Journal of Quantitative Description: Digital Media 6(2026), 1–99 DOI 10.51685/jqd.2026.016

Faker Island: Profiling misinformation sharers on the French Twittersphere

MANON BERRICHE[1]
Sciences Po Paris, France

JEAN-PHILIPPE COINTET

SACHA ALTAY
University of Zurich, Switzerland

ANNE-SOPHIE HACQUIN
Direction Interministérielle de la Transformation Publique, France

BÉATRICE MAZOYER
Sciences Po Paris, France

BENJAMIN OOGHE-TABANOU

HUGO MERCIER
École Normale Supérieure, France

DOMINIQUE CARDON
Sciences Po Paris, France

Copyright © 2026 (Manon Berriche, Jean-Philippe Cointet, Sacha Altay, Anne-Sophie Hacquin, Béatrice Mazoyer, Benjamin Ooghe-Tabanou, Hugo Mercier, Dominique Cardon). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at: http://journalqd.org

What are the distinctive characteristics of misinformation sharers? So far, most studies have focused on the U.S., a two-party system with a highly polarized media landscape, leaving open the question of what misinformation sharers look like in other socio-political contexts. Research has also prioritized testing causal links between cognitive or political variables and misinformation sharing, while paying less attention to fine-grained descriptions of who these users are. To address this gap, we analyze the French Twittersphere using ideological inference methods and a matching procedure. Drawing on four million tweets, we compare 1,907 users who shared at least one URL flagged as false or misleading by fact-checkers with 958 ideologically matched control users who did not. We show that misinformation sharers are more likely to be located at the extremes of the political spectrum, express hostility toward ruling elites and institutions, display political markers in their bios, use pseudonyms, share media content, engage in high-intensity and retweet-oriented activity, and express negative emotions. They also disproportionately circulate topics related to immigration, terrorism, and criticism of mainstream media. However, their language does not indicate lower analytical thinking. By shifting the focus from cognitive vulnerability alone to political self-presentation, anger, and amplification practices, this study contributes to ongoing theoretical debates and calls for further research on the activist practices and party affiliations of misinformation sharers.

Keywords: *Anti-establishment; Fake News; Twitter; Partisanship; Ideological Inference; Matching Procedure; Misinformation.*

[1] Manon Berriche: manon.berriche@sciencespo.fr
Date submitted: 2026-02-12

1. Introduction

In recent years, numerous quantitative studies have measured the proportion of users who share misinformation on social media and identified their main characteristics (for an overview of studies conducted on Twitter, see Supplementary Table 1). By treating URLs, domains, or news items rated as unreliable or misleading by fact-checkers as misinformation, these studies identify the users who spread such content, collect their digital traces, and analyze their behavior. They find that misinformation consumption and sharing are concentrated among a small subset of users, who tend to cluster at the political extremes, especially on the far right (Gaumont et al., 2018; Grinberg et al., 2019; Osmundsen et al., 2021; Kennedy et al., 2022; DeVerna et al., 2022; Mosleh and Rand, 2022; Hopp et al., 2020; de Oliveira et Albuquerque, 2021; Rathje et al., 2022; Baribi-Bartov et al., 2024). Although these findings have been corroborated by several studies, three main limitations stand out.

First, most studies have focused on the United States, a country characterized by a bipartisan political system and a highly polarized media landscape (Benkler et al., 2018). Large-scale investigations in other socio-political contexts are scarce, limiting our understanding of whether misinformation sharers' patterns hold in multi-party systems. In European countries, other ideological dimensions, beyond left-right positions, might be important to consider (Benoit and Laver, 2012; Bakker et al., 2012). Several studies have highlighted the importance of individuals' positions on issues related to European integration, immigration, and globalization in explaining different behaviors (Hix et al., 2006; Ramaciotti Morales et al., 2021). In France, for example, research suggests that ideological dimensions beyond left-right polarization, such as anti-elite sentiment, play a crucial role in social movements like the Yellow Vests (Cointet et al., 2021a) or during the COVID-19 health crisis (Berriche et al., 2020). Recent comparative work on the United States and France also suggests that misinformation and affective polarization may reinforce each other across contexts, while taking different political forms depending on

national party systems and cleavage structures (Temporão, 2026). This raises the following research question:

RQ1: What are the political orientations of users who share misinformation in France, and to what extent are they concentrated on the far right, as prior research on the United States would suggest?

Second, existing studies based on social media data are largely correlational and cannot establish whether misinformation sharing is primarily driven by political motivations or by cognitive biases such as a lack of analytical reasoning. Cognitive scientists Gordon Pennycook and David Rand (2021) argue that misinformation sharers may not only have partisan motives but also weaker critical thinking. After reviewing data from 14 experimental studies, in which American participants were asked to assess the credibility of various news items (true or false; pro-Democratic or pro-Republican), they demonstrated that the effect of political congruence was weaker than the effect of information accuracy. Participants considered true but politically incongruent information to be more credible than false but politically aligned information. This suggests that analytical reasoning plays a greater role than ideology in determining whether people accept or reject misinformation. However, experimental studies face ecological validity issues and rely on self-reported data, which may not accurately reflect real-world behavior (Baumeister et al., 2007). Conversely, observational studies provide a more accurate depiction of social media dynamics, without interfering with them, but do not allow for causal claims, as they cannot determine which variables influence one another (Lazer et al., 2009; Guess et al., 2019). This raises a methodological challenge: how can observational data be used to rigorously compare misinformation sharers with other users and disentangle differences associated with political positioning from those associated with cognitive processing? While a matching procedure (Rubin, 1973) cannot establish whether political motivations or cognitive limitations causally drive misinformation sharing, it can, however, control for political position by comparing misinformation sharers with

ideologically matched control users. This makes it possible to assess whether the two groups still differ in indicators associated with cognitive processing and affective expression, as well as in other characteristics. We therefore ask:

RQ2: When political positioning is held constant through a matching procedure, do misinformation sharers differ from ideologically matched control users in indicators associated with cognitive processing and affective expression, as well as in other characteristics?

Third, most academic research has so far tended to explain the sharing of misinformation by the theory of inattention (Pennycook & Rand, 2021) or by the theory of ideological polarization (Osmundsen et al., 2021, Ramaciotti Morales et al., 2023). As such, academic studies on misinformation have prioritized testing causal links between cognitive or political variables and misinformation sharing but have overlooked more descriptive analyses of who these users are (Wagner & Boczkowski, 2019). This article addresses this gap by adopting an in-depth descriptive approach in order to reveal which variables beyond analytical reasoning or political orientation differentiate users who share misinformation from ideologically matched control users:

RQ3: What other characteristics distinguish misinformation sharers from ideologically matched control users?

To answer these three research questions, we analyze the French Twittersphere using ideological inference and a matching procedure. We first detail how we identify misinformation sharers and infer their ideological position before examining the specific traits that distinguish them from matched control users.

2. Identifying misinformation sharers and inferring their ideological position

2.1. Identification of misinformation sharers

To identify users who shared at least one misinformation item on the French Twittersphere, two steps were necessary.

First, we used a dataset representing the French Twittersphere, i.e. all users posting tweets in France. This dataset was built by a team of researchers who, starting from a curated list of 420 prominent French media outlets, then collected every tweet containing a URL pointing to any of those 420 media websites by relying on the open-source Gazouilloire tool (Ooghe-Tabanou et al., 2022), which leverages both Twitter's live-streaming and search APIs. This collection process yielded over 184 million tweets posted by 3.7 million unique user accounts between May 2018 and May 2020 (for full methodological details, see Cardon et al., 2019; Cointet et al., 2021b).[2]

Next, we used two databases, each containing a list of URLs that had been evaluated as misinformation by journalists specializing in fact-checking: (1) the database of *Les Décodeurs* from *Le Monde*[3]; and (2) Facebook's Privacy Protected Shared URLs database, made available to several research teams through the Social Science One consortium (Messing et al., 2020)[4]. Within these two databases, 1,548 URLs from one of

[2] It is important to note that this is a very restrictive definition of the French Twittersphere, as on Twitter or other social networks, the number of people who share media information is very low (Odabaş, 2022).
[3] Les Décodeurs. (2017, December 19). Fausses informations : les données du Décodex. *Le Monde*. https://www.lemonde.fr/le-blog-du-decodex/article/2017/12/19/fausses-informations-les-donnees-du-decodex-en-2017_5231605_5095029.html
[4] This database contains 38 million URLs that were publicly shared at least 100 times on Facebook between January 1, 2017, and July 31, 2019. Each URL is accompanied by various metadata, including a specific variable that indicates whether it was fact-checked and provides its final classification, with labels such as "mixture or false headline," "satire," "false," "true," among others. In France, the French-speaking third-party fact-checkers collaborating with Meta include AFP Factuel, *Les Décodeurs* from *Le Monde*, *Les Observateurs* from *France 24*, and *Fake Off* from *20 Minutes*.

the 420 French media outlets mentioned above were labeled as misinformation.[5] To better characterize these URLs, we examined their distribution across the four media categories identified by Cointet et al. (2021b): mainstream media, partisan media, counter-information media, and peripheral media (e.g. tabloid or local press). These categories are not fact-checking labels. They result from a structural analysis of hyperlink relations between French media websites, hierarchically clustered using a stochastic block model. Among the 1,548 fact-checked misinformation URLs in our dataset, 72.5% come from counter-information media, 16% from peripheral media, 6% from partisan media, and 5.5% from mainstream media. This distribution shows that misinformation URLs are concentrated among counter-information outlets with low authority in the French media space, while also capturing misinformation items that occasionally originate from mainstream or peripheral media.

