TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 00 (2023) 1–19

ISSN: 1756-8765 online
DOI: 10.1111/tops.12702

This article is part of the topic "The Dynamical Hypothesis Three Decades Later: Advances, Critiques, and Prospects for a Dynamical Cognitive Science," Luis H. Favela and Vicente Raja (Topic Editors).

Understanding Collective Human Behavior in Social Media Networks Via the Dynamical Hypothesis: Applications to Radicalization and Conspiratorial Beliefs

Aaron Necaise,[a,†] Jingjing Han,[b,c,d] Hana Vrzáková,[e] Mary Jean Amon[a,†]

[a] *School of Modeling, Simulation, and Training, University of Central Florida*
[b] *School of Journalism, Fudan University*
[c] *Institute for Global Communications and Integrated Media, Fudan University*
[d] *Shanghai Key Laboratory of Data Science, Fudan University*
[e] *School of Computing, University of Eastern Finland*

Received 23 March 2023; received in revised form 2 October 2023; accepted 3 October 2023

Abstract

The dynamical hypothesis has served to explore the ways in which cognitive agents can be understood dynamically and considered dynamical systems. Originally used to explain simple physical systems as a metaphor for cognition (i.e., the Watt governor) and eventually more complex animal systems (e.g., bird flocks), we argue that the dynamical hypothesis is among the most viable approaches to understanding pressing modern-day issues that arise from collective human behavior in online social networks. First, we discuss how the dynamical hypothesis is positioned to describe, predict, and explain the time-evolving nature of complex systems. Next, we adopt an interdisciplinary perspective to describe how online social networks are appropriately understood as dynamical systems. We introduce a dynamical modeling approach to reveal information about emergent properties in social media, where radicalized conspiratorial beliefs arise via coordination between user-level and community-level

[†] Equal contribution/first authorship.
Correspondence should be sent to Mary Jean Amon, School of Modeling, Simulation, and Training, University of Central Florida, 3100 Technology Parkway, Orlando, FL 32826, USA. E-mail: mjamon@ucf.edu

comments. Lastly, we contrast how the dynamical hypothesis differs from alternatives in explaining collective human behavior in social networks.

Keywords: Dynamical hypothesis; Complex systems; Social networks; Wavelet analysis; Radicalization; Sentiment; Coordination

1. Introduction

The dynamical hypothesis in cognitive science forwards the notion that cognitive agents both qualify as dynamical systems and are properly understood dynamically (van Gelder, 1998). In the three decades from which this hypothesis was forwarded, cognitive science has incorporated dynamical accounts of cognitive activity ranging from linguistics, motor control, psychophysics, and sensation and perception (for review, see Favela, 2020). Yet, the dynamical hypothesis has been largely constrained to explaining single human cognitive agents, with some social psychological research extending the hypothesis to dyads or small groups (e.g., Marzoratti & Evans, 2022; Vrzáková, Amon, & D'Mello, 2021). Even so, the dynamical hypothesis is broadly applicable, not just to single agents or small groups, but to collective human behavior arising by virtue of social networks. Given the increased complexity of social psychological phenomenon in large networks of people, the dynamical hypothesis supports new, foundational understandings of collective human behavior.

The need for dynamical accounts of collective human behavior is illustrated by the issue of social media radicalization. Social media radicalization is the process of developing extreme and violent ideologies, which often occurs when users immerse themselves within a narrow range of communities that share similar beliefs (Borum, 2011). Radicalization comes in many forms (e.g., so-called "incel culture" [Salojärvi, Rantanen, Nieminen, Juote, & Hanhela, 2020], or conspiratorial beliefs about science), but a point of consistency is that radicalization is inherently a time-evolving process (Alimi, Bosi, & Demetriou, 2012; Rousseau, Hassan, & Oulhote, 2017) that progressively leads to extremism (Borum, 2011; Crossett & Spitaletta, 2010; Fernandez, Gonzalez-Pardo, & Alani, 2019; Schmid, 2013). Despite this fact, radicalization theories are typically described qualitatively and rarely empirically or dynamically modeled (Necaise et al., 2021; Odag, Leiser, & Boehnke, 2019; Fernandez et al., 2019). Consequently, little is known about individual users' radicalistic trajectories or how user practices and platform politics coshape intergroup radicalistic perspectives as a situated interaction (Rousseau et al., 2017, Matamoros-Fernández & Farkas, 2021; Smith, Blackwood, & Thomas, 2020). This critical knowledge gap leads to no consensus as to which radicalization theories are most promising (Dalgaard-Nielsen, 2010), and methods for detecting radicalization material remain imprecise (Panetta, 2022). Recent work has argued that our lack of understanding about complex social networks is so vast that the study of collective human behavior should be escalated to a crisis discipline like medical or climate sciences (Bak-Coleman et al., 2021). Social media radicalization is but one example of how the study of collective human behavior can be expanded by the dynamical hypothesis and provides new applications of the hypothesis originally forwarded 30 years ago.

In this paper, we use social media radicalization as an entry point for exploring the applicability of the dynamical hypothesis to collective behavior expressed in social media. We examine how the coupling between user and community text sentiment resonates within extremist social media networks. Using cross-spectrum wavelet analysis as one type of dynamical analysis, we empirically reveal facets of large-scale sentiment coupling between community members and generate novel insights on social media radicalization and conspiratorial beliefs. Cross-spectrum wavelet analysis is typically framed as exploring synchrony; however, social media interactions are inherently asynchronous. Thus, we refer to our analysis as investigating the coupling between the text sentiment of individual users and a radicalized online community to acknowledge the temporal ambiguity of online interactions. We address the following research questions (RQs):

RQ1. Do social media users demonstrate sentiment coupling with their online communities as a collective during text-based discussions?
RQ2. Does the degree of sentiment coupling between a user and an online community correspond to levels of extremist community engagement and the use of language associated with conflict?
RQ3. To what extent does coupling correspond to changes in language sentiment?
RQ4. To what extent do leader-follower dynamics between user and online community vary based on community engagement?

2. Related work

2.1. Leveraging the dynamical hypothesis to understand complex systems

Van Gelder (1995; 1998) proposed the dynamical hypothesis as an alternative to the computational approach to cognitive science, in which he argues that cognitive agents *are* not digital computers but dynamical systems (the nature hypothesis) and *can* and *should* be studied dynamically (the knowledge hypothesis). Such cognitive agents can form multiple dynamical systems as their properties interact with one another and with the surroundings in complex ways. For example, the complex interactions between neurons in the brain can be viewed as one dynamical system, while the interaction between the mind, the body, and its environment can be considered as another. Finally, the complex interactions between cognitive agents themselves form a higher-level dynamical system from which collective behavior emerges.

In this context, the dynamical hypothesis is a framework for understanding both individual cognition and collective behaviors. Dynamical systems underlie numerous mechanical and animal systems, for example, the Watt governor, human brain activity, or ant colonies (Olfati-Saber, 2006). Van Gelder used the Watt governor as a metaphor for cognition as it does not require a central computer to send commands to individual parts. The Watt governor is a device with several interlocked components that interact to regulate the speed of steam engines, and the system's behavior is the result of local "decisions" made by each part as opposed to a central controller (Favela, 2020). Similarly, bird flocks do not have a central commander; rather each bird changes its route based on the state of its neighbors, leading to a collective behavior. In other words, collective behavior is an emergent pattern that arises from interactions among individual parts through the process of self-organization (Strogatz, 2004). Consequently, analyzing isolated individuals in highly interactive environments does not fully describe complex system behaviors. The dynamical hypothesis recognizes the nature of collective behavior as an emergent property of a dynamical system and, therefore, directs researchers to ask questions about how local behaviors can give rise to collective patterns. Second, its epistemological proposal points out valuable analytical tools that researchers can use to describe and predict collective behavior, such as dynamical modeling.

