaar E (2023) The affordances of extreme speech. *Big Data & Society* 10(2).

Douglas KM, Uscinski JE, Sutton RM, et al. (2019) Understanding conspiracy theories. *Political Psychology* 40(S1): 3–35.

Doyle M and Dolan M (2006) Predicting community violence from patients discharged from mental health services. *The British Journal of Psychiatry* 189(6): 520–526.

Ebner J, Kavanagh C and Whitehouse H (2023) Assessing violence risk among far-right extremists: a new role for natural language processing. *Terrorism and Political Violence* 36(7): 1–18.

Franks B, Bangerter A, Bauer MW, et al. (2017) Beyond "Monologicality"? Exploring conspiracist worldviews. *Frontiers in Psychology* 8.

Glick P (2005) Choice of scapegoats. In: *On the Nature of Prejudice: Fifty Years after Allport*. Malden: Blackwell Publishing, pp. 244–261.

Hafez M and Mullins C (2015) The radicalization puzzle: a theoretical synthesis of empirical approaches to homegrown extremism. *Studies in Conflict & Terrorism* 38(11): 958–975.

Harambam J and Aupers S (2021) From the unbelievable to the undeniable: Epistemological pluralism, or how conspiracy theorists legitimate their extraordinary truth claims. *European Journal of Cultural Studies* 24(4): 990–1008.

Heller J (2017) AP WAS THERE: Black men untreated in Tuskegee Syphilis Study. *AP News*, 10 May. Washington, DC. Available at: https://apnews.com/article/business-science-health-race-and-ethnicity-syphilis-e9dd07eaa4e74052878a68132cd3803a (accessed 22 November 2023).

Ingram D and Brooks A (2023) Parler social media app to relaunch in 2024 ahead of election. *NBC News*, 18 December. Available at: https://www.nbcnews.com/tech/tech-news/parler-app-relaunch-2024-new-ownership-download-trump-rcna130287 (accessed 16 September 2024).

Iyer A, Schmader T and Lickel B (2007) Why individuals protest the perceived transgressions of their country: the role of anger, shame, and guilt. *Personality and Social Psychology Bulletin* 33(4): 572–587.

Jakubik J, Vössing M, Pröllochs N, et al. (2023) Online emotions during the storming of the U.S. capitol: evidence from the social media network Parler. *Proceedings of the International AAAI Conference on Web and Social Media* 17: 423–434.

Jasser G, McSwiney J, Pertwee E, et al. (2021) "Welcome to #GabFam": Far-right virtual community on Gab. *New Media & Society*: 146144482110245.

Jolley D and Paterson JL (2020) Pylons ablaze: Examining the role of 5G COVID-19 conspiracy beliefs and support for violence. *British Journal of Social Psychology* 59(3): 628–640.

Kaye LK, Rodriguez-Cuadrado S, Malone SA, et al. (2021) How emotional are emoji?: Exploring the effect of emotional valence on the processing of emoji stimuli. *Computers in Human Behavior* 116: 106648.

Kennedy B, Atari M, Davani AM, et al. (2022) Introducing the Gab Hate Corpus: defining and applying hate-based rhetoric to social media posts at scale. *Language Resources and Evaluation* 56(1): 79–108.

Kou Y, Gui X, Chen Y, et al. (2017) Conspiracy talk on social media: collective sensemaking during a public health crisis. *Proceedings of the ACM on Human-Computer Interaction* 1(CSCW): 1–21.

Kruglanski AW, Molinario E, Ellenberg M, et al. (2022) Terrorism and conspiracy theories: A view from the 3N model of radicalization. *Current Opinion in Psychology* 47: 101396.

Lazarus RS (1991) Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist* 46(8). US: American Psychological Association: 819–834.

Matsumoto D, Frank MG and Hwang HC (2015) The role of intergroup emotions in political violence. *Current Directions in Psychological Science* 24(5): 369–373.

Matsumoto D and Hwang S (2012) The role of emotion in predicting violence. *FBI Law Enforcement Bulletin* 81(1): 12.

Matsumoto D, Hwang HC and Frank MG (2013) Emotional language and political aggression. *Journal of Language and Social Psychology* 32(4): 452–468.

McCauley C and Moskalenko S (2008) Mechanisms of political radicalization: pathways toward terrorism. *Terrorism and Political Violence* 20(3): 415–433.

Miceli M and Castelfranchi C (2018) Contempt and disgust: the emotions of disrespect. *Journal for the Theory of Social Behaviour* 48(2): 205–229.

Nelson LK (2020) Computational grounded theory: a methodological framework. *Sociological Methods & Research* 49(1): 3–42.