We were then able to identify around 85K tweets (i.e., 0.05% of the initial 184M) containing at least one misinformation item and around 31.5K accounts that had shared them (i.e., 0.9% of the initial 3.7 M). Consistent with findings from the United States, misinformation sharing on the French Twittersphere is concentrated within a small minority of users. Moreover, most users (over 90%) shared fewer than five items, and only 0.02% shared more than 50. While these figures highlight the rarity of misinformation sharing on the French Twittersphere, the operational definition of misinformation used in this study likely leads to an underestimation of its prevalence. Indeed, fact-checkers do not have the resources to assess all tweets containing questionable or false information. It is therefore highly probable that a larger number of users have already shared misinformation on Twitter. That being said, this underestimation may be counterbalanced by our restrictive definition of the French Twittersphere. Indeed, we identified a group of misinformation

[5] It is important to clarify that this identification strategy operates at the level of individual URLs rather than at the level of entire domains or websites. In other words, we do not classify media outlets as "misinformation websites," nor do we assume that all content published by a given outlet is misleading. Instead, misinformation is defined as the sharing of specific URLs corresponding to individual articles that were independently evaluated as false or misleading by fact-checkers. This distinction is important because a given media outlet may publish both accurate and inaccurate content. Our operational definition therefore concerns the specific content shared, not the overall reliability of the source.

sharers among Twitter users who had shared at least one piece of information from a French media source. However, on Twitter, as on many other social networks, the majority of users are lurkers, meaning internet users who browse others' posts but rarely publish content themselves (Odabaş, 2022). If these lurkers had been included in our reconstruction of the French Twittersphere, the percentage of misinformation sharers would likely have been much lower.

2.2. Inference of ideological position

To answer RQ1, we sought to estimate the political position of users who disseminated misinformation on the French Twittersphere. To do so, we leveraged the latent ideological space constructed through computational social science research on the French Twittersphere (Cointet et al., 2021a; Ramaciotti Morales et al., 2021). This research builds on the methodological approach developed by Pablo Barberá (2015), and inspired by a study by Robert Bond and Solomon Messing (2015) on Facebook. It is based on a classic homophily hypothesis: two individuals with similar political opinions are likely to follow the same politicians on Twitter.

In their research, Cointet et al. (2021a) and Ramaciotti Morales et al. (2021) identified all French Members of Parliament (MPs) active on Twitter (812 out of 925), collected their followers, and filtered out those following fewer than three MPs and/or having fewer than 25 followers. This process enabled them to infer the political positions of 368,831 Twitter users in France. The originality of their approach, compared to pioneering studies by Robert Bond and Solomon Messing (2015) or Pablo Barberá (2015), lies in the identification of two main axes for interpreting the latent ideological space: the first represents the traditional left-right divide, while the second captures users' sentiments and attitudes toward the ruling elites, which can be defined as a small group of individuals or institutions that wield political, economic, social, or cultural power capable of influencing political decisions and shaping public policies (Mudde, 2004).

To ensure the reliability of this inferred space, the method was validated using a twofold approach. First, political survey data from instruments such as the Chapel Hill Expert Survey (CHES) were employed to compute and interpret the meanings of the underlying spatial dimensions. Second, the inferred ideological positions were cross-validated by comparing them with text-based analyses of users' self-descriptions in their bios. This dual validation confirmed that the estimated spatial dimensions accurately reflect the attitudinal and ideological cues conveyed in users' bios (Ramaciotti Morales et al., 2021; Ramaciotti Morales and Muñoz Zolotoochin, 2022).

To estimate the political positions of users spreading misinformation on the French Twittersphere, we identified which of them were among the 368,831 Twitter users classified in the latent ideological space constructed by Cointet et al., 2021a and Ramaciotti Morales et al., 2021. We then analyzed their distribution within it. Out of the initial 31,760 misinformation sharers, we were able to locate 10,910 accounts, representing just over 34%. This proportion is significant, considering that only 10% of the 3.7M users in the database representing the French Twittersphere were mapped within the latent ideological space. This suggests that misinformation sharers are significantly more politicized than the average user, as following MPs on Twitter indicates a strong interest in politics. Of course, this does not imply that the remaining 20,850 misinformation sharers lack political opinions—only that their political stance could not be inferred from their behavioral traces on Twitter because they did not follow enough MPs or had too few followers.

To examine the distribution of misinformation sharers in the latent ideological space, we adopted the 16-cell partitioning used by Jean-Philippe Cointet and colleagues (2021a) in their study on the Yellow Vests movement. This partitioning divides the ideological space into rectangular sections of varying sizes, each containing an equal number of accounts. The boundaries for this partitioning were determined using a k-dimensional tree (kd-tree), which successively bisects the dataset so that each resulting cell contains an equal number of accounts. At each step, the kd-tree divides the space along one dimension, seeking to balance the number of accounts in each section. This process

continues until the dataset is fully segmented. The 16-cell division was chosen because it provides politically meaningful cells while ensuring large enough sample sizes such that the empirical error on the estimated proportion of misinformation sharers (i.e., the 95% confidence interval) is at most 0.6 percentage points. This method facilitates the visualization of density variations among misinformation sharers within an initially heterogeneous space, where some areas are densely populated while others have fewer accounts.

While the entire latent ideological space contains 368,831 accounts, each rectangle, regardless of its size, represents approximately 23,000 users. If misinformation sharers were equally and indiscriminately distributed across all political communities, the 10,910 misinformation sharers in our sample would be evenly distributed across the 16 rectangles. In other words, we would expect approximately 682 misinformation sharers in each rectangle, representing 3% of all accounts in each political community. However, as shown in Figure 1, misinformation sharers are distributed very unevenly across the latent ideological space: they are highly concentrated in the lower part of the space, with an even higher density in the lower-right rectangles and, to a lesser extent, in the lower-left rectangles. Specifically, the density is 8.6% at the far left and 10% at the far right. In contrast, the lowest density is found in the center-right rectangles, where it ranges between 0.7% and 0.8%. In France, therefore, the sharing of misinformation is predominantly concentrated among internet users positioned on the far right, as observed in the United States; however, it also significantly affects a considerable portion of users on the far left.

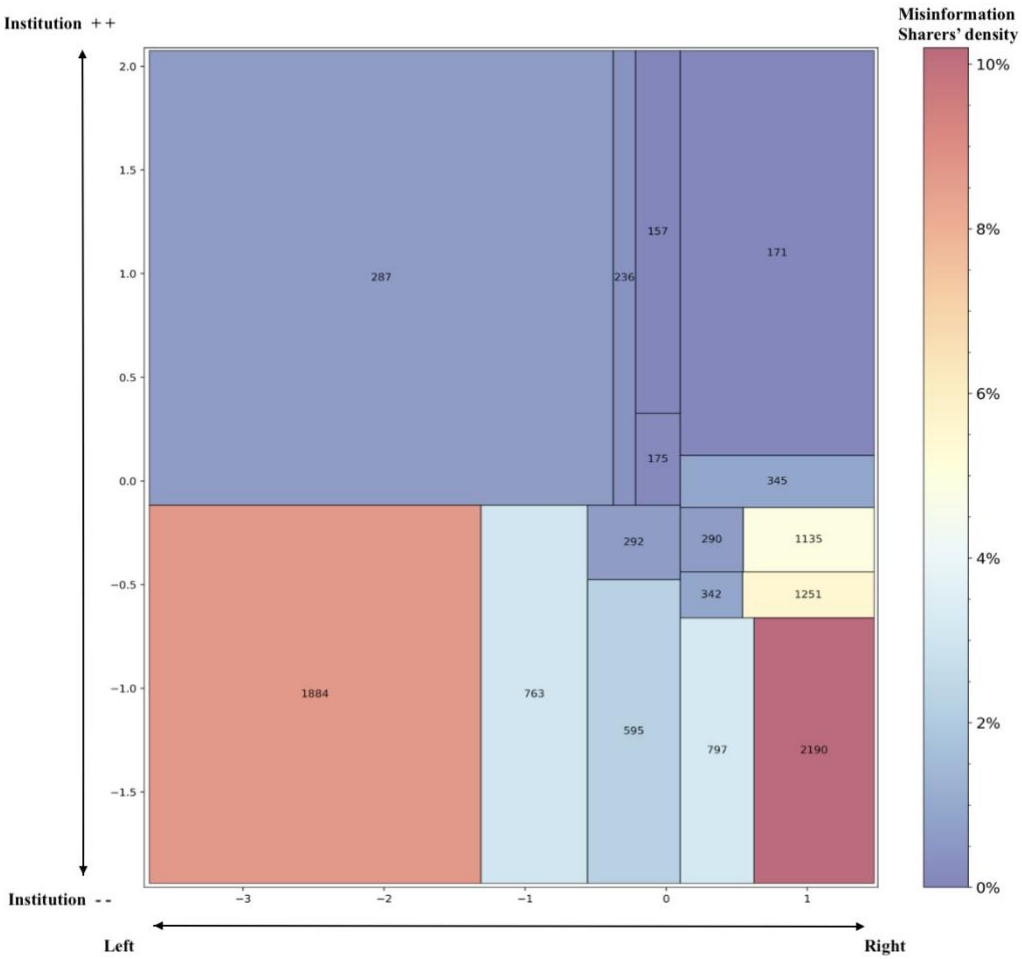

Figure 1. Distribution of Misinformation Sharers in a Latent Ideological Space split in equi-populated cells

Note. Each rectangle, regardless of its perimeter, represents approximately 23,000 Twitter accounts. The number displayed in each rectangle indicates the number of accounts that have shared at least one piece of misinformation. The warmer the color of a rectangle (orange and red), the higher the presence of misinformation sharers. Conversely, the cooler the color (blue and purple), the lower the presence of misinformation sharers.