2.2. Applications of nonlinear dynamics to complex social media networks

The growth of social media and access to archival data from the internet has provided an opportunity to examine collective behaviors at a scale that was not previously possible (Bak-Coleman et al., 2021). Given the importance of social media for socializing and information seeking (Akram & Kumar, 2017; Bak-Coleman et al., 2021), understanding the dynamics underlying social media usage is highly valuable. Online social interactions play a role in shaping our political (Tucker et al., 2018), medical (e.g., Allington, Duffy, Wessely, Dhavan, & Rubin, 2021; Benoit & Mauldin, 2021), and societal views (Akram & Kumar, 2017). Thus, the dynamics driving those influences have implications for public health and safety.

The common approach utilizes network analysis to quantify the structure and dynamics of relationships within networks forming in social media (Blonder, Wey, Dornhaus, James, & Sih, 2012; Bródka, Kazienko, Musiał, & Skibicki, 2012; Farine, 2018). With social network analysis, researchers have described how information disseminates on social media (Luo & Zhong, 2015; Panzarasa, Opsahl, & Carley, 2009), as well as how popular internet figures exert influence (del Fresno García, Daly, & Segado Sanchez-Cabezudo, 2016). Similar methods (i.e., dynamic network analysis) have been used by Aydin and Perdahci (2019) to depict the growth of an online health community and determine when that community transitioned into a stable state in terms of the density of user relationships. Although network analysis is a viable approach to understanding complex behaviors, it is limited to user relationships and struggles with modeling network evolution (Farine, 2018).

Beyond network analysis, the application of dynamical analyses to social media data has been relatively sparse, especially those that characterize nonlinear aspects of change. As with social systems in the real-world, online social networks appear to be fractal with respect to how they are structured (Tsugawa & Ohsaki, 2014) and how they attract activity (Muchnik et al., 2013; Johnson, Faraj, & Kudaravalli, 2014). Systems that are fractal are geometrically or statistically similar across multiple scales of behavior, and fractal analysis is a method for quantifying that self-similarity (for an introduction to fractal analysis, see Brown & Liebovitch, 2010). Leveraging insights about fractals forming in nature, fractal analysis has been used to make inferences about the organizational complexity of social media communities (e.g., Oh, 2022). For example, Aguilera, Morer, Barandiaran, and Bedia (2013) applied detrended fluctuation analysis, a method for quantifying fractals in time series, to examine political movements on Twitter based on hashtag frequency, finding that political movements with fractal scaling tended to be more robust and longer lasting than those without. Although dynamical analysis of social media behaviors has the potential to offer new perspectives, relatively few studies (e.g., Necaise et al., 2021) have explored the topic of online radicalization.

2.3. *Social media radicalization as a dynamical system*

The importance of understanding and addressing dynamics within social media platforms is exemplified by the problem of online radicalization. Social networks support radicalism to an unprecedented degree by providing low publication thresholds, worldwide audiences, and increased anonymity (Agarwal & Sureka, 2015; Bertram, 2016; Koehler, 2014; Mondal, Silva, & Benevenuto, 2017; Tehrani, Manap, & Taji, 2013; Von Behr, 2013). Theories of radicalization highlight its insidious and dynamic nature, whereby beliefs change gradually through exposure to extremist material (Schmid, 2013).

In general, it is well-established that people alter their language to different contexts, often to promote feelings of rapport or to gain attention (Eckert & Rickford, 2001; Giles, Taylor, & Bourhis, 1973; Nguyen & Rose, 2011). Although changes in communication patterns may seem trivial, cognition is often scaffolded on dialogical processes. People find out what they think and feel by hearing themselves in conversation and having it reflected by others (Alfano, 2013; Alfano, Carter, & Cheong, 2018; Doris, 2015; Wong, 2006), which can result in normalizing attitudes and behaviors (Bjelopera, 2013). Research has extended these general sociolinguistic insights to understand how users adapt in online communities, for example, by becoming more uniform in language style, topics, and strategies over time (Cassell & Tversky, 2005; Nguyen & Rose, 2011). User adaptation to online community standards is driven by combinations of viewing behavior, feedback, and replies (Lampe & Johnston, 2005; Sharma & De Choudhury, 2018). Although this previous work is foundational to understanding influence in extremist networks, most research on radicalization focuses on content versus process. However, to fully understand what pushes individuals toward extremist ideologies, there remains a need for dynamical modeling of dialectical and other germane social media processes (Borum, 2011).

The lack of dynamical modeling limits the development and validation of radicalization theory. To date, a variety of theories outline unique stages of radicalization (Regehr, 2022; Borum, 2011; Dalgaard-Nielsen, 2010; Decety, Pape, & Workman, 2018; Gill, 2007; Hafez & Mullins, 2015; Moghaddam, 2005; Sageman, 2004; Silber, Bhatt, & Analysts, 2007; Sinai, 2012; Smith et al., 2016; Torok, 2013; Wiktorowicz, 2005; also see Crossett & Spitaletta, 2010 for review), and the diversity of these models highlight different aspects of radicalization that must be examined empirically. Although these theories differ in many respects, they describe the process of radicalization as having core similarities. Radicalization is described as (1) a time-evolving process (dynamic; Alfano et al., 2018; Alimi et al., 2012; Ayanian, Böckler, Doosje, & Zick, 2018; Baumann et al., 2020); (2) that is "not necessarily linear" (Hardy, 2018); (3) emphasizing interconnectedness between components and scales, such as users and online communities (systems); and (4) accounting for multifactorial processes of interaction (complex; see Fig. 1).