Nouri L, Lorenzo-Dus N and Watkin A-L (2021) Impacts of radical right groups' movements across social media platforms – a case study of changes to Britain first's visual strategy in its removal from Facebook to Gab. *Studies in Conflict & Terrorism*: 1–27.

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.

Parler (2021) Parler Community Guidelines. Available at: https://legal.parler.com/documents/guidelines.pdf (accessed 27 June 2022).

Parveen N and Waterson J (2020) UK phone masts attacked amid 5G-coronavirus conspiracy theory. *The Guardian*, 4 April. Available at: https://www.theguardian.com/uk-news/2020/apr/04/uk-phone-masts-attacked-amid-5g-coronavirus-conspiracy-theory (accessed 22 November 2023).

Reed DA, Ingram DHJ and Whittaker J (2017) *Countering Terrorist Narratives*. Study. European Policy Department for Citizens' Rights and Constitutional Affairs.

Rogers R (2020) Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. *European Journal of Communication* 35(3): 213–229.

Rozin P, Haidt J and McCauley CR (2008) Disgust. In: *Handbook of Emotions, 3rd Ed*. New York, NY, US: The Guilford Press, pp. 757–776.

Scrivens R, Wojciechowski TW, Freilich JD, et al. (2023) Comparing the online posting behaviors of violent and non-violent right-wing extremists. *Terrorism and Political Violence* 35(1): 192–209.

Seidman L (2022) Man attacks workers at Calif. COVID-19 vaccine clinic, calling them "murderers." *The Seattle Times*, 5 January. Seattle. Available at: https://www.seattletimes.com/nation-world/nation/man-attacks-workers-at-calif-covid-19-vaccine-clinic-calling-them-murderers/ (accessed 22 November 2023).

Simi P and Windisch S (2020) The culture of violent talk: an interpretive approach. *Social Sciences* 9(7): 120.

Sipka A, Hannak A and Urman A (2022) Comparing the language of QAnon-related content on Parler, Gab, and Twitter. In: *14th ACM Web Science Conference 2022*, Barcelona Spain, 26 June 2022, pp. 411–421. ACM. Available at: https://dl.acm.org/doi/10.1145/3501247.3531550 (accessed 23 October 2023).

Smith A (2022) Stop employees from "Mask Shaming" colleagues. In: *SHRM*. Available at: https://www.shrm.org/resourcesandtools/legal-and-compliance/employment-law/pages/stop-mask-shaming.aspx (accessed 22 November 2023).

Smith LGE, Piwek L, Hinds J, et al. (2023) Digital traces of offline mobilization. *Journal of Personality and Social Psychology* 125(3): 496–518.

Southern District of California (2021) Santa Barbara Man Indicted in San Diego for Killing his Children in Mexico. Available at: https://www.justice.gov/usao-sdca/pr/santa-barbara-man-indicted-san-diego-killing-his-children-mexico (accessed 22 November 2023).

Sternberg RJ (2003) A duplex theory of hate: development and application to terrorism, massacres, and genocide. *Review of General Psychology* 7(3): 299–328.

Sullivan D, Landau MJ and Rothschild ZK (2010) An existential function of enemyship: Evidence that people attribute influence to personal and political enemies to compensate for threats to control. *Journal of Personality and Social Psychology* 98(3): 434–449.

Sunstein CR and Vermeule A (2009) Conspiracy Theories: Causes and Cures*. *Journal of Political Philosophy* 17(2): 202–227.

Sutton RM and Douglas KM (2022) Rabbit hole syndrome: Inadvertent, accelerating, and entrenched commitment to conspiracy beliefs. *Current Opinion in Psychology*: 101462.

Sweney M and Waterson J (2020) Arsonists attack phone mast serving NHS Nightingale hospital. *The Guardian*, 14 April. Available at: https://www.theguardian.com/technology/2020/apr/14/arsonists-attack-phone-mast-serving-nhs-nightingale-hospital (accessed 13 December 2023).

Törnberg A and Törnberg P (2023) White supremacists anonymous: how digital media emotionally energize far-right movements. *Journal of Information Technology & Politics*. Advance Online Publication. DOI: 10.1080/19331681.2023.2262459.

van Buuren J (2013) Spur to Violence?: *Anders Behring Breivik and the Eurabia conspiracy. Nordic Journal of Migration Research* 3(4): 205–215.

van der Vegt I, Mozes M, Kleinberg B, et al. (2021) The Grievance Dictionary: Understanding threatening language use. *Behavior Research Methods* 53(5): 2105–2119.