The originality of this result lies in adding an additional dimension to the analytical framework used by American studies, highlighting that misinformation sharers, whether from the far right or far left, are characterized by a strong sense of hostility towards the ruling elites and institutions that wield political, economic, social, or cultural power. It is therefore relevant to examine the socio-political dynamics fueling the circulation of misinformation within the French Twittersphere: does this phenomenon primarily result from a traditional left-right divide, or does it reflect a broader opposition to the ruling elites and institutions?

By summing the values of the rectangles located on either side of the horizontal and vertical axes, it appears that users positioned on the right of the political spectrum (n=6,521) are 1.5 times more likely than those on the left (n=4,389) to share misinformation. Furthermore, those expressing the most pronounced anti-elite attitudes (n=9,539) are 7 times more likely to share misinformation than those least opposed to elites (n=1,371). These results highlight that the axis measuring users' attitudes toward elites is more associated with the likelihood of sharing misinformation than the left-right axis. However, this single dimension does not fully explain the positions of misinformation sharers. Indeed, the highest misinformation-sharing rates are found among individuals expressing strong anti-elite sentiments, but only when this opposition is accompanied by strong polarization along the left-right axis.

A qualitative analysis of misinformation content (see Supplementary Tables 2.1 and 2.2) reinforces the idea that misinformation sharing is associated with partisan dimensions, particularly with the political right, but even more strongly with anti-elite and identity-based rhetoric. Political topics account for 55% of the misinformation items analyzed, while respectively 48% and 23% employ anti-elite and identity-based rhetoric. For instance, claims such as “€120,000 misappropriated by Macron: The internet exposes his dirty laundry because the media won’t” combine accusations against political elites with allegations of media concealment. Identity-based claims similarly defend a restrictive conception of national or European belonging, as illustrated by the statement: “Yasmine,

who wears a hijab, cannot receive the Young European Award; she cannot be European." Together, these examples show that the circulation of misinformation is frequently embedded in partisan conflict, but also in broader opposition to political, media, and cultural elites.

These findings suggest that the spread of misinformation on Twitter can be interpreted as a symptom of a crisis in political representation, or even of representative democracy. The issue is not only dissatisfaction with particular parties or governments, but a broader sense of disconnection between citizens' concerns and the institutions that are supposed to represent them (Mineur, 2010). From this perspective, misinformation sharing can be understood as part of a broader repertoire of contestation directed not only against specific political opponents, but against the established political and media order. This interpretation also echoes the notion of a *need for chaos*, defined as "the desire for a fresh start through the destruction of order and established structures" (Petersen et al., 2023, p. 1489), and associated with feelings of social marginalization. Political disruption may thus provide a way of reclaiming status by rejecting the system as a whole rather than merely opposing one camp within it.

Misinformation sharers may therefore be characterized not only by partisan positioning, but also by a more general oppositional stance toward institutions and governing elites. This hypothesis invites further investigation into their degree of political affiliation: although they are located closer to particular political movements, especially on the far right and far left, are they formally affiliated with political parties? Do they feel represented by elected officials? Do they adhere to the substantive positions promoted by these parties, or are they primarily attracted by their anti-establishment rhetoric? Drawing on different models of partisan affiliation (Duverger, 1951; Scarrow, 2014; Gibson et al., 2018), recent qualitative research based on online ethnography and in-depth interviews identifies distinct profiles of misinformation sharers according to their degree of party commitment and identification, including elected officials, party members, activists, and independents. It suggests that integration into formal political organizations may constrain

misinformation sharing by imposing reputational costs on public expression, whereas users without partisan ties tend to share substantially more misinformation than officeholders or formally affiliated supporters (Berriche, 2024).

3. Beyond Ideology: Comparing Misinformation Sharers and Ideologically Matched Control Users

We have shown that identified misinformation sharing concerns only a minority of Twitter accounts in France and is concentrated at the extremes of the political spectrum, particularly among users strongly opposed to elites and institutions. However, political positioning alone may not account for all the differences observed between misinformation sharers and other users. Misinformation sharers may also display distinctive cognitive, affective, behavioral, informational, and identity-related characteristics.

To examine these differences, we use a matching procedure based on the 16 ideological cells described above (Rubin, 1973). This approach improves comparability by contrasting misinformation sharers with control users located in the same areas of the ideological space. It therefore allows us to assess which characteristics remain associated with misinformation sharing once political position is held approximately constant. However, the design does not support causal inference: it cannot establish whether these characteristics cause misinformation sharing, result from it, or reflect other unobserved factors.

Concretely, we first selected approximately equal numbers of misinformation sharers and potential control users across the 16 ideological cells. Potential control users had shared none of the fact-checked misinformation URLs identified in our dataset. To ensure that both groups were active in the same information-sharing environment, we then applied the same additional eligibility criterion to both groups, retaining only accounts that had shared at least one URL from one of the 420 French media outlets between March 31 and October 1, 2020. This criterion excluded a larger proportion of potential control users

than misinformation sharers, yielding a final sample of 1,907 misinformation sharers and 958 ideologically matched control users. We retained all accounts remaining after these selection steps rather than imposing a 1:1 ratio, which would have required discarding eligible misinformation sharers and reduced both statistical power and coverage of the ideological distribution.

In October 2020, we systematically collected up to the 3,200 most recent tweets from each account through the user-timeline endpoint of Twitter API v1.1[6], using the web-mining tool Minet (Plique et al., 2019).[7] This yielded 3,644,421 tweets from misinformation sharers and 459,422 tweets from control users. We also assessed whether the observed activity patterns could be driven by automated accounts (see Supplementary Section 3).

Because positioning users in the latent ideological space requires observable connections to political actors, the main sample disproportionately includes politically attentive and potentially highly active users. Some of the observed differences could therefore reflect intensive political engagement rather than misinformation sharing specifically (Valenzuela et al., 2019; Bridgman, 2024). To address this concern and provide alternative comparison baselines, we conducted three robustness checks, following the comparative logic emphasized by Baribi-Bartov et al. (2024). First, we replicated the analyses among one-time misinformation sharers to assess whether the findings were driven by recurrent sharers. Second, we constructed a sample matched one-to-one on both ideology and daily activity. Third, we examined a broader sample of ideologically unclassified users drawn from the same media-sharing ecosystem. These analyses clarify which findings extend across samples and which are more specific to politically classifiable misinformation sharers (see section 3.8 and Supplementary Section 4).

[6] https://web.archive.org/web/20210302210417/https://developer.twitter.com/en/docs/twitter-api/v1/tweets/timelines/api-reference/get-statuses-user_timeline
[7] https://doi.org/10.5281/zenodo.4679824 https://github.com/medialab/minet/blob/0.51.0/docs/cli.md#user-tweets

We then constructed measures of users' online identity, informational practices, Twitter behavior, online visibility and engagement, pseudonymous self-presentation, and psycholinguistic characteristics. The following subsections describe these measures and present the corresponding results.

3.1. Online identity

A first variable was used to examine whether users who spread misinformation on Twitter are more politicized than control users. Traditionally, the concept of politicization refers to "the attention paid to the functioning of the political field" (Gaxie, 1978, p. 240). In this case, the analysis focused on the role of politics in the digital identity presented by users (Greene, 2004).

To explore whether users' political identity takes precedence over their other social affiliations (Duchesne & Scherrer, 2003), at least on Twitter, we reviewed the bios of accounts. When opening a Twitter account, each user is invited to introduce themselves by writing a short biography of up to 160 characters. Generally, users take advantage of this space to share their interests, mention their family situation, state their profession, or share an inspiring quote. This self-description is then publicly accessible to other users. It allows users to indicate how they wish to be perceived by their audience and which aspects of their identity they want to highlight (Rogers & Jones, 2021).

After an exploratory reading of about one hundred bios, we decided to focus on the presence of political markers. A bio was considered to contain political markers if a user expressed their affiliation with a political party or support for a political figure; highlighted their ideological values or beliefs in general; or criticized opposing opinions or figures (see Figure 2 for examples of bios with and without political markers). The first author hand-coded all the bios in our sample after excluding those without any text (n=837), which were automatically categorized as non political, given the absence of content to evaluate. Subsequently, three research assistants each annotated 200 bios. To assess inter-rater reliability, we calculated Cohen's kappa coefficients. Percent agreement between raters

ranged from 89% to 90%, and kappa values ranged between 0.780 and 0.801, indicating substantial agreement.

