Inside the echo chamber: Linguistic underpinnings of misinformation on Twitter

Xinyu Wang
The Pennsylvania State University
University Park, PA, USA
xzw5184@psu.edu

Jiayi Li
The Pennsylvania State University
University Park, PA, USA
jpl6207@psu.edu

Sarah Rajtmajer
The Pennsylvania State University
University Park, PA, USA
smr48@psu.edu

ABSTRACT

Social media users drive the spread of misinformation online by sharing posts that include erroneous information or commenting on controversial topics with unsubstantiated arguments often in earnest. Work on echo chambers has suggested that users' perspectives are reinforced through repeated interactions with like-minded peers, promoted by homophily and bias in information diffusion. Building on long-standing interest in the social bases of language and linguistic underpinnings of social behavior, this work explores how conversations around misinformation are mediated through language use. We compare a number of linguistic measures, e.g., in-/out-group cues, readability, and discourse connectives, within and across topics of conversation and user communities. Our findings reveal increased presence of group identity signals and processing fluency within echo chambers during discussions of misinformation. We discuss the specific character of these broader trends across topics and examine contextual influences.

CCS CONCEPTS

• **Information systems** → **Social networks**; • **Applied computing** → *Sociology*; *Psychology*; • **Computing methodologies** → **Discourse, dialogue and pragmatics**.

KEYWORDS

Social Media, Misinformation, Echo Chamber, Social Networks, Socio-linguistics, Group Identity, Processing Fluency Theory

1 INTRODUCTION

In recent years, Online Social Networks (OSNs) have developed into a primary source of news for a growing majority of the US population [39]. As OSNs have expanded, so have concerns about the role of social media in promoting misinformation and deepening social fissures [4, 81], the consequences of which were particularly tangible during the COVID-19 pandemic and response [1, 13, 16].

User interactions in OSNs are scaffolded by platforms' algorithms and moderation practices. Optimized primarily for engagement, these algorithms promote interaction with like-minded peers and isolate users from disagreeable content. While this can create safe space for individuals to share opinions and find validation, it can also drive polarization [8, 19, 59]. Some have suggested that the effects of so-called *filter bubbles* and *echo chambers* have been overstated or oversimplified and have argued for a more nuanced characterization based on, e.g., ideology or topic of discussion [9, 15, 33]. Irrespective, these phenomena speak to the broader, well-studied phenomenon of homophily in social group formation [23, 45]. In the online setting, user-generated content facilitates aggregation around shared interests and worldviews, which in turn serves to support the deepening and reinforcement of social ties [25].

At the heart of OSN communities is discourse rich with linguistic signals of individual and relational identities. That is, users of language perform their identities within the uses of language [11, 79, 80]. This is both heightened and complicated by the affordances of social media [67]. Identities are partially performed through alignment with communities, opinions, and cultural issues [67]. Online, this can be enacted as direct engagement and participation in groups, or through indirect interactions, e.g., hashtagging, sharing memes [77, 79], or even simple repetition [62].

Several schools of thought have driven work to understand the co-evolution of language, social ties, and the environment [10, 32, 42, 55, 61]. Languages adapt to the constraints and biases of their learners, which include social, physical and technological context [24, 43]. Foundational work in sociolinguistics, for example, has studied the ways in which language change correlates with social network dynamics [12, 46]. Shared linguistic features can enhance so-called *processing fluency*, making communications more effective and reinforcing community bonds [17, 48]. Within the landscape of social media, both notions of community and language practices are flexible and interactively constructed [67].

Our work is a study of the relationships between social media users' linguistic patterns and the community structures in which they are embedded. We have selected misinformation as our point of inquiry because prior work has suggested that misinformation spreads particularly quickly [25, 36, 37, 63, 75] and because a shared vocabulary has emerged around these topics. We are motivated by two primary research questions.

RQ1: Is there an observable increase in users' signaling of group identity through publicly shared content within echo chambers? Is this phenomenon more pronounced during discussions that involve topics related to misinformation?

RQ2: Is there heightened processing fluency in publicly shared content within echo chambers? Is this phenomenon more pronounced during discussions that involve topics related to misinformation?

To address these questions, we identify echo chambers from user interaction networks and study their linguistic characteristics. Our analysis suggests that while there are some shared linguistic patterns among certain topics, linguistic characteristics observed within echo chambers do not exhibit uniformity across all discussions of misinformation. Rather, these patterns appear to be contingent upon contextual factors (e.g., political nature of the topic). This interplay between linguistic uniformity and contextual particularity constitutes is at the core of our work, contributing to deeper understanding of the linguistic underpinnings of echo chambers, particularly in the context of misinformation-related topics.

2 RELATED WORK

2.1 Misinformation dynamics

The extensive presence of user-generated content on online social media platforms enables the formation of communities that unite around shared beliefs and narratives. Echo chambers, where beliefs are amplified and reinforced by repeated internal validation, play a role in the dynamics of misinformation within online platforms [25]. The architecture of modern social media algorithms, which prioritizes user engagement, can inadvertently foster echo-chambered communities [82]. Within these environments, misinformation thrives and evolves, taking on new narratives that further reinforce the community's existing beliefs [75].

2.2 Social identity formation

Social identity formation emerges as a critical process underlying group dynamics and individual affiliations in social networks. In this study, social identities are performed and measured through a combination of explicit identity cues, as explained by social identity theory, and implicit identity cues via ambient affiliation.

2.2.1 Social identity theory. Social identity theory posits that individuals tend to see themselves and others in relation to groups [66, 68, 69]. When a group identity becomes prominent, individuals develop their self-concept around paradigmatic ingroup characteristics [14]. They are motivated to differentiate between their ingroup and relevant outgroups [40]. Social identity theory has been engaged to describe political group formation in online social communities, e.g., [31, 71].

Recent work [40] looked at discussions containing the term "fake news" on Twitter and found an uptrend in the use of identity language, as measured by increased frequency of group pronouns. They also found that group pronouns were likely to collocate with words boosting ingroup messaging. These findings dovetail with concerns about polarization. Rathje et al. [54] found out-group animosity to be a predictor of engagement on Facebook and Twitter, due in part to algorithmic designs favoring messages containing outgroup antagonism. Our work operationalizes linguistic indicators of social identity as in-/out-group reference language.

2.2.2 Ambient affiliation. So-called *ambient affiliation* through hashtagging has evolved as a ubiquitous mechanism for social media users to associate and communicate in digital environments, especially on OSNs. It allows for the spontaneous aggregation of like-minded individuals around shared interests, experiences, or sentiments, often leading to the establishment of virtual communities or "micro-groupings" [77]. More than just categorization, the act of hashtagging serves as an instrument of social group identity formation. By selecting specific hashtags, users signal an allegiance or a stance, fostering both inclusion within communities and distinction from others. Consistent patterns of hashtagging behaviors can solidify these group identities, creating online communities unified by shared narratives and ideologies [78, 79].