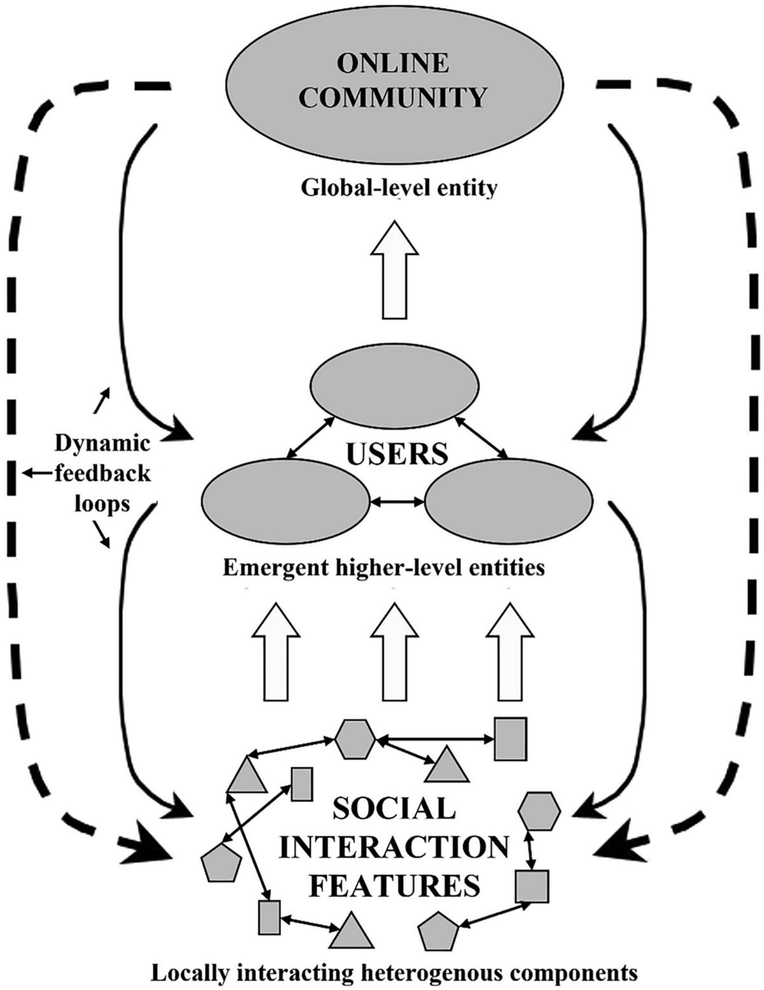

Fig. 1. Complex dynamics of social media intergroup radicalization. *Gray arrows* show the process through which social interaction features fuel user and online community intergroup radicalization. In turn, *unidirectional black arrows* illustrate how the online community constrains user attitudes and social interactions (adapted from Parrott, 2002).

These common points of radicalization theories are aligned with a conceptual and methodological framework of social media radicalization as a dynamical system. Consistent with van Gelder's two-point dynamical hypothesis, radicalized social media networks: (1) qualify as dynamical systems and (2) are properly studied as dynamical systems. Whereas the first point is supported by prior theory and common conceptualizations of radicalization, the dynamics leading to radicalization remain unclear due to the prior emphasis on qualitative and discrete approaches (Baumann et al., 2020). Without empirical and dynamical modeling, it remains unknown exactly which aspects of leading theories apply to social media. In this work, we demonstrate how a method associated with nonlinear dynamical systems theory provides new insights into the dynamics of social media radicalization.

3. Method

3.1. Data collection

We selected a random sample of 246 users who participated in discussions in the /r/Conspiracy Reddit community during 2022. Reddit is a social media platform where users can create and maintain their own forum-based communities (i.e., subreddits) focused on discussing specific topics (Anderson, 2015). We selected Reddit for analysis because it is an established source of social media radicalization (van Raemdonck, 2019) containing verbose and topic-specific conversations (Medvedev, Lambiotte, & Delvenne, 2019). Furthermore, Reddit provides users with pseudonymity, veiling their identity and potentially benefiting the radicalization process (Cilluffo, Cardash, & Whitehead, 2006). The /r/Conspiracy subreddit in specific is a community known for proliferating extreme conspiracy theories about science and politics (Samory & Mitra, 2018). The sample size was determined according to a power analysis for linear regression using G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) with a small effect size (.05) and $\alpha = .05$. Users were required to have at least 50 comments in the community to be included in the study, excluding comments that were no longer accessible. This cutoff provided a large enough sample size for the continuous wavelet transform to achieve a maximum frequency scale of at least 16 (discussed below; Issartel, Bardainne, Gaillot, & Marin, 2015).

Next, we used the Pushshift database and Reddit's application programming interface (API) to download every comment made by users in our sample within /r/Conspiracy since account creation. To collect information about the social interactions these individuals participated in, we also used the Reddit API to download the text content they responded to when making their personal comments. Users could have made their personal comments in response to the topic of a discussion thread or as replies to comments by other users. For simplicity, we refer to comments made by users in our sample as *user comments* and the text they interacted with as *peer comments*. Finally, we removed 12,375 (6.50% of the total) comments with less than three words from the data set, as these comments consisted primarily of emojis and URLs. We define a social interaction on the forum as the pairing of a user's comment with the comment of a peer that they interacted with when posting their personal comment. The final data set contained 86,429 interactions (i.e., 172,858 total comments) within /r/Conspiracy.

3.2. Data processing

Two validated natural language processing (NLP) methods were applied to extract information about the content of each interaction. The Linguistic and Word Count (LIWC-22; Boyd, Ashokkumar, Seraj, & Pennebaker, 2022) dictionary is an NLP tool designed to evaluate the psychological content of text. Using LIWC-22, we calculated the total number of words and the percentage of "conflict" words (e.g., "fight" or "argue") within each user and peer comment. The LIWC-22 dictionary can also estimate the sentiment of text, but the results are reported to be unreliable for short text passages (Boyd et al., 2022). To address this shortcoming, we opted to use the valence-aware dictionary and sentiment reasoner (VADER; Hutto & Gilbert, 2014) toolkit. VADER is a sentence-level analysis developed to analyze short blurbs of text from social media and has been validated against human reviewers (Hutto & Gilbert, 2014). Sentiment scores from VADER are reported on a composite scale from −1 *(most negative)* to 1 (*most positive*) with scores of 0 representing neutral valence. An example of a comment classified as positive is: "Thank you for the earnest response, I really appreciate it." On the other hand, an example of a comment identified as negative is: "He's a shifty-eyed [offensive language redacted], isn't he?" After sentiment analysis, each user in the sample had one set of time series containing the sentiment, word count, and percentage of conflict words in their personal comments (e.g., one time series per data stream), and a second set of time series with the same information of their peers' comments aligned at the point of interaction. The length of each user's time series corresponded to the number of interactions they had within the community.

3.3. Cross-wavelet analysis

Wavelet analysis is a method for examining time series in terms of their underlying frequency components. Frequency describes how quickly the values of a time series fluctuate "up-and-down," and multiple frequency components can be present in a single time series. In our work, high frequencies correspond to rapid fluctuations in sentiment (e.g., a shift from negative to positive) spanning several comments, whereas low frequencies refer to long-term trends in sentiment spanning tens or hundreds of comments. We expect sentiment to change rapidly depending on momentary interactions; however, it is also reasonable to expect long-term trends in sentiment as users entrain to community norms. These slow- versus fast-moving dynamics represent the "frequency components" underlying changes in sentiment, and our goal is to untangle these processes using a multiscale wavelet analysis.

Specifically, we use cross-wavelet analysis to investigate the coupling between a user's comment sentiment on the conspiracy subreddit and the comments of their peers within the community. Using cross-wavelet analysis, we "thin-sliced" (Ambady & Rosenthal, 1992) users' engagement with the online community by extracting their successive interactions to understand broader cognitive processes as they unfold over time. Thin slices are brief samples of activity that can be used to make inferences about more general behaviors and beliefs (Ambady & Rosenthal, 1992). Each user-to-peer interaction (e.g., a user commenting their opinion in a discussion thread) qualifies as a thin slice, as they represent small intersections of engagement with the online community that punctuate more prolonged scrolling, reading, and "lurking" behavior when compiled.