Vanderwee J and Droogan J (2023) Testing the link between conspiracy theories and violent extremism: a linguistic coding approach to far-right shooter manifestos. *Behavioral Sciences of Terrorism and Political Aggression*. Advance Online Publication. DOI: 10.1080/19434472.2023.2258952.

van Prooijen J-W (2011) Suspicions of injustice: the sense-making function of belief in conspiracy theories. In: Kals E and Maes J (eds) *Justice and Conflicts*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 121–132. Available at: http://link.springer.com/10.1007/978-3-642-19035-3_7 (accessed 8 August 2022).

van Zomeren M, Postmes T and Spears R (2008) Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin* 134(4): 504–535.

van Zomeren M, Spears R, Fischer AH, et al. (2004) Put your money where your mouth is! Explaining collective action tendencies through group-based anger and group efficacy. *Journal of Personality and Social Psychology* 87(5): 649–664.

Vegetti F and Littvay L (2022) Belief in conspiracy theories and attitudes toward political violence. *Italian Political Science Review/Rivista Italiana di Scienza Politica* 52(1): 18–32.

Webber D and Kruglanski AW (2017) Psychological factors in radicalization: A "3 N" approach. In: LaFree G and Freilich JD (eds) *The Handbook of the Criminology of Terrorism*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 33–46. Available at: https://onlinelibrary. wiley.com/doi/10.1002/9781118923986.ch2 (accessed 30 December 2022).

Wheeler MA and Laham SM (2016) What we talk about when we talk about morality: deontological, consequentialist, and emotive language use in justifications across foundation-specific moral violations. *Personality and Social Psychology Bulletin*. Epub ahead of print 22 June 2016. DOI: 10.1177/0146167216653374.

Xiao S, Cheshire C and Bruckman A (2021) Sensemaking and the chemtrail conspiracy on the internet: insights from believers and ex-believers. *Proceedings of the ACM on Human-Computer Interaction* 5(CSCW2): 1–28.

Zeng J and Schäfer MS (2021) Conceptualizing "Dark Platforms." Covid-19-related conspiracy theories on 8kun and Gab. *Digital Journalism* 9(9): 1321–1343.

Zhu Y, Cheng EW, Shen F, et al. (2022) An eye for an eye? An integrated model of attitude change toward protest violence. *Political Communication* 39(4): 539–563.

Author biographies

Darja Wischerath is a doctoral researcher at the University of Bath. Her research explores how belief in conspiracy theories can lead to violence and violent extremism through the mediating factors of psychological needs, social networks, and violence legitimating narratives.

Dr. Lukasz Piwek is an associate professor of data science at the University of Bath, specializing in interdisciplinary research centered on behavior analytics and computational social science. His projects range from the identification of psychological markers in digital footprints, the application of machine learning in behavioral analytics, the development of custom devices and software for digital research methods, and the advancement of data visualization techniques.

Dr Jonathan F. Roscoe leads the Automated Cyber Detection and Response research group at BT. His research interests include open-source intelligence, AI/ML, and threat modeling. During the COVID-19 pandemic, his research focused on detecting and understanding the spread of COVID-related disinformation, human behavior, and modeling the rise in attacks against 5G infrastructure.

Dr Brittany I. Davidson is an associate professor of analytics at the University of Bath. Her work is focused on how we can use new and emerging forms of data (e.g., digital traces from social media, smartphones, sensors) to better understand various online and offline behaviors and how this can be fractured across the lifespan and across facets of our lives. She has a particular interest in measurement and how we conceptualize, operationalize, measure, model, and generate inferences about individuals and how they use technology for positive (and sometimes negative) consequences for themselves and others.

Appendix A.
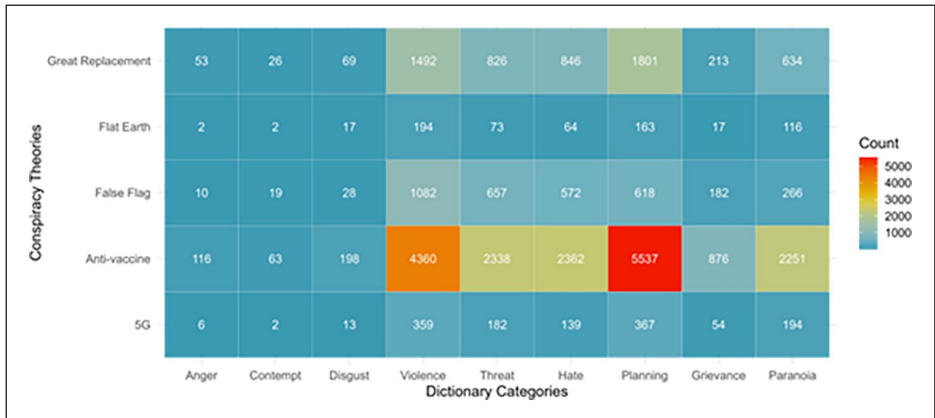
Appendix A. Heatmap of raw counts of dictionary scores.