2.3 Processing fluency

Processing fluency theory suggests that the subjective experience of ease with which people process information influences people's judgments of that information [56, 58]. People are more likely to judge information favorably if they are able to digest it easily. This effect has been shown robust over diverse experimental settings and a variety of formulations of fluency, e.g., conceptual, perceptual, linguistic [5, 47]. Central to this theory is the notion that enhanced fluency can be interpreted in multiple ways: it might signify the benefits of a simplified message; the role of priming and familiarity; or, the importance of explanation and elaboration.

2.3.1 Simplicity bolsters information processing. Recent work has shown that easy-to-read social media posts facilitate processing fluency, resulting in more likes, comments, and shares [7, 48]. Likewise, repeated assertions gain a valence of credibility [34, 57, 73]. The so-called repetition-induced truth effect explains one mechanism by which individuals come to trust obviously false information, a concern magnified by echo chambers [51, 74]. In the context of today's misinformation landscape, scholars have noted that conspiracism is increasingly expressed as bare assertion [62]. New conspiracists repeat simple statements, signaling solidarity with one another through retweets, likes, and shares [62].

2.3.2 Elaboration bolsters information processing. In parallel, there exists a body of psychological theories suggesting that explanations bolster credibility of a claim, namely, the idea of *rational persuasion*. Information-processing theories of human cognition have explored the ways in which we engage in the reception, encoding, storage, and retrieval of information [6, 64]. Systematic information processing is associated with explanations, whereby individuals engage in effortful scrutiny and comparison of information [18, 70, 72]. Relatedly, the Elaboration Likelihood Model (ELM) [52, 53] suggests that when people process information deeply, they are more likely to be influenced by it. Prior studies on the QAnon conspiracy have demonstrated the efficacy of this technique. Individuals aligned with these beliefs use archival practices to enable the construction and diffusion of knowledge within their community. Members are urged to engage in critical thinking and do extensive research using key terms promoted by the group [44].

3 DATA

We collected tweets using the Twitter Developer API and keywords associated with our target topics, namely: anti-vaccination; 2020 US election fraud; and, QAnon. We pulled author ID, tweet ID, and referenced tweets (type and ID) for interaction network construction

and tweet content for linguistic analyses. Tweets were collected using popular keywords and hashtags over date ranges selected for greatest topic relevance. With the exception of discussion surrounding the US election which has a shorter activity period and limited number of tweets, we collect approximately 1 million tweets per topic, adjusting date ranges accordingly. In addition, we collected discussions to serve as comparative for our misinformation-related discourse: the MeToo movement, Black Lives Matter movement (BLM), and #mondaymotivation. While not generally associated with misinformation, these have been topics of widespread social media discourse, representing very different degrees of social sensitivity and polarization.

Pre-processing. We retained only tweets containing interactions within the datasets, as well as original tweets without interactions, to ensure that all users in the social network have representative linguistic content. Dataset statistics are provided in Table 1. We removed mentions, #hashtags, URLs, RT signs, and emojis. For readability tests, we retained punctuation marks to allow sentence length calculation. Hashtags and URLs were extracted directly from the raw data. Tweet IDs and code have been made publicly available and are provided with detailed documentation including a Datasheet [30] and README[1].

4 PRELIMINARIES

We introduce key definitions and measures used throughout.

4.1 Linguistic metrics

Linguistic features were computed using and extending state-of-the-art open-source toolboxes for linguistic analyses. In- and out-group language was used to measure group identity formation. Group identity is expressed and sustained through pronouns; pronominal choices are not just categorical references but also reflect power dynamics and relationships [35, 50]. Specifically, "we-pronouns" are used to invoke ingroup solidarity while "they-pronouns" create distance with the outgroup, often implying the superiority of one's own community [14, 40]. We operationalize simplicity and elaboration as theoretical constructs through discourse connectives, big words, and readability metrics. Research has shown that textually less complex and highly readable content facilitates processing fluency, resulting in more social interactive behaviors [48]. While, discourse connectives may also improve online processing and reading comprehension in both narrative and expository texts [21, 83]. Evidence from these studies forms the basis for our approach.

4.1.1 In-/out-group identity language. In-/out-group language ratio [22] is used as an index of group identity. Exemplars of in-/out-group cues include: we, our, us, they, their, them, themselves.

4.1.2 Big words. Words containing seven or more letters serve as a rough indicator of language complexity. This metric is derived from The Linguistic Inquiry and Word Count (LIWC) [49] software.

4.1.3 Readability. We employ the Flesch reading ease score [28] as a measure of readability based on two factors: *average sentence length* and *average number of syllables in each word.*

---

[1]Github link: https://github.com/XinyuWang1998/Linguistic-Underpinning

4.1.4 Discourse connectives. We consider the incidence of discourse connectives [2, 22] as an indicator of explanation/elaboration. Our analyses consider three families of discourse connectives.
Logical connectives reflect logical explanation, e.g., because, consequently, hence;
Additive connectives reflect agreement and further explanation, e.g., and, besides, further;
Negative connectives, indicate an adversarial argument, e.g., but, alternatively, although.
Connectives ratios are calculated as the number of connective words from the given family divided by total number of words.

4.2 Network representation

We construct user interaction networks for each topic. Users are represented as nodes and interactions–replies, quotes, retweets–are represented as directed, weighted edges. Edge weights indicate number of interactions.

Echo chambers are recognized by coherent content and structural connectivity. Given that our interaction networks are topic-specific by virtue of our data collection, there is coherence in content by design. For structural identification of echo chamber members, we focus on *strongly connected components*, i.e., subgraphs in which there is a path from each vertex to every other, respecting directionality of the edges. This approach follows precedent set in prior literature, e.g., [20], and is particularly suitable for modeling echo chambers on platforms like Twitter where interactions are often non-reciprocal [38, 65]. Following, we identify users within strongly connected components ($n \geq 2$) as echo chamber members and their shared content as echo chamber tweets. User statistics are provided in Table 2.

4.3 Ambient affiliation

Sociolinguists have theorized that Twitter users signal group identity implicitly through hash-tagging. Hashtags serve as a semiotic identifier for the discourse community they establish, as well as a searchable record of discourse for the purpose of establishing new affiliates [77]. We measure **ambient affiliation** through the use of hash-tagging.

5 RQ1: GROUP IDENTITY FORMATION

RQ1 asks whether there are observable disparities in group identity between echo chambers and external groups within discussions surrounding misinformation-related topics. We assess this through two approaches: analyzing explicit signaling of group identity through the use of first and third-person pronouns; and, exploring implicit signaling of group identity through ambient affiliation.