Cross-wavelet analysis is a multiscale approach representing the similarities between two systems by frequency, quantifying patterns otherwise hidden in the raw data. Cross-wavelet analysis offers several advantages compared to simple correlations. First, wavelet methods provide additional information about nonstationary systems where different patterns of

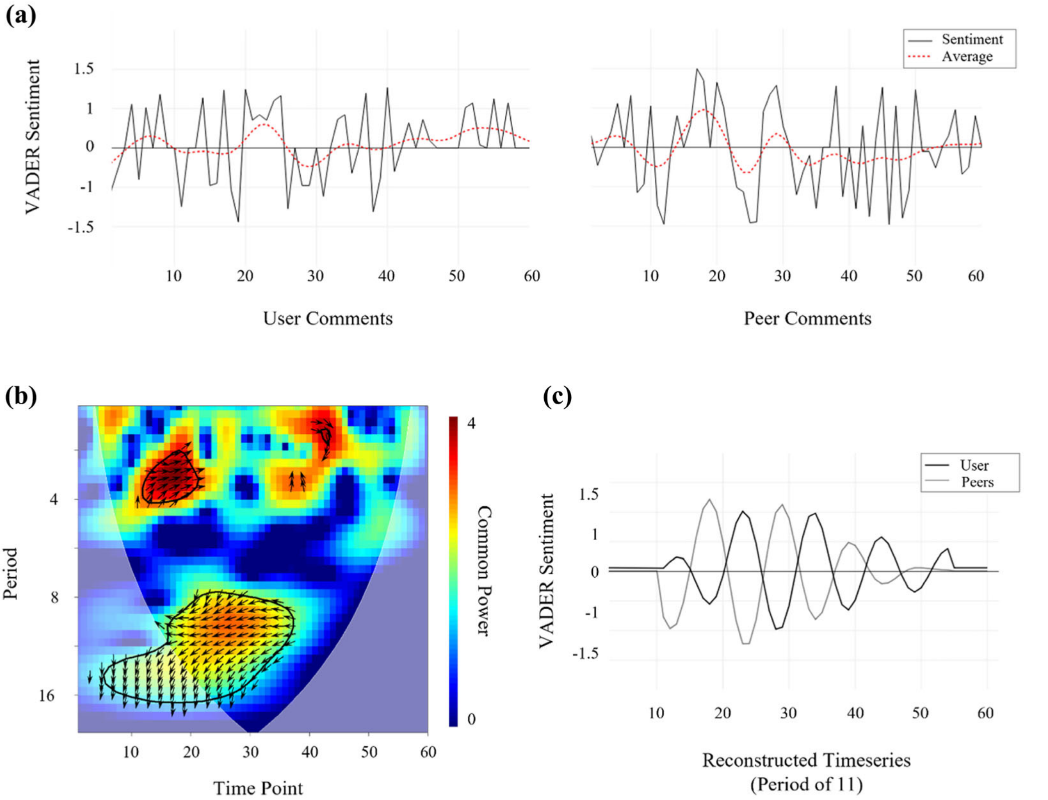

Fig. 2. Example cross-wavelet analysis applied to one user's comment sentiment and the comment sentiment of their peers. Panel A contains the sentiment time series of user comments on the left and the sentiment time series of peer comments on the right. Visual inspection reveals several common fluctuations between user and peer comments, particularly at time points 0–40. Panel B contains the cross-wavelet plot with time on the *x*-axis, period (e.g., frequency; the number of data points needed for a complete cycle) on the *y*-axis, and common power depicted by color intensity. Significant regions of common power are circled in black. According to the left-facing arrows in the plot, there is a slow-moving anti-phase relationship in the early sections of the data. This anti-phase relationship is visualized in panel C by retaining only fluctuations occurring at a period of 11 and reconstructing that specific frequency component.

behavior emerge over time, and the analyses can be implemented using linear or nonlinear wavelet transforms (e.g., Heijmans & Goutsias, 2000). Therefore, it is an analytical tool well-suited for describing nonlinear dynamical systems with frequent shifts in behavior. Second, the analysis describes the magnitude of interaction between two systems by identifying *points of common power* localized in both frequency *and* time. Finally, cross-wavelet analysis provides information about leader-follower dynamics with respect to *phase difference* (i.e., the time-lag between changes in two systems). In our context, the phase difference describes how social media users' sentiment changes in time based on the sentiment expressed by their community peers (see Fig. 2 for a practical example).

We computed the cross-wavelet transform between each user and their peer's sentiment using the "biwavelet" package in R (Gouhier, Grinsted, & Simko, 2018) with a Morlet mother wavelet function ($\omega_0 = 6$). Significance testing was performed with a time-average test to identify points of common high-power (i.e., covariance in time-frequency space) and normalized by variance. These were periods when users and their peers' sentiment fluctuated at similar rates and intensities (see Fig. 2, panel B). To estimate the overall degree of sentiment coupling exhibited by each user, we calculated the average common power within significant regions of each cross-wavelet plot and considered higher common power as a measure of increased coupling. We also extracted a relative phase (RP) angle from each user to describe leader-follower dynamics. An RP angle of 0° would indicate the sentiment of user and peer comments are in-phase (i.e., peaks in user and peer sentiment occur simultaneously), while an RP of 180° represents an anti-phase relationship (i.e., peaks in user sentiment occur a half-cycle prior to peaks in peer sentiment; see Fig. 2, panel C for an example anti-phase relationship). Values between 0° and 180° correspond to other intermediate lags in leading or following. In other words, the RP angle describes whether fluctuations in user sentiment follow fluctuations in peer sentiment (or vice-versa) as well as the time lags of those relationships. One method to describe RP would be to average the RP angles across all significant regions; however, overaggregation can lead to misleading results if the leader-follower dynamics differ by frequency or time. Considering this, we opted to extract a singular RP angle from each user's highest point of common power between periods 8 and 16.

4. Results

4.1. Descriptives

The final sample contained 246 users from /r/Conspiracy with an average of 341.56 interactions per user. User comments were generally neutral, contained a low percentage of conflict words, and had an average length of 43.97 words per comment. Peer comments that users interacted with tended to be more verbose, consisted of a slightly higher percentage of conflict words, and had an increased amount of negative sentiment compared to their own. Full descriptive statistics and correlations are provided in Table 1.

4.2. RQ1. Do users demonstrate sentiment coupling with the extreme subreddit at the group level?

First, we examined whether users of the conspiracy subreddit coupled their sentiment to align with their peers at the community level. We established a randomized baseline by shuffling each user's time series to remove time dependence and repeated the cross-wavelet analysis described previously. We then compared the common power of each user's real and shuffled time series across all frequencies using a paired sample *t*-test. We found that common power for the observed data ($M = 2.53$, $SD = 0.30$) was significantly higher than the shuffled baseline ($M = 2.41$, $SD = 0.37$, $t(245) = 4.19$, $p < .001$, $d = 0.38$). Based on the estimated

Table 1
Descriptive statistics for user interactions in /r/conspiracy

	M (SD)	1	2	3	4	5	6	7
1. % Conflict Words (User)	.46 (.22)	—						
2. % Conflict Words (Peers)	.55 (.21)	.30*	—					
3. Word Length (User)	43.97 (67.63)	.03	−.05	—				
4. Word Length (Peers)	61.52 (159.16)	−.16*	−.08	.31*	—			
5. Sentiment (User)	−.02 (.10)	−.31*	−.28*	−.03	.23*	—		
6. Sentiment (Peers)	−.07 (.06)	−.22*	−.40*	.02	.00	.35*	—	
7. Common Power	2.52 (.30)	.16*	−.09	.20*	.12	.05	.02	—

Note. Pearson correlation coefficients are provided in the numbered columns. * indicates $p < .05$.