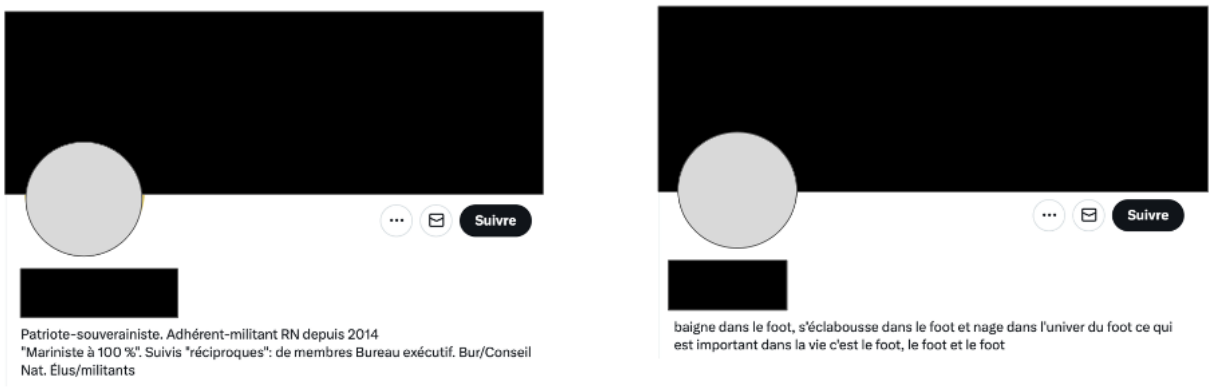

Figure 2. Examples of Twitter bios with (left) and without (right) political markers

Translation. "*Patriot and sovereigntist. RN member-activist since 2014. 100% supporter of Marine Le Pen. Mutual follows with members of the Executive Bureau / National Council. Elected officials/activists"* (bio on the left); "*Lives and breathes football, splashes around in football and swims in the world of football. What matters in life is football, football, and football*" (bio on the right).

The results show that misinformation sharers are more than twice as likely as control users (e.g. 42.1% vs. 19.4%; $\chi^2(1, N = 2{,}865) = 145.27$, $p_{adj} < .001$, $\varphi = 0.225$) to display their political opinions or party affiliation in their bio (see Figure 3). For example, hashtags and keywords referring to candidates and political parties such as Rassemblement National, Reconquête, or La France Insoumise (e.g., #RN2022, #Marine2022, #Zemmour, #JLM2022, #LFI2022) were very common in the bios of misinformation sharers. Some of them, however, seemed to prefer expressing their support for general ideological values such as patriotism or feminism, or for social movements (e.g., the Yellow Vest or the LGBTQ+ community), rather than their affiliation with specific political parties. Several accounts also displayed the hashtag #TeamPatriote in their bio, referring to the Telegram

group of the same name, created to allow users whose accounts are frequently suspended to discuss privately without content moderation.[8]

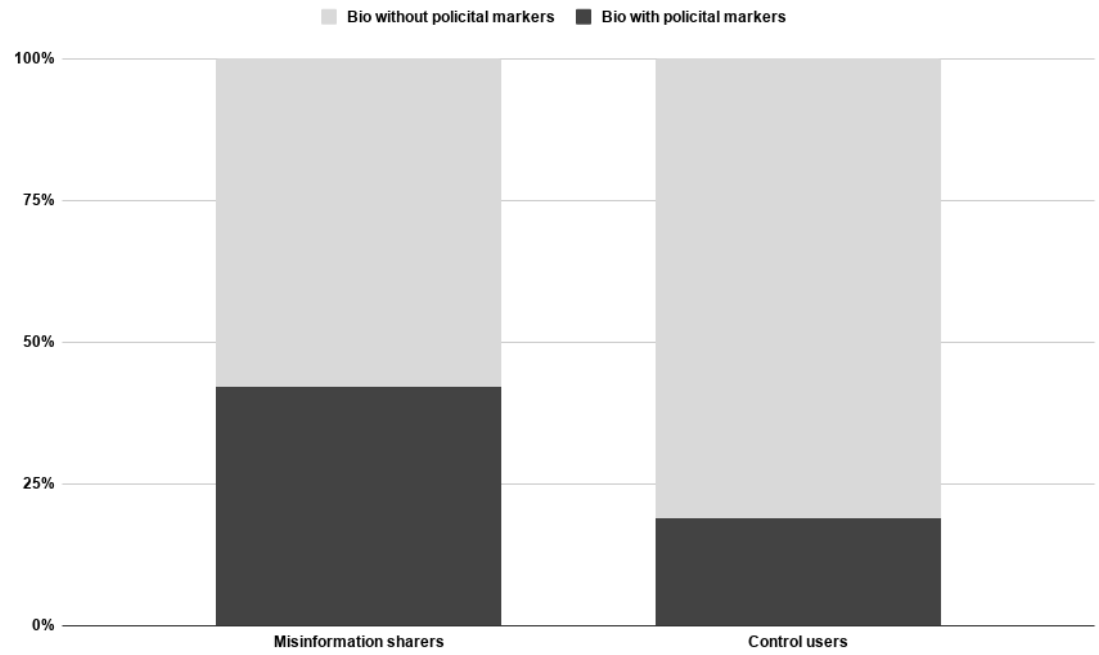

Figure 3. Proportion of Twitter bios with and without political markers depending on whether the accounts have shared at least one misinformation item

The greater tendency of misinformation sharers to highlight their party affiliation or political opinions in their Twitter profile stands out even more when compared to a representative sample of all U.S. Twitter users. A Pew Research Center study shows that only 6% of Americans users on Twitter include words in their bio suggesting opposition or support for a political party, ideology, figure, organization, or political movement (Widjaya, 2022).

[8] Laurent, S. (2020, 21 juillet). Plongée dans la haine en ligne avec « Team patriote », groupe privé de militants nationalistes. *Le Monde.* https://www.lemonde.fr/societe/article/2020/07/21/team-patriote-plongee-dans-la-haine-en- ligne_6046790_3224.html

This result invites future research to explore more in-depth and to draw on Social Identity Theory (Tajfel & Turner, 1986) to test new hypotheses to explain the sharing of misinformation. For instance, it could be relevant to explore the link between online activism and misinformation sharing by looking beyond classic measures of political position and taking into account measures of political participation (e.g. petition signing or participation in demonstrations).

The significant tendency of misinformation sharers to highlight their political opinions in their bios suggests that they already hold strong, pre-established political views. From this perspective, it is plausible that misinformation reinforces preexisting opinions more than it sparks new ones. Social media users might approach information with well-established prior convictions, and misinformation likely serves to consolidate these beliefs rather than modify or sway them (Ecker et al., 2022).

3.2. Informational practices

A second set of variables was used to examine the information-sharing practices of misinformation sharers compared to the control group. First, we measured the volume of tweets containing at least one link to an external site to study the extent to which users posted web content not related to current news. Then, we measured the volume of tweets containing at least one link redirecting to one of the 420 media sites listed in our database to examine users' propensity to post content from the French media ecosystem. The results suggest that misinformation sharers are more active in sharing links to French media outlets than control users. Indeed, 23% of their tweets contain a URL linking to a French media outlet, compared to only 14% of control users (see Table 1). In contrast, the difference in other URLs — which correspond to sites not related to current news (e.g., gaming platforms, e-commerce sites, etc.) — is statistically significant but substantively small (27% vs. 28%). These results suggest that sharing misinformation is not necessarily associated with a lack of access to current news but, on the contrary, with a marked interest in it.

To further investigate, we then analyzed the types of media sources shared by users in our sample. To do so, we used the classification of media sources carried out by a study on the structure of the French media ecosystem (Cointet et al., 2021b). The study measured the authority of 420 French media sites by examining the network of hyperlinks that connect them to each other. Based on the linking patterns among these media websites, four blocks were then distinguished:

(1) mainstream media: this block gathers the main French media of the press, radio and television, which have the largest audience in the French media space (e.g. Le Monde, Le Figaro, Europe 1, BFM TV).
(2) partisan media: it gathers websites which defend a strong political line (either right-wing or left-wing).
(3) counter information: this block hardly receives any links from the rest of the French media ecosystem even though it sends a lot of citations to the entire ecosystem. This pattern indicates that counter-information outlets occupy a marginal position within the French media hyperlink network: they link frequently to other parts of the media ecosystem but receive relatively few incoming links from more central or mainstream outlets.
(4) periphery: brings together different types of specialized media (e.g. Local Press, Technology press, Entertainment Magazines and Health Magazines & Tabloids).

The results show that misinformation sharers share a proportion of counter-information media about twice as high as the control group (15% compared to 8%) as well as slightly more partisan media (19% compared to 16%). In contrast, they share fewer mainstream media sources (52% compared to 62%), and there is little difference when it comes to local press and tabloids (see Table 1). These findings suggest that users who spread misinformation are more distrustful of mainstream media than control users and seek to bypass them by sharing alternative sources. However, they do not seem disconnected from mainstream media, as these sources still make up more than half of the media they share. This finding nuances the idea of a complete isolation in filter bubbles:

although misinformation sharers favor alternative sources, they maintain significant exposure to central public debate. As such, they do not illustrate a phenomenon of strict selective avoidance towards opposing sources of information.

These results are consistent with other studies that have shown that polarized individuals often consume large quantities of mainstream media (Garrett, 2009; Weeks et al., 2016) and do not systematically unsubscribe from sources offering perspectives that contradict their own opinions (Ross Arguedas et al., 2022).