5.1 Explicit identity signaling through in-/out-group cues

5.1.1 Methodology. We employ the zero-inflated beta regression model to examine the relationship between a user's membership to (presence within) an echo chamber and their use of group pronouns. The zero-inflated beta model is particularly suited for this analysis due to the nature of the data, which contains many zeros and is bounded between 0 and 1. We concentrate on the model's non-zero

Category	Topic	keywords (case-insensitive)	Time Frame	Total Tweets	Total Tweets after Filtering
Misinformation	US Election Fraud	voterfraud, discardedballots, cheatingdemocrats, stopvoterfraud, voterfraudbymail, voterfraudisreal, ballotharvasting, ballotvoterfraud	[2020-11-01, 2021-02-28]	256,036	219,872
	Anti-Vaccination	antivax, antivaxx, novaccine, novaccinemandates, pureblood, unvaxxed, unvaccinated, naturalimmunity, nomandatoryvaccines	[2021-11-15, 2021-11-30]	1,224,306	1,139,378
	QAnon	qanon, thegreatawakening, pizzagate, savethechildren, deepstate, thestormisuponus	[2019-11-01, 2020-03-31]	902,862	792,733
Hyper-polarized	MeToo Movement	metoo, hertoo, ustoo, metoocongress, timesup [76]	[2017-10-15, 2017-10-31]	845,472	828,856
	BLM Movement	blacklivesmatter,georgefloyd, icantbreathe, blm, justiceforfloyd, justiceforgeorgefloyd, georgefloydprotests, worldagainstracism, walkwithus, kneelwithus, blackouttuesday, voteouthate, nojusticenopeace, blackwomenmatter, blackgirlsmatter, theshowmustbepaused[27]	[2020-06-01 09:00AM, 2020-06-01 05:00PM]	994,663	968,270
Neutral	Monday Motivation	monday motivation	[2021-09-01, 2021-11-30]	896,350	892,096

Table 1: Dataset statistics

Topic	EC	EXT	Total
Election Fraud	470	97,510	97,980
Antivax	18,200	393,331	411,531
QAnon	8,934	233,166	242,100
MeToo	3,468	449,627	453,095
BLM	1,505	664,930	666,435
Monday Motivation	3,019	339,903	342,922

Table 2: Number of users/nodes engaged with each topic (EC=echo chamber; EXT=external).

component because it remains largely unaffected by extreme outliers, specifically instances where the content is very short thus not containing target vocabularies. The coefficient of the predictor variable in this regression, which represents echo chamber membership, provides an estimate of the log odds ratio of the change in identity measurement in response to echo chamber membership. If this coefficient is positive and statistically significant, it suggests that membership to an echo chamber is associated with higher log odds of using group pronouns. Recognizing that analyses with large datasets often yields small, less informative p-values, we employ a stratified random sampling method [29]. This involves selecting 1000 data points each from echo chamber and non-echo chamber data subsets (2000 total), ensuring a balanced representation of both categories. To address potential small sample biases and enhance the robustness of our findings, we conduct 1000 iterations of this regression analysis, obtaining median values for the coefficient, standard error, T-score, and p-value. Analyses without sampling are also provided, for completeness, alongside datasets and code through the project's Github page.

5.1.2 Findings. **We observe clear, heightened group identity development through the deliberate utilization of target plural pronouns in echo chambers within US election fraud and QAnon discussions.** See Table 3.

Following is a typical tweet from within a QAnon echo chamber that exemplifies pronounced group identity. After mentioning 50 names in a chain, the tweet reads:

> *Totally agreed. We Americans chart our own destiny The Individual is THE most protected of its rights by the Constitution we need to know this, and go forward confidently. DeepState, we will deal w/them in unity w/our President. For such a time like this-we have our fighting man!*

Likewise, strong group identity cues are evident within election fraud echo chambers. For instance:

> *@realDonaldTrump keep fighting like the warrior you are! We are with you! #VoterFraud https://t.co/RuxhsXyO8U*

Note that we observe an inverse coefficient for Monday Motivation. We expect this may be attributed to the inherent characteristics of the challenge, wherein individuals share personal experiences with their communities [60].

5.2 Implicit identity signaling through ambient affiliation

5.2.1 Methodology. Building on our observations of increased explicit group identity formation within echo chambers for misinformation-related topics in particular, we explore differences in ambient affiliation between echo chambers and the remainder of interaction network during discussions of misinformation-related topics.
Top hashtags. We extract all hashtags present inside the dataset. From these, we select the 10 with highest density, standardized
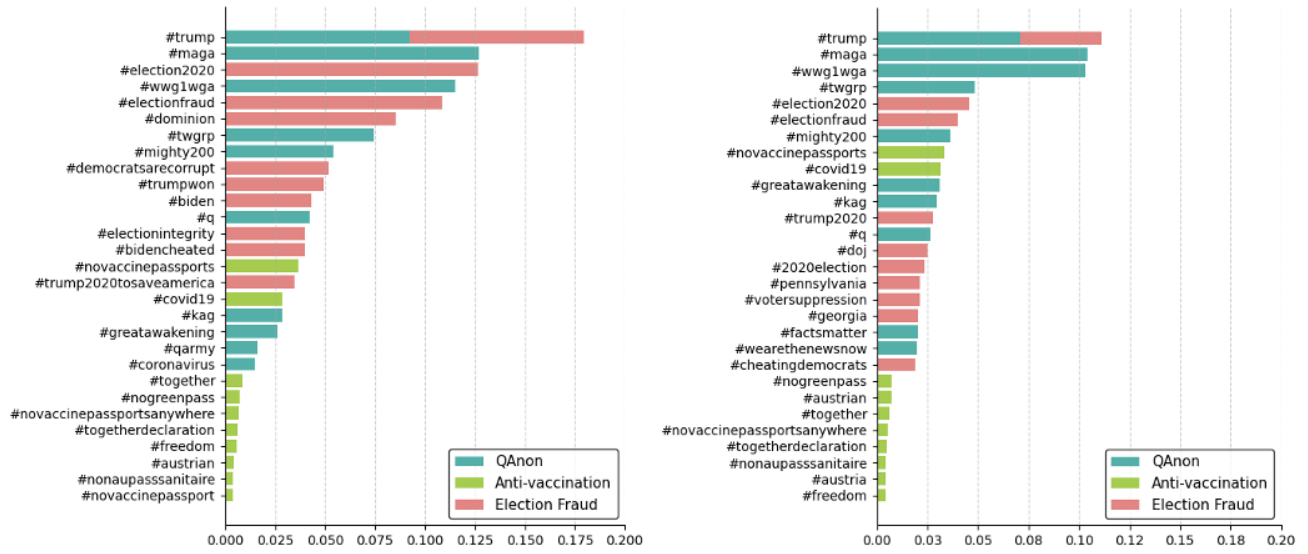
Figure 1: Top 10 hashtags, by density, in echo chambers (left) and overall (right) for misinformation topics.