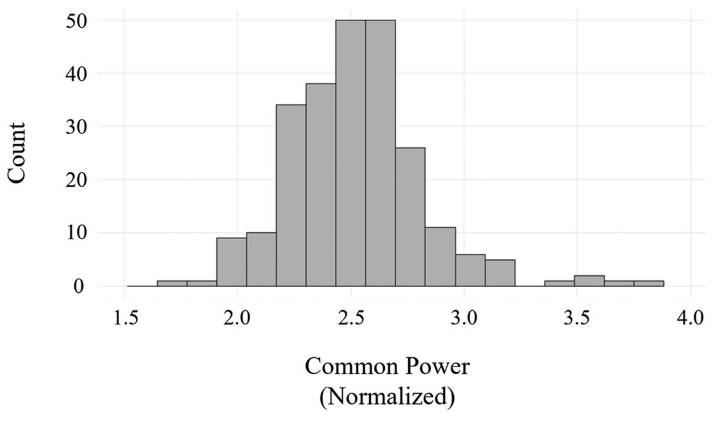

Fig. 3. Histogram showing the distribution of average common power between each user's comments and the comments of their peers.

effect size, there appeared to be a small amount of sentiment coupling at the group-average level. Fig. 3 illustrates the distribution of common power, with some users exhibiting substantial amounts of coupling, while others little-to-none.

4.3. RQ2–3. How does sentiment coupling correspond to engagement levels, conflict language, and longitudinal changes in sentiment?

Next, we investigated how individual differences in sentiment coupling corresponded to changes in user engagement within /r/Conspiracy and changes in sentiment over time. We fit a linear regression model with average common power as the outcome (see Table 2). Total comments and average comment word length were entered into the model as correlates of community engagement. To estimate changes in sentiment over time, we used regression to measure the change of sentiment for each user and peer comment as a function of time and included those standardized coefficients as predictors in the current model. We found greater sentiment coupling was associated with increased total comment count in the extremist

Table 2
Regression model predicting average common power by comment count, language sentiment, usage of conflict words, and word count. P-values lower than .05 are reported in bold.

	Average Common Power		
Predictors	*β*	*Standardized CI*	*p*
(Intercept)	0.00	–0.11 to 0.11	**<.001**
Comment Count	0.29	0.17–0.40	**<.001**
% Conflict Words (Peers)	–0.13	–0.25 to –0.02	**.04**
% Conflict Words (User)	0.19	0.07–0.31	**.002**
Sentiment Trajectory (Peers)	–0.16	–0.28 to –0.04	**.01**
Sentiment Trajectory (User)	–0.05	–0.17 to 0.07	.41
Word Count	0.17	0.05–0.28	**.005**
Observations	246		
R^2/R^2 adjusted	0.20/0.18		

community ($\beta = .29, p < .001$) and those comments were more verbose ($\beta = .17, p = .005$). Increased sentiment coupling was also associated with more frequent usage of conflict words ($\beta = .19, p = .002$) but a lower percentage of conflict words in peer comments ($\beta = -.13, p = .038$). Finally, we found that users with greater sentiment coupling were selectively engaging with increasingly negative peer comments over time ($\beta = -.16, p = .01$).

4.4. RQ4. To what extent do leader-follower dynamics between user and online community vary based on engagement with extremist communities?

Finally, we used the RP angle to determine the degree to which there were leader-follower dynamics between users and peers. First, we binned the RP angles into 20° bins and used a chi-square test to compare the distribution of RP angles in each user's real versus shuffled baselines. We found the distribution of RP angles differed significantly between the two conditions ($X^2 = 28.23$, $p < .001$) with the real data tending to reflect in-phase (i.e., around 0°) relationships and the shuffled data exhibiting a more uniform distribution (Fig. 4). To examine whether RP differed by user engagement with the subreddit, we compared users with high versus low comment counts. We split users into two groups based on median comment count and applied a chi-square test to compare their RP angles. There were no significant differences between users with high versus low comment counts ($p > .05$).

5. Discussion

Even though radicalization is an inherently time-evolving process (Alimi et al., 2012), little research has examined the dynamics of radicalization (Fernandez et al., 2019). We took

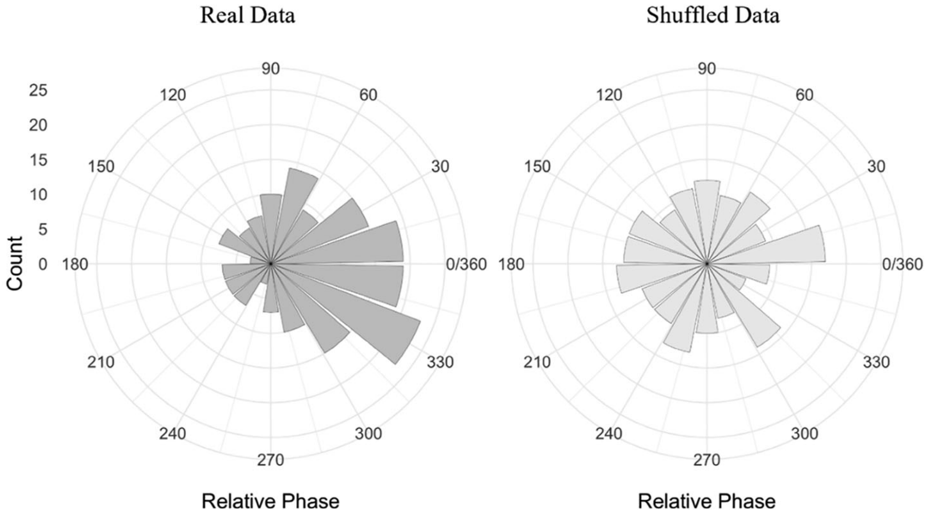

Fig. 4. Distribution of RP angle from moment of peak common power in each user's real (right panel) and shuffled (left panel) data. Points within the quadrant of the unit circle near 0° (e.g., between 45° and −45°) are essentially in-phase and exhibit simultaneous peaks in user and peer sentiment. Points in the quadrant surrounding 180° (e.g., between 135° and 225°) can be thought of as anti-phase, meaning the peaks in user sentiment occurred during the low points of their peers' sentiment. Finally, points near 90° suggest that the community is leading changes in peer sentiment, while points near 270° suggest that users are leading changes in community sentiment.

a step toward doing so, with the aim of elucidating processes underlying social media user's sentiment coupling with an extremist online community centered on conspiratorial beliefs, as well as associations between coupling and indices of user extremism (i.e., degree of engagement with the conspiratorial community and conflict language). Our findings show that users engaged with extremist online communities are generally attracted to more verbose and negative content. At a group level, users of the conspiratorial subreddit adjusted the sentiment of their comments to match that of their peers. Moreover, those who were more attuned with their peers demonstrated increased levels of engagement with the extremist subreddit, interacted with more conflict-oriented content, and became increasingly negative over time.