Table 1. Distribution of Content Types and Sources Shared by Misinformation Sharers and Control Users

Measures	Variables	Misinformation sharers	Control Users	U	p_adj	effect size
Type of content	media URL	23%	14%	1,293,220.5	< .001	0.416
	other URL	27%	28%	1,001,145.5	< .001	0.096
	no URL	50%	58%	679,380.0	< .001	-0.256
Type of media sources	mainstream media	52%	62%	594,164.0	< .001	-0.350
	partisan media	19%	16%	1,300,279.5	< .001	0.424
	counter-information media	15%	8%	1,437,541.5	< .001	0.574
	tabloid/local press	14%	14%	1,122,068.0	< .001	0.228

Note. Statistical tests are conducted at the user level using Mann–Whitney U tests. P-values are adjusted for multiple comparisons using the Benjamini-Hochberg procedure. Effect size is reported as rank-biserial correlation

In summary, these findings indicate that the sharing of misinformation is not associated with a lack of access to current news but, on the contrary, with active information-sharing practices. Far from being confined to "echo chambers," misinformation sharers are exposed to media also consumed by the majority of the public. Not only do they not ignore information that contradicts their opinions, but might seek it out to better criticize it. Their recourse to alternative sources is therefore not motivated by an information deficit but by a distrust of narratives promoted by mainstream media. In other words, sharing misinformation is not symptomatic of a lack of knowledge but of a crisis of confidence in institutionalized information, particularly that of mainstream media. Ultimately, the real question researchers should ask is not so much: "Why do they adhere to misinformation?" but rather "Why are they so distrustful of traditional media?" This shift in perspective allows for a reorientation of the analysis toward dynamics of mistrust and contestation, examining the complex relationship between these users and media institutions, beyond a simple explanation based on a deficit of discernment or informational skills.

3.3. Going beyond URLs: thematic repertoires of circulated tweets

To complement the URL-level analysis, we conducted an exploratory topic-modeling analysis of the tweets posted or reshared by misinformation sharers and matched control users. The aim was to examine whether the two groups circulate different thematic repertoires beyond the URLs they share. Topics were identified using BERTopic, and their prevalence was compared across groups using smoothed log prevalence ratio and chi-square tests. Full details on corpus construction, preprocessing and model specification are provided in Supplementary Section 7.

The analysis identified 17 substantive topics. Topic -1 outliers, corresponding to 32.8% of the analyzed corpus, are not interpreted as a substantive topic. As shown in Figure 4, misinformation sharers were significantly more likely to circulate content related to Islam, Muslims, and terrorism (log-odds = 0.63, $p_{adj} < .001$), criticism of media and journalists (log-odds = 0.40, $p_{adj} < .05$), French politics and immigration (log-odds =

0.34, p_adj < .001), and police, justice, and crime (log-odds = 0.30, p_adj < .001). By contrast, matched control users were more likely to circulate football-related content (log-odds = -1.51, p_adj < .001) and casual or interpersonal conversation (log-odds = -0.44, p_adj < .001).

These results show that misinformation sharers differ not only in the sources they circulate, but also in the thematic repertoires associated with their activity. Their tweets are more strongly centered on immigration, Islam, security, crime, and criticism of mainstream media, whereas matched control users are more strongly associated with sports and everyday conversational content.

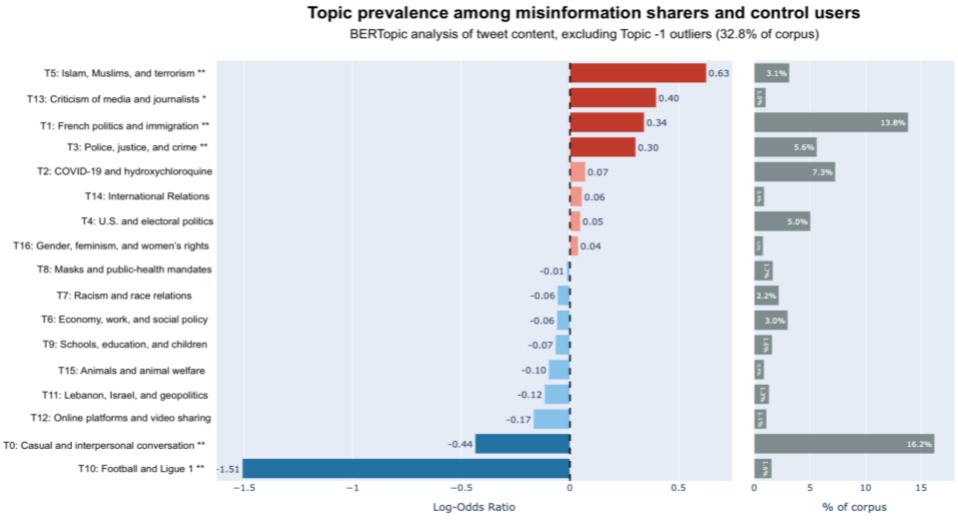

Figure 4. Topic prevalence among misinformation sharers and matched control users.

Note. The left panel reports smoothed log prevalence ratio comparing topic prevalence between misinformation sharers and matched control users; positive values indicate greater prevalence among misinformation sharers, while negative values indicate greater prevalence among matched control users. The right panel reports each topic's share of the full analyzed corpus. Darker bars indicate statistically significant differences based on BH-adjusted p-values. * p_adj < .05; ** p_adj < .001.

3.4. Behavior on Twitter

A fourth set of variables was constructed to compare the Twitter usage patterns of misinformation sharers with those of control users.

The propensity of accounts to post or share tweets over a given period was first assessed by calculating their median number of tweets per day. The results show that misinformation sharers are significantly more active on Twitter than users in the control group, with a median of 22 (re)tweets per day compared to less than one (re)tweet per day for control users. A Mann–Whitney U test confirmed a significant difference between the two groups ($U = 1{,}637{,}278.0$, $p_{adj} < .001$), with a large effect size (rank-biserial $r = 0.79$).[9]

We also constructed an indicator of high-intensity posting, or "bombardment", defined as having posted or shared at least 10 consecutive tweets within a one-hour window. Misinformation sharers stand out sharply on this measure: 92% of them meet this criterion, compared with 34% of control users ($\chi^2(1, N = 2{,}865) = 1{,}042.25$, $p_{adj} < .001$, $\varphi = 0.603$). To assess whether this result depends on the chosen threshold, we replicated the analysis using alternative cutoffs ranging from 5 to 30 tweets posted within one hour. The difference remains large and significant across all cutoffs: for example, 58.6% of misinformation sharers posted at least 30 tweets within one hour, compared with only 6.9% of control users (see Table 2). These findings confirm that misinformation sharers are not only more active on average, but also much more likely to engage in bursts of very high-intensity posting. This level of activity is substantial given that half of U.S. Twitter users post fewer than five tweets per month (Odabaş, 2022). This echoes work showing that the heaviest posters are the most partisan (Wojcik & Hughes, 2019) and that heavy social media use predicts willingness to share misinformation across countries (Valenzuela et al., 2019; Morosoli et al., 2025).

[9] We resorted to this non-parametric statistical test rather than a Student's t-test because the data from the two samples did not follow a normal distribution

Table 2. Robustness of the bombardment indicator across alternative cutoffs

Cutoff: k tweets within 1h	Misinformation sharers	Control users	Difference pp	χ^2	p_adj	φ	OR
5	97.4%	62.2%	35.2	640.43	< .001	0.473	22.35
10	91.6%	34.4%	57.1	1,042.25	< .001	0.603	20.57
15	83.4%	21.8%	61.6	1,034.33	< .001	0.601	17.92
20	75.4%	14.7%	60.7	949.35	< .001	0.576	17.70
25	66.6%	9.4%	57.2	836.74	< .001	0.540	19.13
30	58.6%	6.9%	51.7	703.91	< .001	0.496	19.01

Note. The bombardment indicator identifies users who posted at least k consecutive tweets within a one-hour window. Statistical tests are chi-square tests. P-values are adjusted using the Benjamini-Hochberg procedure. Effect size is reported as φ.

We then examined users' modes of circulation of statements on Twitter by distinguishing between original tweets, quote tweets, and retweets. These modes imply different degrees of enunciative responsibility. Original tweets position users as the direct authors of the statements they circulate; quote tweets allow them to add commentary to another user's statement; and retweets enable them to relay content without comment, acting primarily as intermediaries or amplifiers.

Original tweets account for only 23% of misinformation sharers' posts, compared to 41% for control users. Meanwhile, about 72% of misinformation sharers' posts are simple retweets, compared to 52% for control users. However, little difference is observed in the use of quote tweets, which represent a very small percentage of total posts for both misinformation sharers (5%) and control users (7%).

Table 3. Comparison of Twitter Usage Between the Two Samples

Measures	Variables	Misinformation sharers	Control Users	U	p_adj	Effect size
Tweet types	original tweet	23%	41%	645,511.0	< .001	−0.293
	retweet	72%	52%	1,183,976.0	< .001	0.296
	quote	5%	7%	902,193.5	.585	−0.012
Tweet features	hashtag	35%	31%	979,409.0	.003	0.072
	multimedia	38%	37%	854,265.0	.006	−0.065
Interactivity	reply	16%	28%	765,992.5	< .001	−0.161
	mention	31%	30%	670,055.0	.045	−0.051

Note. Percentages are computed at the tweet level. Statistical tests are conducted at the user level on users' individual proportions. Effect sizes are rank-biserial correlations. P-values are adjusted using the Benjamini-Hochberg procedure. Mentions are measured among pure original tweets only.