Category	Topic	Coefficient	Std. Error	T-score	p-value
Misinformation	**Election Fraud**	**0.162**	**0.049**	**3.272**	**0.001****
	Antivax	0.010	0.038	0.266	0.523
	QAnon	**0.116**	**0.045**	**2.339**	**0.019***
Hyper-polarized	MeToo	0.026	0.044	0.593	0.453
	BLM	-0.058	0.046	-1.265	0.206
Neutral	**Monday Motivation**	**-0.186**	**0.056**	**-3.307**	**0.001****

Table 3: Pairwise zero-inflated beta model median statistics for group identity language using bootstrap sampling. Positive coefficient represents positive association between echo chamber membership and increased group identity language.

over the total number of tweets within each category. We exclude hashtags identical to the keywords used for data collection.

Hashtag networks. We construct undirected common hashtag networks to model ambient affiliation within echo chambers, for all three topics of misinformation. In these networks, nodes represent users within echo chambers and a weighted edge exists between users who share at least one common hashtag. As before, we exclude hashtags identical to keywords utilized during data collection.

5.2.2 Findings. A conspicuous pattern emerges, complementing our findings from in-/out-group identity analyses. Both QAnon and election fraud discussions are characterized by methodical use of ambient affiliation relative to anti-vaccination dataset. When conducting a comparison of hashtag usage within echo chambers versus overall hashtag usage, we observe that **echo chambers have notably higher rate of ambient affiliation.** This suggests members may be rallying around common hashtags to amplify their message and consolidate their identity.

Top hashtags. It is clear that the strategic use of hashtags plays an important role, both within and outside echo chambers, in QAnon-related conversations (see Figure 1). Several hashtags speak directly to group membership, such as #Mighty200 (a group of conservative

or right-wing influencers) and #wearethenewsnow (QAnon supporters use it to suggest that mainstream media cannot be trusted, positioning the QAnon community as the authentic purveyors of news). Some hashtags are slogan-like, e.g., #WWG1WGA ("Where We Go One, We Go All") and #kag ("Keep America Great").

Hashtag networks. Hashtag network statistics are provided in Table 4. In the QAnon echo chamber, the hashtag network exhibits the highest density, suggesting a deliberate and strategic approach to hashtag bulk use, likely for content promotion. Additionally, we see high percentage of ambient affiliation engagement and average common hashtags per paired users for election fraud, which is consistent with the high identity consolidation observed within election fraud discussions. Conversely, we observe limited hashtag usage within anti-vaccination discourse, which is also mirrored in the network structure. These findings highlight the complex nature of online discourse and nuance under the very broad umbrella of misinformation-related conversations.

5.3 Case Study 1. Echo chambers in conversations about US election fraud

We observe strongest group identity formation within US election fraud discourse. To further explore this, we conduct a case study

analyzing interactions within, between, and outside echo chambers. We find that echo chamber members are deeply embedded in the broader interaction network, maintaining active engagements with individuals outside their group. This heightened activity of echo chamber members in external interactions challenges the simplistic view of echo chambers as entirely insular entities and suggests a more nuanced understanding of discourse dynamics.

Figure 2 shows the internal interaction network amongst members of the largest echo chamber (red) and other echo chamber members with which they interact. We observe echo chambers are interconnected rather than isolated. Two nodes, in particular, exhibit highest centrality: Node EC_a, a member of the largest echo chamber; and, EC_b, a bridge between echo chambers. Deeper examination of the content shared by these two nodes reveals distinct roles. Node EC_a emerges as a pivotal resource hub, disseminating information about the election status, often accompanied by a URL. These URLs typically redirect to tweets conveying similar content, supplemented by relevant hashtags.

Furthermore, we notice several indicators of group identity from EC_a that extend beyond our initial examination of first and third-person plural pronouns. This content manifests as expressions of favoritism and support within the group or, conversely, as extreme animosity towards those outside the group. For instance:

> *Good morning @realDonaldTrump. I'm still here, standing with you and 70 million other freedom loving patriots. It time to expose the #voterfraud. #DoNotConcede #DoNotConcedeUnderAnyCircumstances #DoNotGiveUp*
> *GDMRNNG (The vowels have been stolen by Democrats) #Election2020 #VoterFraud #BidenHarris2020 #2020Elections* (Explanation: GDMRNNG is goodmorning without vowels.)

We observe that EC_b mainly interacts with itself. Among 111 tweets posted by EC_b, 93 are retweets or replies to its own tweets. Within this set of tweets, we notice: heavy use of user mentioning, an average of 46.32 mentions per tweet; use of hashtags, an average of 5.7 hashtags per tweet; and ending the post with a URL.

6 RQ2: PROCESSING FLUENCY EFFECTS

We explore the impact of processing fluency on discourse around misinformation-related topics, with particular focus on the variance between echo chambers and external members. This study is not intended to compare the overall simplicity or elaboration between misinformation and control topics, but to identify differences in processing fluency between discussions within echo chambers and those outside of them. We utilize a series of linguistic metrics to measure processing fluency, complemented by an analysis of URLs, to understand these dynamics in detail.

6.1 Processing fluency through linguistic simplicity

6.1.1 Methodology. Simplicity is captured through readability and word length measurements.

Readability. As discussed in Preliminaries, we utilize the Flesch readability ease score. It considers not just word-level but also
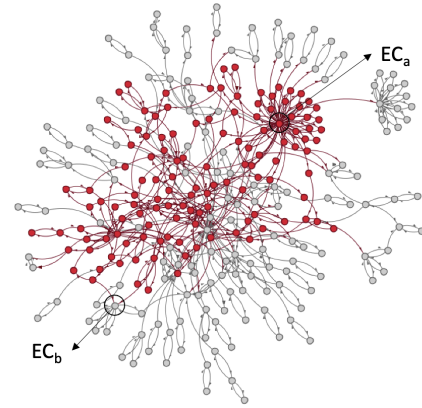

Figure 2: US election fraud. Internal interaction network amongst members of the largest echo chamber (red) and other echo chamber members with which they interact.

sentence-level complexity. Text processing follows as above, retaining all punctuation marks. Calculation of the Flesch score requires text at least 100 words, so our approach deviated from standard data-point level comparison as follows. To address potential sampling bias, we employed a stratified sampling method wherein we selected 100 posts each from both the echo chamber and external post datasets and aggregated them for readability score calculation. This was repeated for 100 iterations. Given that the datasets exhibited a near-normal distribution, as depicted in Figure 3, we utilize the T-test for analyzing mean differences. However, we observe non-uniform variance in certain instances, such as with the BLM topic. To reinforce the validity of our comparisons under these conditions, we additionally applied the non-parametric Mann-Whitney U test. Notably, the results from both testing methods converged.