Taken together, dynamics of user-to-community sentiment coupling reveal processes of radicalization that follow from engagement with extremist material, where conflict is a primary situational factor associated with developing extremism (McCauley & Moskalenko, 2008; Odag et al, 2019). The relationship uncovered between conflict language and sentiment coupling is notable, as radicalization theory highlights how loneliness can be transformed into anger through engrossment in echo chambers (Regehr, 2022). Moreover, conflict management strategies that help individuals think critically about their moral conflicts with other groups have been proposed as a deradicalization method (Feddes, Mann, & Doosje, 2015). The current findings also suggest new possibilities for deradicalization using interventions aimed at disrupting the dynamics underlying social media radicalization, for example, by increasing the visibility of more positive and lower-conflict material. Although more research is needed to test potential mitigating strategies, understanding the dynamics of radicalization opens new avenues for advancements in deradicalization.

Our study provides a brief demonstration of applying the dynamical hypothesis to understanding collective human behavior in social media networks. Consistent with the dynamical hypothesis, radicalization is both defined as dynamic and aptly studied in terms of its state-space evolution, where dynamical methods empirically support previously proposed theory (van Gelder, 1998). Consistent with the original Watt governor problem used to illustrate the dynamical hypothesis (van Gelder, 1998), our conceptualization of radicalization did not require appeal to representations or computations. Instead, the dynamical approach emphasizes *coupling* between user and community within a *system* of variables that change interdependently (see Fig. 1) and based on system *parameters* including linguistic features. Also contrary to a mainstream computational approach to cognitive science that often shelves problems of embeddedness, the dynamical hypothesis allows social cognition to be contextualized and understood within the environment that shapes information-gathering and resulting beliefs (van Gelder, 1998). With the dynamical hypothesis previously applied to nonhuman animal group behavior such as bird flocks and schools of fish (Olfati-Saber, 2006), as well as a variety of human psychological and cognitive phenomenon, it follows that the dynamical hypothesis can and should be applied to the new frontier of human behavior arising within complex social media networks.

While this study focuses on multiresolution dynamic analysis of user engagement with an extremist online community, the approach taken could be applied to study aspects of social dynamics in many different domains. Future studies on social media radicalization could investigate how extremist political ideologies and attitudes toward controversial topics are related to changes in comment sentiment and user engagement at both the user and community levels. The methods introduced here are also applicable to topics outside of social media radicalization, such as exploring the spread of misinformation in online communities or examining the influence of online peer support on participant attitudes. Beyond measuring affective coupling with cross-spectrum wavelet analysis, future research investigating social media interactions through the dynamical hypothesis could analyze the qualitative state of comments, how they evolve, and whether online communities converge to stable attractors. Social media provides convenient access to high-resolution social interaction data; however, there are emerging opportunities to explore collective behaviors in real-world contexts as well. Given the increasing accessibility of sensor-based measures, the dynamical hypothesis can be leveraged to understand how groups of people interact outside of social media, with the potential to enrich human−computer interaction, social psychology, and urban analytics research (Amon, 2015). These are just a few examples of how the dynamical hypothesis can deepen our understanding of collective human behavior.

6. Limitations and future directions

There are several points to make about our applications of wavelet analysis on social media data. First, although VADER has been referred to as the gold standard (Bonta & Janardhan,

2019) for lexicon-based sentiment analysis on social media, it does not have perfect accuracy. The benefit of VADER was the ability to code large amounts of text, but our method could be applied using machine learning for sentiment analysis or with objective metrics like word count. Another consideration should be made for how we modeled conversations on the forum as a series of inputs (peer comments) and responses (user comments), emphasizing the iterative influence of social interactions. Modeling in this way can make it difficult to predict the effect of missing data. We attempted to overcome this by relying on an archival database to retrieve comments removed since the time of their posting. Other approaches could be applied that mitigate against the effects of missing data, such as by averaging sentiment by day to examine global trends.

References

Agarwal, S., & Sureka, A. (2015). Applying social media intelligence for predicting and identifying on-line radicalization and civil unrest oriented threats. *arXiv preprint*. https://doi.org/10.48550/arXiv.1511.06858

Aguilera, M., Morer, I., Barandiaran, X., & Bedia, M. (2013). Quantifying political self-organization in social media. Fractal patterns in the Spanish 15M movement on Twitter. In *ECAL 2013: The 12th European Conference on Artificial Life* (pp. 395–402). MIT Press. https://doi.org/10.7551/978-0-262-31709-2-ch057

Akram, W., & Kumar, R. (2017). A study on positive and negative effects of social media on society. *International Journal of Computer Sciences and Engineering*, *5*(10), 351–354. https://doi.org/10.26438/ijcse/v5i10.351354

Alfano, M. (2013). *Character as moral fiction*. Cambridge University Press.

Alfano, M., Carter, J., & Cheong, M. (2018). Technological seduction and self-radicalization. *Journal of the American Philosophical Association*, *4*(3), 298–322. https://doi.org/10.1017/apa.2018.27

Alimi, E., Bosi, L., & Demetriou, C. (2012). Relational dynamics and processes of radicalization: A comparative framework. *Mobilization: An International Quarterly*, *17*(1), 7–26. https://doi.org/10.17813/maiq.17.1.u7rw348t8200174h

Allington, D., Duffy, B., Wessely, S., Dhavan, N., & Rubin, J. (2021). Health-protective behaviour, social media usage and conspiracy belief during the COVID-19 public health emergency. *Psychological Medicine*, *51*(10), 1763–1769. https://doi.org/10.1017/S003329172000224X

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*(2), 256. https://psycnet.apa.org/doi/10.1037/0033-2909.111.2.256

Amon, M. J. (2015). Visual attention in mixed-gender groups. *Frontiers in Psychology: Personality and Social Psychology*, *5*(1569), 1–10. https://doi.org/10.3389/fpsyg.2014.01569

Anderson, K. E. (2015). Ask me anything: What is Reddit? *Library Hi Tech News*, *32*(5), 8–11. https://doi.org/10.1108/LHTN-03-2015-0018

Ayanian, A. H., Böckler, N., Doosje, B., & Zick, A. (2018). Processes of radicalization and polarization in the context of transnational Islamist terrorism. *International Journal of Conflict and Violence*, *12*, a648. https://doi.org/10.4119/ijcv-3098

Aydin, M. N., & Perdahci, N. Z. (2019). Dynamic network analysis of online interactive platform. *Information Systems Frontiers*, *21*, 229–240. https://doi.org/10.1007/s10796-017-9740-8

Bak-Coleman, J. B., Alfano, M., Barfuss, W., Bergstrom, C. T., Centeno, M. A., Couzin, I. D., Donges, F.D., Galesic, M., Gersick, A. S., Jacquet, J., Kao, A. B., Moran, R. E., Romanczuk, P., Rubenstein, D. I., Tombak, K. J., Van Bavel, J. J., & Weber, E. U. (2021). Stewardship of global collective behavior. *Proceedings of the National Academy of Sciences*, *118* (27), 1–10. https://doi.org/10.1073/pnas.2025764118

Baumann, F., Lorenz-Spreen, P., Sokolov, I. M., & Starnini, M. (2020). Modeling echo chambers and polarization dynamics in social networks. *Physical Review Letters*, *124*(4), 1–6. https://doi.org/10.1103/PhysRevLett.124.048301

Benoit, S. L., & Mauldin, R. F. (2021). The “anti-vax” movement: A quantitative report on vaccine beliefs and knowledge across social media. *BMC Public Health*, *21*(1), 1–11. https://doi.org/10.1186/s12889-021-12114-8

Bertram, L. (2016). Terrorism, the Internet and the social media advantage: Exploring how terrorist organizations exploit aspects of the Internet, social media and how these same platforms could be used to counter-violent extremism. *Journal for Deradicalization*, *17*, 225–252.