Furthermore, when examining how misinformation is disseminated, we find that the majority of such content is simply retweeted (80%), while only 17% originates from an original tweet. This suggests that rather than crafting their own statements to express their opinions, misinformation sharers primarily act as prolific and highly active relayers of

existing content. On the one hand, this mode of enunciation indicates a form of rapid and unreflective dissemination that prioritizes the circulation of pre-existing information. On the other hand, it suggests that misinformation sharers do not fully assume responsibility for the content they spread on Twitter—especially misinformation. To some extent, the retweeting mode can be seen as a way of distancing themselves from the shared statements, thereby avoiding direct accountability, particularly when it comes to misinformation. Moreover, while retweeting is often interpreted as an act of endorsement, this equivalence is not always straightforward. Some users even make an effort to clarify this in their bios by including statements such as "RT is not endorsement." These practices raise questions about users' awareness of the questionable nature of the information they share and suggest that retweeting could be a deliberate strategy to avoid being directly held accountable for the content they propagate.

A key feature of misinformation sharers is thus their intense and systematic use of retweets. They often retweet in rapid succession—sometimes posting up to a hundred times per day—from accounts linked to political figures, prominent polemicists, alternative media, and editorialists. Such practices reflect a commitment to promoting the political agenda of their ideological camp. They also invite us to broaden the scope of misinformation studies by examining how information—whether accurate or not—can be strategically mobilized to ideologically frame public debates and distort perceptions of reality, notably by elevating certain issues while marginalizing others on the public agenda.

This suggests that misinformation sharing on French Twitter should be understood primarily as a practice of amplification rather than production. Most misinformation sharers in our sample do not appear to create the identified misleading content themselves; instead, they contribute to its visibility by repeatedly redistributing content produced elsewhere. This pattern parallels Baribi-Bartov et al. (2024), who show that misinformation circulation on U.S. Twitter was highly concentrated among a small group of supersharers relying largely on persistent manual retweeting. This distinction has important theoretical implications: it shifts attention from the production of false claims to the collective

mechanisms through which such content is circulated and amplified within political information networks. In this sense, retweeting should not be treated as a passive or secondary behavior, but as one of the central mechanisms through which misinformation gains visibility and reaches wider audiences.

We also compare the use of different tweet features between the two groups. Misinformation sharers use hashtags slightly more frequently than control users: hashtags appear in 35% of their tweets, compared with 31% among control users. At the user level, this difference is statistically significant after correction, but the effect size is small ($U = 979{,}409.0$, $p_{adj} = .003$, rank-biserial $r = 0.072$), indicating only a weak association between hashtag use and misinformation sharing. By contrast, multimedia use is very similar across the two groups. Multimedia content appears in 38% of misinformation sharers' tweets and 37% of control users' tweets. Although the user-level test is statistically significant ($U = 854{,}265.0$, $p_{adj} = .006$, rank-biserial $r = -0.065$), the effect is very small, suggesting that multimedia use is not a distinctive feature of misinformation sharing.

Finally, we constructed interactivity indicators to assess users' propensity to engage with other accounts, either by replying to their tweets or by mentioning them when composing original tweets, notably through the use of the @ symbol, which serves as an interaction marker (Honey & Herring, 2009). The results show that misinformation sharers engage much less through replies: only 16% of their tweets are replies, compared with 28% among control users ($U = 765{,}992.5$, $p_{adj} < .001$, rank-biserial $r = -0.161$). Mentions, by contrast, do not reveal a substantively meaningful difference. Among pure original tweets, mentions appear in 31% of misinformation sharers' posts and 30% of control users' posts. The user-level difference is statistically significant but very small and goes in the opposite direction once individual-level distributions are taken into account ($U = 670{,}055.0$, $p_{adj} = .045$, rank-biserial $r = -0.051$). We therefore interpret reply behavior, rather than mention use, as the more meaningful indicator of reduced interactional engagement among misinformation sharers.

Overall, misinformation sharers appear to prefer broadcasting tweets indiscriminately to an undefined audience rather than addressing specific recipients to engage in conversation. This suggests that their primary goal is to make their voice heard and position themselves in the public sphere rather than engage in direct and interpersonal interactions. This stance is reminiscent of the "ego-visible" users on Facebook (Bastard et al., 2017), who do not use social networks to create a broad and open conversational space but rather to display their identity and share content without fully engaging in interactive exchanges that they do not control. This low level of interactivity highlights a particular expressive dynamic among misinformation sharers: they seem less interested in dialogue and more focused on broadcasting their opinions, reflecting a unilateral conception of online communication—closer to proclamation than conversation.

In summary, misinformation sharers are substantially more active than ideologically matched control users and rely heavily on retweeting, suggesting that misinformation sharing operates largely through amplification. By contrast, differences in hashtag use, multimedia content, and mentions are statistically significant but negligible, and should not be interpreted as central features of misinformation sharing. The lower proportion of replies, however, points to a more meaningful difference in interactional style: misinformation sharers appear less oriented toward direct conversation and more oriented toward one-way broadcasting.

3.5. Online visibility and engagement

A fifth set of variables compares the online visibility and engagement of misinformation sharers and control users on Twitter.

Misinformation sharers have larger typical audiences than control users. Their median follower count is 852, compared with 94.5 among control users. Although mean follower counts are similar because of highly influential outliers in the control group (1,995.9 vs. 2,074.2), the user-level difference is large and statistically significant (U =

1,514,837.0, p_adj < .001, rank-biserial r = 0.658). Misinformation sharers also follow more accounts, with a median following count of 1,195 compared with 358.5 among control users (U = 1,401,425.0, p_adj < .001, r = 0.534). Their follower/following ratio is likewise higher at the median (0.771 vs. 0.253; U = 1,394,663.5, p_adj < .001, r = 0.527), suggesting that they occupy a more visible position in the network.

However, this visibility does not translate into stronger endorsement through likes. Misinformation sharers receive fewer likes per original tweet than control users (median = 0.192 vs. 0.310; U = 809,913.0, p_adj < .001, r = -0.113), and a smaller proportion of their original tweets receive at least one like (7.4% vs. 14.3%; U = 748,727.5, p_adj < .001, r = -0.180). These effects are modest, but they indicate that misinformation sharers are not more positively endorsed by their audience.

Retweet-based engagement follows a different pattern. Among original tweets, misinformation sharers receive more retweets at the median than control users (0.276 vs. 0.120; U = 1,063,667.0, p_adj < .001, r = 0.270), and a larger proportion of their original tweets receive at least one retweet (14.3% vs. 7.8%; U = 1,056,121.0, p_adj < .001, r = 0.261). Thus, while misinformation sharers are not more liked, their own tweets are more likely to be recirculated.

These results point to a distinction between visibility, endorsement, and recirculation. Misinformation sharers have larger typical audiences and follow more accounts, but their original tweets receive fewer likes than those of control users. Their original tweets, however, are more likely to be retweeted, suggesting that their activity circulates more through retweet-based diffusion than through positive endorsement.

3.6. Use of pseudonyms

A sixth variable was introduced to compare the propensity of misinformation sharers to conceal their identity behind a pseudonym compared to control users. All accounts were manually annotated by the first author of the article. We considered that a user disclosed their full name when both a first name and a last name appeared on their profile. In contrast, when an account used only a first name, a nickname, an abbreviation, an object, a flower, an animal name, or an adjective, we classified it as using a pseudonym. For example, usernames such as "Pierre Durand" or "Latifa Mansouri" were categorized as full names, while "Revolutionary415" or "MusicJunkie87" were classified as pseudonyms. It is important to note that some accounts may have deliberately chosen first and last names different from their legal identity, introducing a certain margin of error in our annotations.

While nearly two-thirds of users who shared misinformation used a pseudonym, this was the case for less than half of control users (65.0% vs. 45.1%; $\chi^2(1, N = 2{,}865) = 104.22$, $p_{adj} < .001$, $\varphi = 0.191$). These findings indicate that Twitter accounts responsible for sharing misinformation are more likely than control users to conceal their identity behind a pseudonym. This practice can be seen as a risk-management strategy, allowing them to engage vigorously in online debates without endangering their reputation. Indeed, by decoupling their online statements from their civil identity, misinformation sharers reduce personal accountability and guard themselves against potential repercussions.