Big words. We employ LIWC to calculate the "bigword" score for each tweet. To avoid biases, we apply this analysis to cleaned texts, excluding hashtags, mentions, URLs, and similar elements that could be erroneously identified as "big words." Given that resulting scores are no longer bounded between 0 and 1, the zero-inflated beta model proves unsuitable for our comparative tasks. As the datasets being compared exhibit approximately normal distribution and similar variances, we employ an independent two-sample T-test. Given our relatively large dataset, this test tends to yield very small p-values, hence less informative. To mitigate this effect or potential non-normality in distribution, we perform bootstrapping by randomly sampling 1000 data points from both the echo chamber and non-echo chamber groups and conducting the t-test over 1000 iterations [3, 26, 41]. We select the median as the representative outcome.

6.1.2 Findings. **Our findings demonstrate clearly and consistently that simplicity, as a measure of processing fluency, is amplified within echo chambers in misinformation-related discussions.**

Topic	hashtag Network Density	Ambient Affiliation Engagement	Average Common Hashtags per Paired Users
Election Fraud	0.0721	52.34% (246/470)	1.2327
Anti-vax	0.0716	8.37% (1523/18200)	1.0310
QAnon	0.0949	33.15% (2962/8934)	1.2123

Table 4: Hashtag network statistics for three misinformation topics.

Readability Ease. Outcomes are visualized in Figure 3. Content within echo chambers consistently demonstrates higher readability ease across all three misinformation topics when compared to external content. Conversely, in the cases of Metoo and Monday Motivation, a contrasting pattern emerges, indicating a higher level of complexity within echo chambers. No significant differences were observed in discussions surrounding BLM. Here, we emphasize the contrast between echo chambers and external members rather than making broad comparisons of readability ease between misinformation and control topics. The apparent lower overall readability ease of misinformation topics is a reflection of the topics' inherent characteristics, e.g., the complexity of these discussions or the use of specialized vocabulary.

Big words. The results of all independent two-sample T-tests are presented in Table 5. We also report the percentage of iterations having mean EC smaller than mean EXT and the percentage of p-value<0.05 in the table. Our findings consistently indicate that all misinformation topics yield a lower average percentage of big words in echo chamber posts compared to non-echo chamber posts, with all p-values less than 0.05. This suggests that echo chambers exhibit less linguistic complexity, possibly due to the development of a shared knowledge base within their community.
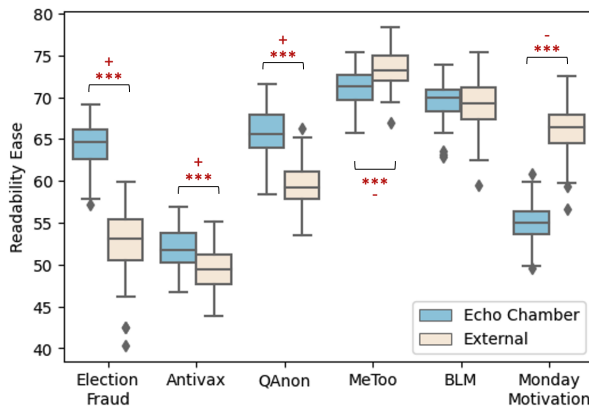

Figure 3: Readability ease score box plot for 100 samples of aggregated text. ***p<0.001, +significance in the positive direction, -significance in the negative direction.

6.2 Processing fluency through elaboration

6.2.1 Methodology. Elaboration is captured through the use of discourse connectives. Similar to our analysis of in-/out-group language, we employ a zero-inflated beta regression model with bootstrap sampling to examine the relationship between echo chamber membership and the use of discourse connectives.

6.2.2 Findings. Median statistics are shown in Table 6. We observe that US election fraud discourse shows a positive relationship between echo chamber membership and additive and negative connectives. QAnon exhibits a positive significant coefficient for logical and additive connectives. Conversely, in the context of the control topics Monday Motivation and BLM, an entirely opposite behavior is observed; echo chambers tend to exhibit a lower prevalence of discourse connectives, aligning with our findings above that suggest greater simplicity in these communities.

Generally speaking, we might expect fewer discourse connectives in echo chambers due to greater preexisting consensus amongst members. However, the nature of some misinformation-related topics requires more elaboration and explanation, and might explain the increased use of connectives in these contexts.

In line with prior work on QAnon [44], **we also note cues of explanation-oriented behavior within QAnon echo chambers**, co-occuring with straightforward vocabulary. QAnon proponents often use phrases like "do the research." These phrases appear to steer users toward an archive of QAnon-related information, likely intended for persuasive or reinforcing purposes. For instance:

> There are many sides to any issue. It doesńt matter what others think. What matters is what you think. Do the research. Dig. Come to your own conclusion. Seek the truth knowing that you will never know what TRUTH really is, but come as close as you possibly can. #QAnon https://t.co/W68boAmWW3
>
> People no longer have the luxury of claiming ignorance is why they believe the things they believe. We are living in a time where the truth can be easily accessed via documentation, Youtube, QAnon and even if you don't want to do the research their are many who do and give.

A pattern in many such tweets is the inclusion of URLs, guiding users to external sources, e.g., news articles, videos, or other tweets to "validate" claims. While the strategic use of archival information may vary by topic, it complements the linguistic characteristics we have identified and can be an asset for debunking misinformation.

Category	Topic	Avg.EC	Avg. EXT	T-score	p-value	% Avg.EC <Avg.EXT	%p-value <0.05
Misinformation	**Election Fraud**	**22.973**	**26.638**	**-5.440**	**5.993e-8*****	**100%**	**99.8%**
	Antivax	**25.895**	**26.638**	**-2.630**	**0.009***	**99.9%**	**75.0%**
	QAnon	**22.044**	**23.731**	**-2.692**	**0.007***	**99.9%**	**75.0%**
Hyper-polarized	MeToo	19.520	19.043	0.854	0.351	19.6%	12.4%
	BLM	18.612	19.482	-1.156	0.120	95.6%	34.6%
Neutral	**Monday Motivation**	**24.133**	**19.767**	**7.239**	**6.412e-13*****	**0.0%**	**94.7%**

Table 5: Independent two-sample T-test median statistics for big words after 1000 iterations of bootstrap sampling. A significant p-value, coupled with a positive T score, indicates that echo chambers exhibit a greater utilization of big words.

Category	Topic	Linguistic metrics	Coefficient	Std. Error	T-score	p-value
Misinformation	Election Fraud	Logical	0.066	0.046	1.435	0.149
		Additive	**0.171**	**0.043**	**4.040**	**5.600e-5*****
		Negative	**0.198**	**0.072**	**2.721**	**0.007***
	Antivax	Logical	-0.043	0.032	-1.340	0.180
		Additive	0.014	0.033	0.411	0.513
		Negative	-0.046	0.040	-1.148	0.248
	QAnon	**Logical**	**0.083**	**0.042**	**1.999**	**0.046***
		Additive	**0.106**	**0.040**	**2.646**	**0.008***
		Negative	0.094	0.066	1.427	0.154
Hyper-polarized	MeToo	Logical	-0.021	0.035	-0.592	0.453
		Additive	0.044	0.048	0.915	0.340
		Negative	0.015	0.048	0.305	0.512
	BLM	**Logical**	**-0.126**	**0.041**	**-3.060**	**0.002****
		Additive	**-0.085**	**0.038**	**-2.264**	**0.024***
		Negative	**-0.118**	**0.059**	**-1.997**	**0.046***
Neutral	Monday Motivation	**Logical**	**-0.219**	**0.040**	**-5.420**	**0.000*****
		Additive	**-0.104**	**0.035**	**-2.960**	**0.003****
		Negative	**-0.211**	**0.066**	**-3.182**	**0.001****

Table 6: Pairwise zero-inflated beta model median statistics of discourse connectives after bootstrap sampling(positive coefficient represents a positive association between membership in echo chambers and increased use of discourse connectives).