Bjelopera, J. P. (2013). *American jihadist terrorism: Combating a complex threat*. DIANE Publishing.

Blonder, B., Wey, T. W., Dornhaus, A., James, R., & Sih, A. (2012). Temporal dynamics and network analysis. *Methods in Ecology and Evolution*, *3*(6), 958–972. https://doi.org/10.1111/j.2041-210X.2012.00236.x

Bonta, V., & Janardhan, N. K. N. (2019). A comprehensive study on lexicon based approaches for sentiment analysis. *Asian Journal of Computer Science and Technology*, *8*(S2), 1–6. https://doi.org/10.51983/ajcst-2019.8.s2.2037

Borum, R. (2011). Radicalization into violent extremism I: A review of social science theories. *Journal of Strategic Security*, *4*(4), 7–36. http://doi.org/10.5038/1944-0472.4.4.1

Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). *The development and psychometric properties of LIWC-22*. Austin, TX: University of Texas at Austin.

Bródka, P., Kazienko, P., Musiał, K., & Skibicki, K. (2012). Analysis of neighbourhoods in multi-layered dynamic social networks. *International Journal of Computational Intelligence Systems*, *5*(3), 582–596. https://doi.org/10.1080/18756891.2012.696922

Brown, C., & Liebovitch, L. (2010). *Fractal analysis* (Vol. 165). Sage.

Cassell, J., & Tversky, D. (2005). The language of online intercultural community formation. *Journal of Computer-Mediated Communication*, *10*(2), 1–11. https://doi.org/10.1111/j.1083-6101.2005.tb00239.x

Cilluffo, F. J., Cardash, S. L., & Whitehead, A. J. (2006). Radicalization: Behind bars and beyond borders. *Brown Journal of World Affairs*, *13*, 113.

Crossett, C., & Spitaletta, J. (2010). Radicalization: Relevant psychological and sociological concepts. John Hopkins University.

Dalgaard-Nielsen, A. (2010). Violent radicalization in Europe: What we know and what we do not know. *Studies in Conflict & Terrorism*, *33*(9), 797–814. https://doi.org/10.1080/1057610X.2010.501423

Decety, J., Pape, R., & Workman, C. I. (2018). A multilevel social neuroscience perspective on radicalization and terrorism. *Social Neuroscience*, *13*(5), 511–529. https://doi.org/10.1080/17470919.2017.1400462

del Fresno García, M., Daly, A. J., & Segado Sanchez-Cabezudo, S. (2016). Identifying the new influences in the Internet era: Social media and social network analysis. *Revista Española de Investigaciones Sociológicas*, *153*, 23–42. https://doi.org/10.5477/cis/reis.153.23

Doris, J. M. (2015). *Talking to our selves: Reflection, ignorance, and agency*. OUP Oxford.

Eckert, P., & Rickford, J. R. (2001). *Style and sociolinguistic variation*. Cambridge University Press.

Farine, D. R. (2018). When to choose dynamic vs. static social network analysis. *Journal of Animal Ecology*, *87*(1), 128–138. https://doi.org/10.1111/1365-2656.12764

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Favela, L. H. (2020). Dynamical systems theory in cognitive science and neuroscience. *Philosophy Compass*, *15*(8), e12695. https://doi.org/10.1111/phc3.12695

Feddes, A. R., Mann, L., & Doosje, B. (2015). Increasing self-esteem and empathy to prevent violent radicalization: A longitudinal quantitative evaluation of a resilience training focused on adolescents with a dual identity. *Journal of Applied Social Psychology*, *45*(7), 400–411. https://doi.org/10.1111/jasp.12307

Fernandez, M., Gonzalez-Pardo, A., & Alani, H. (2019). Radicalisation influence in social media. *Journal of Web Science*, *6*, 1–15. https://doi.org/10.34962/jws-70

Giles, H., Taylor, D. M., & Bourhis, R. (1973). Towards a theory of interpersonal accommodation through language: Some Canadian data[1]. *Language in Society*, *2*(2), 177–192. https://doi.org/10.1017/S0047404500000701

Gill, P. (2007). A multi-dimensional approach to suicide bombing. *International Journal of Conflict and Violence*, *1*(2), 142–159. https://ssrn.com/abstract=1030045

Gouhier, T. C., Grinsted, A., & Simko, V. (2018). R package biwavelet: Conduct univariate and bivariate wavelet analyses. *Version 0.20*, *19*. https://github.com/tgouhier/biwavelet

Hafez, M., & Mullins, C. (2015). The radicalization puzzle: A theoretical synthesis of empirical approaches to homegrown extremism. *Studies in Conflict & Terrorism*, *38*(11), 958–975. https://doi.org/10.1080/1057610X.2015.1051375

Hardy, K. (2018). Comparing theories of radicalisation with countering violent extremism policy. *Journal for Deradicalization*, *15*, 76–110.

Heijmans, H. J., & Goutsias, J. (2000). Nonlinear multiresolution signal decomposition schemes. II. Morphological wavelets. *IEEE Transactions on Image Processing*, *9*(11), 1897–1913.

Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media* (pp. 216–225). https://doi.org/10.1609/icwsm.v8i1.14550

Issartel, J., Bardainne, T., Gaillot, P., & Marin, L. (2015). The relevance of the cross-wavelet transform in the analysis of human interaction–A tutorial. *Frontiers in Psychology*, *5*, 1566. https://doi.org/10.3389/fpsyg.2014.01566

Johnson, S., Faraj, S., & Kudaravalli, S. (2014). Emergence of power laws in online communities. *MisQuarterly*, *38*(3), 795–808. https://doi.org/10.25300/misq/2014/38.3.08

Koehler, D. (2014). The radical online: Individual radicalization processes and the role of the Internet. *Journal for Deradicalization*, *1*, 116–134.

Lampe, C., & Johnston, E. (2005). Follow the (slash) dot: Effects of feedback on new members in an online community. In *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work* (pp. 11–20). https://doi.org/10.1145/1099203.1099206

Luo, Q., & Zhong, D. (2015). Using social network analysis to explain communication characteristics of travel-related electronic word-of-mouth on social networking sites. *Tourism Management*, *46*, 274–282. https://doi.org/10.1016/j.tourman.2014.07.007

Marzoratti, A., & Evans, T. M. (2022). Measurement of interpersonal physiological synchrony in dyads: A review of timing parameters used in the literature. *Cognitive, Affective, & Behavioral Neuroscience*, *22*(6), 1215–1230. https://doi.org/10.3758/s13415-022-01011-1

Matamoros-Fernández, A., & Farkas, J. (2021). Racism, hate speech, and social media: A systematic review and critique. *Television & New Media*, *22*(2), 205–224. https://doi.org/10.1177/1527476420982230

McCauley, C., & Moskalenko, S. (2008). Mechanisms of political radicalization: Pathways toward terrorism. *Terrorism and Political Violence*, *20*(3), 415–433. https://doi.org/10.1080/09546550802073367