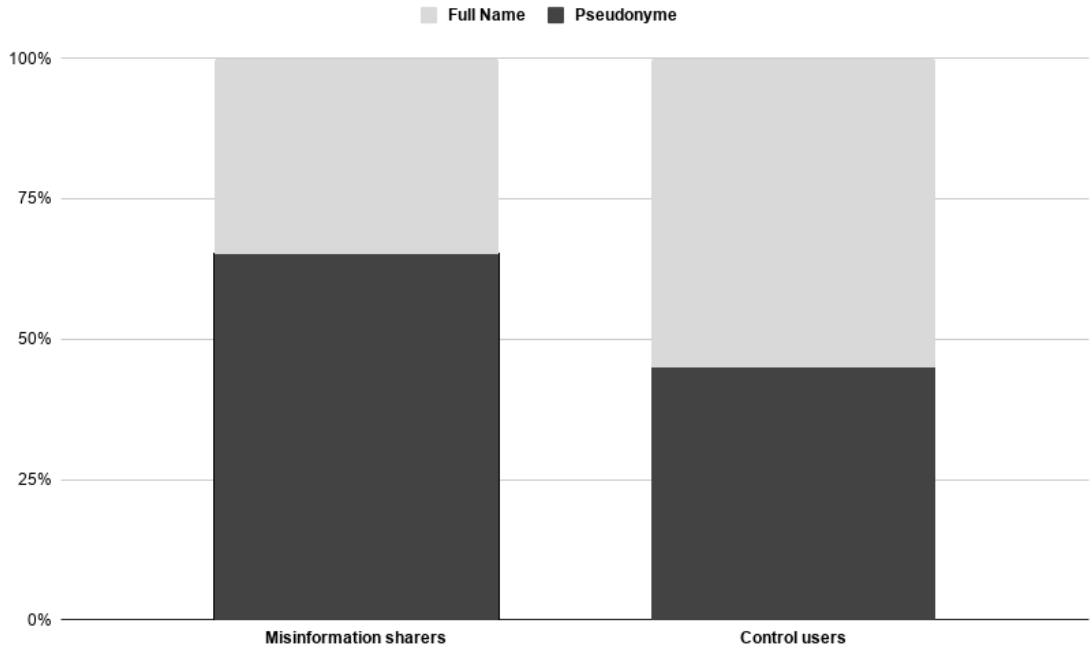

Figure 5. Proportion of Twitter accounts providing a full name or using a pseudonym, depending on whether the accounts have shared at least one misinformation item

3.7. Psycholinguistic traits

A seventh set of variables was defined to compare the language used by misinformation sharers with that of users in the control group. For this purpose, the LIWC (Linguistic Inquiry and Word Count) software was employed. Developed by psycholinguists, LIWC quantifies the relative frequency of words assigned to predefined linguistic and psychologically relevant lexical categories (Pennebaker et al., 2001; Pennebaker et al., 2015). The software compares the words contained in a text with predefined dictionaries grouping words into theoretically meaningful categories, such as affective or cognitive processes, and computes the percentage of words associated with each category. We used the French LIWC dictionary developed by Piolat et al. (2011), derived from LIWC 2007, which remains the French LIWC dictionary most clearly documented in the academic literature.

Although LIWC has important limitations when applied to short-form social media texts (e.g. it may fail to capture contextual meaning, irony, sarcasm, slang, emojis, or platform-specific conventions, Franklin, 2015; Schwartz & Ungar, 2015), it remains widely used in psychology, communication, and computational social science (for examples, see Bail et al., 2017; Vosoughi et al., 2018). More specifically, previous studies of social media communication and misinformation-related phenomena have used LIWC to examine the language of conspiracy-minded users (Samory & Mitra, 2018; Fong et al., 2021), accounts sharing unreliable sources (Mosleh et al., 2021), and users commenting on or reacting to fake news (Jiang & Wilson, 2018; Metzger et al., 2021). Using LIWC therefore allows our analysis to speak directly to this literature and to assess whether similar linguistic patterns can be observed among misinformation sharers in the French Twittersphere. Furthermore, LIWC has also been used and validated in French-language social media contexts closely related to the present study. Garcia and Rimé (2019), for instance, validated the applicability of the French LIWC dictionary to Twitter data by comparing LIWC scores with manually annotated French tweets posted after the November 2015 Paris terrorist attacks. Other studies have similarly applied LIWC to French Twitter

data, including analyses of #BalanceTonPorc and #MeToo (Lopez et al., 2019) and of the mental-health effects of COVID-19 lockdowns in France (Colella et al., 2023). These studies suggest that, despite its limitations, LIWC can provide useful lexical indicators for the analysis of French social media data, provided that its results are interpreted cautiously.

Following this literature, we focused on two broad families of LIWC categories: cognitive-process indicators and affective indicators. The former capture the relative presence of terms associated with reasoning, causality, certainty, or insight, while the latter capture the relative presence of positive and negative emotional language. Examples of words associated with each category are provided in Table 4. To minimize interpretative bias, LIWC analyses were conducted only on original tweets authored by the users themselves — as in the study by Fong and colleagues (2021). Retweets and quote tweets were excluded in order to avoid attributing to users the linguistic properties of content written by others. This filtering reduced the dataset from approximately 4 million tweets to 1.3 million original tweets, and from 2,865 to 2,759 users. LIWC scores were then computed on these original tweets and used to compare misinformation sharers with users in the matched control group.

Table 4. Examples of words for each of the LIWC categories used

Psychological processes	Category	Examples
Cognitive processes	Cognition	exactly, precisely, right
	Insight	think, know, consider
	Causation	because, effect, hence
	Certainty	always, never
	Inhibition	block, constrain
Affective processes	Emotional tone	incredible, wonderful, lol
	Affect	cry, ugly, bitter
	Positive Emotions	happy, pretty, good
	Negative Emotions	hate, worthless, enemy
	Anxiety	nervous, afraid, tense
	Anger	hate, kill, pissed
	Sadness	sad, grief, crying

The analysis of linguistic measures reveals nuanced differences between misinformation sharers and control users.

In terms of cognition, misinformation sharers show lower overall cognitive-process scores than control users, although the difference is substantively small (median = 9.33 vs. 9.82; U = 779,215.5, p_adj = .005, rank-biserial r = −0.070). Among specific cognitive indicators, insight, causation, and certainty do not differ significantly after correction.

Misinformation sharers display higher inhibition scores than control users, suggesting somewhat greater linguistic markers of restraint or caution (median = 0.65 vs. 0.565; U = 962,634.5, p_adj < .001, r = 0.149), as well as marginally higher analytic thinking scores (median = 30.91 vs. 30.00; U = 886,098.0, p_adj = .019, r = 0.058). By contrast, agency is lower among misinformation sharers (median = 36.04 vs. 38.60; U = 720,734.5, p_adj < .001, r = −0.139). Overall, these cognitive-process differences are statistically detectable but modest. They should therefore be interpreted cautiously: rather than pointing to a distinctive cognitive style among misinformation sharers, they suggest broad similarity with control users on most cognitive LIWC indicators. This pattern is consistent with survey evidence showing limited differences in cognitive processing between users who are more and less willing to share misinformation (Morosoli et al., 2025).

By contrast, affective indicators show clearer and more substantively interpretable differences. Misinformation sharers exhibit a markedly lower emotional tone than control users (median = 0.11 vs. 0.63; U = 598,896.5, p_adj < .001, r = −0.285) and lower positive emotion scores (median = 2.04 vs. 2.36; U = 699,789.5, p_adj < .001, r = −0.164). They display higher levels of negative emotion (median = 1.88 vs. 1.565; U = 1,014,601.5, p_adj < .001, r = 0.212), anxiety (median = 0.23 vs. 0.15; U = 1,047,939.0, p_adj < .001, r = 0.251), anger (median = 0.64 vs. 0.46; U = 1,049,446.5, p_adj < .001, r = 0.253), and sadness (median = 0.52 vs. 0.42; U = 1,016,819.0, p_adj < .001, r = 0.214). The distinction between anxiety and anger is especially relevant, since prior work suggests that anger may be more closely associated with sharing dynamics, whereas anxiety is more strongly tied to information-seeking behavior (Han et al., 2023). In this respect, the higher anger scores observed among misinformation sharers are consistent with the idea that misinformation sharing is linked not only to exposure to negative content, but also to emotionally charged forms of political circulation.

Overall, the LIWC results show a clear asymmetry between cognitive and affective indicators. Cognitive-process differences are either non-significant or statistically significant with very small effect sizes, and should therefore not be treated as substantively central. In contrast, affective indicators — especially negative emotion, anger, anxiety, and

emotional tone — display larger and more interpretable differences. The main psycholinguistic distinction between misinformation sharers and ideologically matched control users is therefore primarily affective rather than cognitive.

Table 5. Comparison of the psycholinguistic styles of the two samples

Measures	Variables	Misinformation sharers	Control users	U	p_adj	Effect size
Cognitive style	cognition	9.33	9.82	779,215.5	.005	−0.070
	insight	1.35	1.34	855,222.0	.428	0.021
	causation	1.46	1.49	843,628.0	.755	0.007
	certainty	1.35	1.40	830,525.0	.755	−0.008
	inhibition	0.65	0.565	962,634.5	< .001	0.149
	analytic thinking	30.91	30.00	886,098.0	.019	0.058
	agency	36.04	38.60	720,734.5	< .001	−0.139
	emotional tone	0.11	0.63	598,896.5	< .001	−0.285

Measures	Variables	Misinformation sharers	Control users	U	p_adj	Effect size
Affect and emotion	affect	4.03	4.14	797,739.5	.054	−0.048
	positive emotion	2.04	2.36	699,789.5	< .001	−0.164
	negative emotion	1.88	1.565	1,014,601.5	< .001	0.212
	anxiety	0.23	0.15	1,047,939.0	< .001	0.251
	anger	0.64	0.46	1,049,446.5	< .001	0.253
	sadness	0.52	0.42	1,016,819.0	< .001	0.214

Note. Values reported for both groups are medians of user-level LIWC scores computed from original tweets only. Statistical tests are Mann–Whitney U tests conducted at the user level. Effect sizes are rank-biserial correlations. P-values are adjusted using the Benjamini–Hochberg procedure.