6.3 Case Study 2: URL usage in discussions of misinformation

Including URLs or referencing news sources can appear to enhance the credibility of information presented online without the need for extensive explanation. We explore whether echo chamber members exhibit increased URL usage, as compared to external nodes. To examine this, we assess URL density in interactions within echo chambers, outside echo chambers, and in standalone original tweets across our three misinformation topics, as shown in Figure 4. Notably, we observe a consistently higher URL density in isolated posts across all misinformation-related topics. Conversely, echo chamber interactions tend to utilize URLs less frequently, perhaps suggesting that users within echo chambers may not feel the need to persuade each other with new information, given their synchronized beliefs.

In between-topic comparisons, QAnon discussions have most frequent URL usage, followed by election fraud. In contrast, discussions about anti-vaccination do not engage as much with URLs. Among QAnon tweets, 73.68% contains at least one URL, in comparison to Election Fraud (69.81%) and Antivax (44.45%). This trend

is accentuated when considering tweets with at least two URLs. Within QAnon discourse, a substantial 10.82% of tweets fall into this category, as compared to 3.51% for US election fraud and 3.91% for anti-vaccination. Extensive use of URLs in QAnon discussions suggests a deliberate strategy to strengthen main assertions with multiple references.

7 DISCUSSION AND CONCLUSION

Our work is motivated by longstanding theories from the socio-psychologial and socio-linguistics literatures that describe phenomena like collective belonging through social identity and the importance of familiarity and fluency in information credibility. We operationalize these theories through a suite of computational metrics and analyses, which let us test these theories in the wild through social media with particular focus on discussions around misinformation-related topics. This approach also allows us to differentially explore echo chambers and understand their linguistic underpinnings.
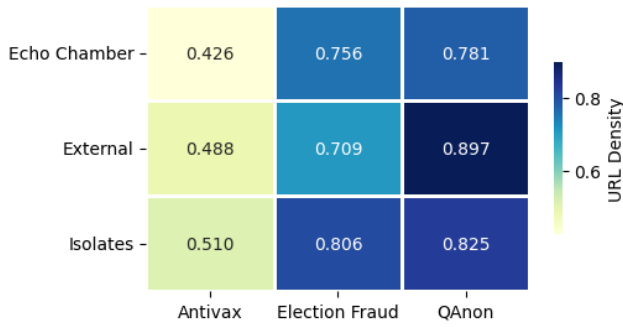
Figure 4: Heatmap of URL density of echo chamber interactions, non-echo chamber interactions (excluding isolates), and isolated tweets for misinformation topics

We do not see uniform behavior across all misinformation topics. Perhaps surprisingly, anti-vaccination discourse shows distinct linguistic patterns from US election fraud and QAnon discussions. This may signify the differential impact of *political* group identities, in particular, on the linguistic characteristics of politically-charged topics. We also note that the neutral topic we selected (#monday-motivation) consistently displays behaviors contrary to those of misinformation. This underscores that echo chamber dynamics are not uniform but are influenced by discourse.

These insights might enable identification of social communities prone to influence through anomalous linguistic signs. Moreover, understanding the anatomy of echo chambers offers a blueprint for designing more effective and targeted moderation strategies. By studying the linguistic underpinning of social communities around misinformation-related discussions, we offer a fresh perspective on dynamics of information dissemination in the digital age.

8 LIMITATIONS AND FUTURE WORK

Our methodology necessitates the selection of a data collection time frame coincident with peak popularity of the respective topics. Given our intentional control over the total volume of tweets, it is an intrinsic outcome that different topics span different lengths of time. Future work should consider controlling for temporal variables to better understand the influence of time on our findings.

Additionally, here we employed a relatively strict, network structure-based definition of echo chambers using strongly connected components within topic-specific discourse. It is possible that other approaches to defining echo chambers would yield different insights.

Finally, our study selected the zero-inflated beta model to analyze the relationship between echo chamber membership and linguistic patterns as it is suitable for the data distribution. The extensive size of our dataset results in the models being hyper-sensitive, detecting even the slightest deviations in values. This heightened sensitivity is not unique to this model but is observed in other tests we employed, such as permutation tests. As a remedy, we adopted bootstrapping across 1000 iterations for all tests, using the median as a representative measure. Nevertheless, it is the direction and magnitude of the coefficients that offer deeper insights

into the interplay between echo chamber dynamics and linguistic underpinnings in discussions.

9 ETHICAL CONSIDERATIONS

This study primarily involves the comprehensive analysis of data collected from social media users on a large scale. The data was obtained from the Twitter Developer API is solely utilized for the purposes of this research. We ensure that the analyses conducted do not display any identifiable information of users. Tweet IDs are shared with the public for reproducibility of the analyses in this work. Our findings are expected to support ongoing conversations within data ethics about misinformation and polarization on social media, and potential interventions.