Medvedev, A. N., Lambiotte, R., & Delvenne, J. C. (2019). The anatomy of Reddit: An overview of academic research. In Ghanbarnejad, F., Saha Roy, R., Karimi, F., Delvenne, JC., Mitra, B. (eds) *Dynamics on and of complex networks III: Machine learning and statistical physics approaches* (pp. 183–204). https://doi.org/10.1007/978-3-030-14683-2_9

Moghaddam, F. M. (2005). The staircase to terrorism: A psychological exploration. *American Psychologist*, *60*(2), 161. https://doi/10.1037/0003-066X.60.2.161

Mondal, M., Silva, L. A., & Benevenuto, F. (2017). A measurement study of hate speech in social media. In *Proceedings of the 28th ACM Conference on Hypertext and Social Media* (pp. 85–94). http://doi.org/10.1145/3078714.3078723

Muchnik, L., Pei, S., Parra, L. C., Reis, S. D., Andrade Jr, J. S., Havlin, S., & Makse, H. A. (2013). Origins of power-law degree distribution in the heterogeneity of human activity in social networks. *Scientific Reports*, *3*(1), 1783. https://doi.org/10.1038/srep01783

Necaise, A., Williams, A., Vrzáková, H., & Amon, M. J. (2021). Regularity versus novelty of users' multimodal comment patterns and dynamics as markers of social media radicalization. In *Proceedings of the 32nd ACM Conference on Hypertext and Social Media* (pp. 237–243). https://doi.org/10.1145/3465336.3475095

Nguyen, D., & Rose, C. (2011). Language use as a reflection of socialization in online communities. In *Proceedings of the Workshop on Language in Social Media* (pp. 76–85).

Odag, Ö., Leiser, A., & Boehnke, K. (2019). Reviewing the role of the Internet in radicalization processes. *Journal for Deradicalization*, *21*, 261–300.

Oh, G. (2022). Multifractal analysis of social media use in financial markets. *Journal of the Korean Physical Society*, *80*(6), 526–532. https://doi.org/10.1007/s40042-022-00448-4

Olfati-Saber, R. (2006). Flocking for multi-agent dynamic systems: Algorithms and theory. *IEEE Transactions on Automatic Control*, *51*(3), 401–420. https://doi.org/10.1109/tac.2005.864190

Panetta, G. (2022). *Twitter reportedly won't use an algorithm to crack down on white supremacists because some GOP politicians could end up getting barred too*. Business Insider. https://www.businessinsider.com/twitter-algorithm-crackdown-white-supremacy-gop-politicians-report-2019-4

Panzarasa, P., Opsahl, T., & Carley, K. M. (2009). Patterns and dynamics of users' behavior and interaction: Network analysis of an online community. *Journal of the American Society for Information Science and Technology*, *60*(5), 911–932. https://doi.org/10.1002/asi.21015

Parrott, L. (2002). Complexity and the limits of ecological engineering. *Transactions of the ASAE*, *45*(5), 1697–1702. https://doi.org/10.13031/2013.11032

Regehr, K. (2022). In (cel) doctrination: How technologically facilitated misogyny moves violence off screens and on to streets. *New Media & Society*, *24*(1), 138–155. https://doi.org/10.1177/1461444820959019

Rousseau, C., Hassan, G., & Oulhote, Y. (2017). And if there were another way out? Questioning the prevalent radicalization models. *Canadian Journal of Public Health*, *108*, e633–e635. https://doi.org/10.17269/CJPH.108.6233

Sageman, M. (2004). *Understanding terror networks*. University of Pennsylvania Press.

Salojärvi, E., Rantanen, M., Nieminen, E., Juote, A., & Hanhela, H. (2020). The 'incel' phenomenon in the digital era–How echo chambers have fueled the incel movement. https://philarchive.org/archive/SALTIP

Samory, M., & Mitra, T. (2018). 'The government spies using our webcams': The language of conspiracy theories in online discussions. In *Proceedings of the ACM on Human−Computer Interaction*. https://doi.org/10.1145/3274421

Schmid, A. P. (2013). Radicalisation, de-radicalisation, counter-radicalisation: A conceptual discussion and literature review. *ICCT Research Paper*. https://doi.org/10.19165/2013.1.02

Sharma, E., & De Choudhury, M. (2018). Mental health support and its relationship to linguistic accommodation in online communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3173574.3174215

Silber, M. D., Bhatt, A., & Analysts, S. I. (2007). *Radicalization in the West: The homegrown threat*. New York: Police Department.

Sinai, J. (2012). Radicalisation into extremism and terrorism. *Intelligencer: Journal of US Intelligence Studies*, *19*(2), 21–25.

Smith, B. L., Snow, D. A., Fitzpatrick, K., Damphousse, K. R., Roberts, P., Tan, A., & Klein, B. (2016). Identity and framing theory, precursor activity, and the radicalization process. *National Criminal Justice Reference Service*.

Smith, L. G., Blackwood, L., & Thomas, E. F. (2020). The need to refocus on the group as the site of radicalization. *Perspectives on Psychological Science*, *15*(2), 327–352. https://doi.org/10.1177/1745691619885870

Strogatz, S. (2004). *Sync: The emerging science of spontaneous order*. Hyperion.

Tehrani, P. M., Manap, N. A., & Taji, H. (2013). Cyber terrorism challenges: The need for a global response to a multi-jurisdictional crime. *Computer Law & Security Review*, *29*(3), 207–215. https://doi.org/10.1016/j.clsr.2013.03.011

Torok, R. (2013). Developing an explanatory model for the process of online radicalisation and terrorism. *Security Informatics*, *2*, 1–10. https://doi.org/10.1186/2190-8532-2-6

Tsugawa, S., & Ohsaki, H. (2014). Emergence of fractals in social networks: Analysis of community structure and interaction locality. In *2014 IEEE 38th Annual Computer Software and Applications Conference* (pp. 568–575). IEEE. https://doi.org/10.1109/compsac.2014.80

Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., & Nyhan, B. (2018). *Social media, political polarization, and political disinformation: A review of the scientific literature*. https://doi.org/10.2139/ssrn.3144139 van Gelder, T. (1995). What might cognition be, if not computation? *Journal of Philosophy*, *92*(7), 345–381. https://doi.org/10.2307/2941061

van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, *21*(5), 615–628. https://doi.org/10.1017/S0140525×98001733

van Raemdonck, N. (2019). The echo chamber of anti-vaccination conspiracies: Mechanisms of radicalization on Facebook and Reddit. Institute for Policy, Advocacy and Governance (IPAG) Knowledge Series, Forthcoming.

Von Behr, I. (2013). Radicalisation in the digital era: The use of the Internet in 15 cases of terrorism and extremism. https://www.rand.org/content/dam/rand/pubs/research_reports/RR400/RR453/RAND_RR453.pdf

Vrzáková, H., Amon, M. J., & D'Mello, S. K. (2021). Looking for a deal! Visual social attention during negotiations via mixed media videoconferencing. In *Proceedings of the Association for Computing Machinery: Computer Supported Cooperative Work (CSCW)*. https://doi.org/10.1145/3434169

Wiktorowicz, Q. (2005). *Radical Islam rising: Muslim extremism in the West*. Rowman & Littlefield Publishers.

Wong, D. B. (2006). *Natural moralities: A defense of pluralistic relativism*. Oxford University Press.