ACKNOWLEDGMENTS

REFERENCES

[1] Elissa M Abrams and Matthew Greenhawt. 2020. Mitigating misinformation and changing the social narrative. *The Journal of Allergy and Clinical Immunology: In Practice* 8, 10 (2020), 3261–3263.
[2] Aseel Addawood, Adam Badawy, Kristina Lerman, and Emilio Ferrara. 2019. Linguistic cues to deception: Identifying political trolls on social media. In *Proceedings of the international AAAI conference on web and social media*, Vol. 13. 15–25.
[3] Nor Aishah Ahad, Suhaida Abdullah, Choo Heng Lai, and Nazihah Mohd Ali. 2012. Relative power performance of t-test and bootstrap procedure for two-sample. *Pertanika Journal of Science & Technology* 20, 1 (2012), 43–52.
[4] Hunt Allcott, Matthew Gentzkow, and Chuan Yu. 2019. Trends in the diffusion of misinformation on social media. *Research & Politics* 6, 2 (2019), 2053168019848554.
[5] Adam L Alter and Daniel M Oppenheimer. 2009. Uniting the tribes of fluency to form a metacognitive nation. *Personality and social psychology review* 13, 3 (2009), 219–235.
[6] John Robert Anderson and Jane Crawford. 1980. *Cognitive psychology and its implications.* wh freeman San Francisco.
[7] Richard C Bailey. 2013. Increasing processing fluency in the classroom instructional system. *CELE Journal* 21 (2013), 43–52.
[8] Pablo Barberá. 2020. Social media, echo chambers, and political polarization. *Social media and democracy: The state of the field, prospects for reform* 34 (2020).
[9] Pablo Barberá, John T Jost, Jonathan Nagler, Joshua A Tucker, and Richard Bonneau. 2015. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological science* 26, 10 (2015), 1531–1542.
[10] Christian Bentz, Dan Dediu, Annemarie Verkerk, and Gerhard Jäger. 2018. The evolution of language families is shaped by the environment beyond neutral drift. *Nature Human Behaviour* 2, 11 (2018), 816–821.
[11] Bethan Benwell. 2006. *Discourse and identity.* Edinburgh University Press.
[12] Jan Blommaert. 2010. *The sociolinguistics of globalization.* Cambridge University Press.
[13] J Scott Brennen, Felix M Simon, Philip N Howard, and Rasmus Kleis Nielsen. 2020. *Types, sources, and claims of COVID-19 misinformation.* Ph. D. Dissertation. University of Oxford.
[14] Marilynn B Brewer et al. 1999. The psychology of prejudice: Ingroup love or outgroup hate? *Journal of social issues* 55 (1999), 429–444.
[15] Axel Bruns. 2017. Echo chamber? What echo chamber? Reviewing the evidence. In *6th Biennial Future of Journalism Conference (FOJ17).*
[16] Leonardo Bursztyn, Aakaash Rao, Christopher P Roth, and David H Yanagizawa-Drott. 2020. *Misinformation during a pandemic.* Technical Report. National Bureau of Economic Research.
[17] Rossella Canestrino, Pierpaolo Magliocca, and Yang Li. 2022. The Impact of Language Diversity on Knowledge Sharing Within International University Research Teams: Evidence From TED Project. *Frontiers in Psychology* 13 (2022), 879154.
[18] Shelly Chaiken and Alison Ledgerwood. 2012. A theory of heuristic and systematic information processing. *Handbook of theories of social psychology* 1 (2012), 246–266.
[19] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. 2021. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences* 118, 9 (2021).
[20] Wesley Cota, Silvio C Ferreira, Romualdo Pastor-Satorras, and Michele Starnini. 2019. Quantifying echo chamber effects in information spreading over political communication networks. *EPJ Data Science* 8, 1 (2019), 35.

[21] Ludivine Crible, Mathis Wetzel, and Sandrine Zufferey. 2021. Lexical and structural cues to discourse processing in first and second language. *Frontiers in psychology* 12 (2021), 685491.

[22] Scott A Crossley, Kristopher Kyle, and Danielle S McNamara. 2016. The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior research methods* 48, 4 (2016), 1227–1237.

[23] Sergio Currarini, Matthew O Jackson, and Paolo Pin. 2009. An economic model of friendship: Homophily, minorities, and segregation. *Econometrica* 77, 4 (2009), 1003–1045.

[24] Rick Dale and Gary Lupyan. 2012. Understanding the origins of morphological diversity: The linguistic niche hypothesis. *Advances in complex systems* 15, 03n04 (2012), 1150017.

[25] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H Eugene Stanley, and Walter Quattrociocchi. 2016. The spreading of misinformation online. *Proceedings of the National Academy of Sciences* 113, 3 (2016), 554–559.

[26] Jorge Faber and Lilian Martins Fonseca. 2014. How sample size influences research outcomes. *Dental press journal of orthodontics* 19 (2014), 27–29.

[27] Anjalie Field, Chan Young Park, Antonio Theophilo, Jamelle Watson-Daniels, and Yulia Tsvetkov. 2022. An analysis of emotions and the prominence of positivity in# BlackLivesMatter tweets. *Proceedings of the National Academy of Sciences* 119, 35 (2022), e2205767119.

[28] Rudolph Flesch. 1948. A new readability yardstick. *Journal of applied psychology* 32, 3 (1948), 221.

[29] John Fox and Sanford Weisberg. 2002. Bootstrapping regression models. *An R and S-PLUS Companion to Applied Regression: A Web Appendix to the Book. Sage, Thousand Oaks, CA. URL http://cran. r-project. org/doc/contrib/Fox-Companion/appendix-bootstrapping. pdf* (2002).

[30] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. Datasheets for datasets. *Commun. ACM* 64, 12 (2021), 86–92.

[31] Steven Greene. 2004. Social identity theory and party identification. *Social Science Quarterly* 85, 1 (2004), 136–153.

[32] "Five Graces Group", Clay Beckner, Richard Blythe, Joan Bybee, Morten H Christiansen, William Croft, Nick C Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman, et al. 2009. Language is a complex adaptive system: Position paper. *Language learning* 59 (2009), 1–26.

[33] Andrew Guess, Brendan Nyhan, Benjamin Lyons, and Jason Reifler. 2018. Avoiding the echo chamber about echo chambers. *Knight Foundation* 2 (2018), 1–25.

[34] Lynn Hasher, David Goldstein, and Thomas Toppino. 1977. Frequency and the conference of referential validity. *Journal of verbal learning and verbal behavior* 16, 1 (1977), 107–112.

[35] Isabel Íñigo-Mora. 2004. On the use of the personal pronoun we in communities. *Journal of Language and Politics* 3, 1 (2004), 27–52.

[36] Natascha A Karlova and Karen E Fisher. 2013. A social diffusion model of misinformation and disinformation for understanding human information behaviour. (2013).

[37] KP Krishna Kumar and G Geethakumari. 2014. Detecting misinformation in online social networks using cognitive psychology. *Human-centric Computing and Information Sciences* 4, 1 (2014), 1–22.

[38] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. 2010. What is Twitter, a social network or a news media?. In *Proceedings of the 19th international conference on World wide web*. 591–600.

[39] Ro'ee Levy. 2021. Social media, news consumption, and polarization: Evidence from a field experiment. *American economic review* 111, 3 (2021), 831–70.

[40] Jianing Li and Min-Hsin Su. 2020. Real talk about fake news: Identity language and disconnected networks of the US public's "fake news" discourse on Twitter. *Social Media+ Society* 6, 2 (2020), 2056305120916841.

[41] Thomas Lumley, Paula Diehr, Scott Emerson, and Lu Chen. 2002. The importance of the normality assumption in large public health data sets. *Annual review of public health* 23, 1 (2002), 151–169.

[42] Gary Lupyan and Rick Dale. 2010. Language structure is partly determined by social structure. *PloS one* 5, 1 (2010), e8559.

[43] Gary Lupyan and Rick Dale. 2016. Why are there different languages? The role of adaptation in linguistic diversity. *Trends in cognitive sciences* 20, 9 (2016), 649–660.

[44] Alice E Marwick and William Clyde Partin. 2022. Constructing alternative facts: Populist expertise and the QAnon conspiracy. *New Media & Society* (2022), 14614448221090201.

[45] Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* 27, 1 (2001), 415–444.

[46] Lesley Milroy and James Milroy. 1992. Social network and social class: Toward an integrated sociolinguistic model1. *Language in society* 21, 1 (1992), 1–26.

[47] Daniel M Oppenheimer. 2008. The secret life of fluency. *Trends in cognitive sciences* 12, 6 (2008), 237–241.

[48] Ethan Pancer, Vincent Chandler, Maxwell Poole, and Theodore J Noseworthy. 2019. How readability shapes social media engagement. *Journal of Consumer Psychology* 29, 2 (2019), 262–270.

[49] Pennebaker Conglomerates, Inc. 2022. INTRODUCING LIWC-22 A NEW SET OF TEXT ANALYSIS TOOLS AT YOUR FINGERTIPS. https://www.liwc.app/

[50] Alastair Pennycook. 1994. The politics of pronouns. (1994).

[51] Gordon Pennycook, Tyrone D Cannon, and David G Rand. 2018. Prior exposure increases perceived accuracy of fake news. *Journal of experimental psychology: general* 147, 12 (2018), 1865.

[52] RE Petty and DT Wegener. 1999. The elaboration likelihood model: Current status and controversies In Dual-process theories in social psychology (pp. 37–72). *New York, NY: Guilford Press.[Google Scholar]* (1999).

[53] E Petty Richard and John T Cacioppo. 1986. Communication and persuasion: Central and peripheral routes to attitude change.

[54] Steve Rathje, Jay J Van Bavel, and Sander van der Linden. 2021. Out-group animosity drives engagement on social media. *Proceedings of the National Academy of Sciences* 118, 26 (2021).

[55] Limor Raviv, Antje Meyer, and Shiri Lev-Ari. 2020. The role of social network structure in the emergence of linguistic structure. *Cognitive Science* 44, 8 (2020), e12876.

[56] Rolf Reber. 2012. Processing fluency, aesthetic pleasure, and culturally shared taste. *Aesthetic science: Connecting minds, brains, and experience* (2012), 223–249.

[57] Rolf Reber and Christian Unkelbach. 2010. The epistemic status of processing fluency as source for judgments of truth. *Review of philosophy and psychology* 1, 4 (2010), 563–581.

[58] Rolf Reber, Piotr Winkielman, and Norbert Schwarz. 1998. Effects of perceptual fluency on affective judgments. *Psychological science* 9, 1 (1998), 45–48.

[59] Samuel C Rhodes. 2022. Filter bubbles, echo chambers, and fake news: how social media conditions individuals to be less critical of political misinformation. *Political Communication* 39, 1 (2022), 1–22.

[60] Catherine M Ridings, David Gefen, and Bay Arinze. 2002. Some antecedents and effects of trust in virtual communities. *The journal of strategic information systems* 11, 3-4 (2002), 271–295.

[61] Seán Roberts and James Winters. 2012. Social structure and language structure: The new nomothetic approach. *Psychology of Language and Communication* 16, 2 (2012), 89–112.

[62] Nancy L Rosenblum and Russell Muirhead. 2019. 1. Conspiracy without the Theory. In *A Lot of People Are Saying*. Princeton University Press, 19–41.

[63] Jieun Shin, Lian Jian, Kevin Driscoll, and François Bar. 2018. The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior* 83 (2018), 278–287.

[64] Herbert A Simon. 1978. Information-processing theory of human problem solving. *Handbook of learning and cognitive processes* 5 (1978), 271–295.

[65] Andrew N Smith, Eileen Fischer, and Chen Yongjian. 2012. How does brand-related user-generated content differ across YouTube, Facebook, and Twitter? *Journal of interactive marketing* 26, 2 (2012), 102–113.

[66] Jan E Stets and Peter J Burke. 2000. Identity theory and social identity theory. *Social psychology quarterly* (2000), 224–237.

[67] Caroline Tagg and Philip Seargeant. 2014. *The language of social media: Identity and community on the internet*. Palgrave Macmillan.

[68] Henri Tajfel and John C Turner. 2004. The social identity theory of intergroup behavior. (2004).

[69] Henri Tajfel, John C Turner, William G Austin, and Stephen Worchel. 1979. An integrative theory of intergroup conflict. *Organizational identity: A reader* 56, 65 (1979), 9780203505984–16.

[70] Alexander Todorov, Shelly Chaiken, and Marlone D Henderson. 2002. The heuristic-systematic model of social information processing. *The persuasion handbook: Developments in theory and practice* 23 (2002), 195–211.

[71] Sabine Trepte and Laura S Loy. 2017. Social identity theory and self-categorization theory. *The international encyclopedia of media effects* (2017), 1–13.

[72] Craig W Trumbo. 1999. Heuristic-systematic information processing and risk judgment. *Risk Analysis* 19, 3 (1999), 391–400.

[73] Christian Unkelbach and Sarah C Rom. 2017. A referential theory of the repetition-induced truth effect. *Cognition* 160 (2017), 110–126.

[74] Christian Unkelbach and Felix Speckmann. 2021. Mere Repetition Increases Belief in Factually True COVID-19-Related Information. *Journal of Applied Research in Memory and Cognition* (2021).

[75] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science* 359, 6380 (2018), 1146–1151.

[76] Ying Xiong, Moonhee Cho, and Brandon Boatwright. 2019. Hashtag activism and message frames among social movement organizations: Semantic network analysis and thematic analysis of Twitter during the# MeToo movement. *Public relations review* 45, 1 (2019), 10–23.

[77] Michele Zappavigna. 2011. Ambient affiliation: A linguistic perspective on Twitter. *New media & society* 13, 5 (2011), 788–806.

[78] Michele Zappavigna. 2012. *Discourse of Twitter and social media: How we use language to create affiliation on the web*. Vol. 6. A&C Black.

[79] Michele Zappavigna. 2014. Enacting identity in microblogging through ambient affiliation. *Discourse & Communication* 8, 2 (2014), 209–228.

[80] Justine Zhang, William Hamilton, Cristian Danescu-Niculescu-Mizil, Dan Jurafsky, and Jure Leskovec. 2017. Community identity and user engagement in a multi-community landscape. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 11.

[81] Melissa Zimdars and Kembrew McLeod. 2020. *Fake news: understanding media and misinformation in the digital age.* MIT Press.

[82] Fabiana Zollo, Petra Kralj Novak, Michela Del Vicario, Alessandro Bessi, Igor Mozetič, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. 2015. Emotional dynamics in the age of misinformation. *PloS one* 10, 9 (2015), e0138740.

[83] Sandrine Zufferey, Willem Mak, Liesbeth Degand, and Ted Sanders. 2015. Advanced learners' comprehension of discourse connectives: The role of L1 transfer across on-line and off-line tasks. *Second Language Research* 31, 3 (2015), 389–411.