3.8. Robustness across alternative comparison samples

Three robustness checks qualify the findings from the main ideology-matched sample (see Supplementary Section 4). First, among one-time misinformation sharers, differences in political self-presentation, media-oriented information sharing, high-intensity posting, and negative affect remain visible, although generally with smaller effect

sizes. Retweet orientation is attenuated and no longer significant after correction for multiple comparisons, while the composition of media sources is closer to that of control users.

Second, in the sample of 454 pairs matched on both ideology and daily activity, where median activity is nearly identical across groups (3.17 vs. 3.12 tweets per day), misinformation sharers remain more oriented toward media, partisan, and counter-informational sources, more reliant on retweet-based amplification, less likely to reply, and more likely to display political and pseudonymous self-presentation, engage in high-intensity posting, and use negative-affect language.

Third, the ideologically unclassified sample provides a broader comparison baseline for behavioral, informational, identity, and engagement measures. Differences in political self-presentation, source repertoires, and retweet-oriented behavior remain, but the average daily activity gap disappears, high-intensity posting differences are substantially weaker, and the association with pseudonymous self-presentation reverses direction.

Taken together, these checks indicate that informational repertoires and several amplification-related practices extend most consistently across samples, whereas average hyperactivity and pseudonym use are more dependent on the population and comparison baseline considered.

4. Discussion

4.1. Main findings and contributions

Our analyses reveal a gap between how misinformation sharers are portrayed in public discourse and the profile that emerges from behavioral data. Misinformation sharing does not appear to be driven by a broad crowd of gullible users. Rather, in the French Twittersphere, it is concentrated among a small minority of accounts located at the extremes of the political spectrum and, above all, strongly opposed to ruling elites and institutions. Compared with ideologically matched control users, misinformation sharers display a more pronounced political identity, more intensive information-sharing practices, higher levels of activity, stronger retweet-oriented behavior, and more frequent negative emotional expression. Their language, however, does not indicate lower analytical thinking. These findings contribute to the literature in two ways.

Methodologically, the article shows how digital trace data can be used to describe the distinctive characteristics of misinformation sharers with greater precision. Building on computational social science approaches to ideological inference (Cointet et al., 2021a; Ramaciotti Morales et al., 2021), we estimate users' political positions from their proximity to political actors, a strategy that is particularly relevant for studying politically attentive Twitter users (Tufekci, 2014; Hargittai, 2015; Barberá & Rivero, 2015; Wojcik & Hughes, 2019; Wojcieszak et al., 2022). The matching procedure then allows us to compare misinformation sharers with ideologically matched control users, thereby holding political positioning constant and identifying which characteristics remain distinctive beyond ideological location.

Theoretically, these findings move the analysis beyond explanatory models centered solely on cognitive variables, such as deficits in analytical reasoning. They are

more consistent with approaches emphasizing motivated reasoning, political polarization, and anti-elite hostility, while also showing the value of "mere description" for refining explanatory models of misinformation sharing (Gerring, 2012; Munger et al., 2021). More specifically, our results suggest that the political motivations of misinformation sharers are not necessarily reducible to partisan attachment. They may also reflect a broader oppositional stance toward ruling elites and institutions. This invites further research on the degree of political affiliation of misinformation sharers: whether they are party members, feel represented by elected officials, adhere to the ideas promoted by these political groups, or primarily support them because of their opposition to the established political order. Such questions are central to understanding the socio-political mechanisms that promote or limit misinformation sharing, as suggested by qualitative work showing that integration into political and social spaces can reduce misinformation sharing by imposing reputational constraints on public expression (Berriche, 2024).

4.2. Limitations and future research

The contribution of this study should be interpreted within a clearly defined empirical scope. The dataset reconstructs the French media-sharing Twittersphere from users who circulated URLs from 420 French media outlets. It therefore captures an important information-sharing ecosystem, but does not represent all Twitter users in France and does not capture misinformation circulated without external links, for example through text, images, videos, screenshots, or memes. Our findings consequently concern identified misinformation sharing within this media-sharing ecosystem rather than the entire French Twitter population.

Our identification strategy is conservative rather than exhaustive. Misinformation sharers are users who circulated at least one individual URL previously assessed as false or misleading by professional fact-checkers; entire outlets or domains are not classified as misinformation sources. This criterion identifies demonstrable instances of misinformation

sharing but necessarily misses claims that were not fact-checked or did not contain external links.

The ideological-inference procedure also shapes the primary population examined. The main comparison concerns politically classifiable, media-active users whose follow networks allowed them to be located in the latent ideological space. This provides a demanding comparison baseline, since control users were active in the same media-sharing ecosystem and located in the same ideological cells as misinformation sharers, but less politically visible or occasional sharers may display different patterns. As shown in Section 3.8 and Supplementary Section 4, the alternative samples partly address this scope condition: informational repertoires and several amplification-related practices extend across comparison baselines, whereas average activity and pseudonymous self-presentation are more population-specific.

The matching procedure improves comparability but does not establish causality. Matching within ideological cells allows us to examine which characteristics remain associated with misinformation sharing once political position is held approximately constant. It cannot determine whether these characteristics cause misinformation sharing, result from it, or reflect unobserved factors. The findings therefore do not definitively adjudicate between political and cognitive explanations. Rather, they show that political self-presentation, informational practices, amplification behavior, and affective language distinguish misinformation sharers more consistently than lexical indicators associated with cognitive processing.

The psycholinguistic findings should also be interpreted cautiously. LIWC measures the relative frequency of words belonging to predefined lexical categories; it does not directly measure cognitive abilities, motivations, or emotional states. Dictionary-based methods are also limited in their treatment of slang, irony, sarcasm, emojis, hashtags, and politically coded language. The small and inconsistent differences in cognitive-process categories should therefore not be treated as evidence against cognitive explanations of

misinformation sharing. The affective pattern is more consistent, including among one-time sharers and after matching on activity, but future research should combine lexical indicators with contextual language models and manual validation.

Finally, the data describe Twitter before its transformation into X. Subsequent changes in ownership, data access, verification, moderation, recommendation systems, and user composition limit direct extrapolation to the current platform. Replication on X and comparative research across platforms would help determine whether the concentration of misinformation sharing, the importance of anti-elite positioning, and the predominance of amplification over original production persist under different technical and governance conditions.

Overall, these scope conditions delimit rather than undermine the contribution of the study. The article does not propose a universal profile of everyone who may share misinformation. It provides a detailed description of identified misinformation sharers within the French media-sharing Twittersphere and uses alternative comparison samples to distinguish the most stable findings from those that are more population-specific.

Conclusion

This article provides a descriptive account of misinformation sharing in the French Twittersphere. By combining ideological inference and a matching procedure, it shows that misinformation sharing is concentrated among a small minority of users located at the political extremes and, above all, strongly opposed to ruling elites and institutions.

Compared with ideologically matched control users, misinformation sharers display stronger political self-presentation, a more media-oriented and less mainstream-centered informational repertoire, greater reliance on retweet-based amplification, higher levels of activity, more frequent pseudonymous self-presentation, and more negative emotional expression.

These findings shift attention away from explanations centered solely on individual gullibility and toward the political and communicational contexts in which misinformation is circulated. They also point to amplification, rather than original production, as a central mechanism of misinformation sharing. Future research should examine how these patterns vary across political systems, user populations, platforms, and the transformed technical and governance architecture of X and other contemporary social media.

Acknowledgements

We would like to thank the two anonymous reviewers for their generous and constructive feedback, and Andy Guess for a particularly thoughtful and efficient editorial process. We are also grateful to the many colleagues and researchers who read earlier versions of the manuscript or attended presentations of this work at seminars and workshops, and whose comments and suggestions helped us improve the paper.

Funding Information

This work has been partially funded by the Reboot Foundation and by project De Facto Observatory of Information (Connecting Europe Facility - 2020-FR-IA-0291).

References

Bail, C. A., Brown, T. W., & Mann, M. (2017). Channeling hearts and minds: Advocacy organizations, cognitive-emotional currents, and public conversation. *American Sociological Review*, *82*(6), 1188-1213.

Bakker, R., Jolly, S., & Polk, J. (2012). Complexity in the European party space: Exploring dimensionality with experts. *European Union Politics*, *13*(2), 219-245.

Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political analysis*, *23*(1), 76-91.

Barberá, P., & Rivero, G. (2015). Understanding the political representativeness of Twitter users. *Social Science Computer Review*, *33*(6), 712-729.

Baribi-Bartov, S., Swire-Thompson, B., & Grinberg, N. (2024). Supersharers of fake news on Twitter. *Science*, *384*(6699), 979-982.

Bastard, I., Cardon, D., Charbey, R., Cointet, J. P., & Prieur, C. (2017). Facebook, pour quoi faire?. *Sociologie*, *8*(1), 57-82.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior?. *Perspectives on psychological science*, *2*(4), 396-403.

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.

Benoit, K., & Laver, M. (2012). The dimensionality of political space: Epistemological and methodological considerations. *European Union Politics*, *13*(2), 194-218.

Berriche, M., Do, S., & Tainturier, B. (2020). Consentir et critiquer. Les réactions sur Twitter face à la crise de la Covid-19 en France. In *Le monde d'aujourd'hui. Les sciences sociales au temps de la Covid* (pp. 223-240). Presses de Sciences Po.

Berriche, M. (2024). *Tu crois que c'est vrai? Diversité des régimes d'énonciation face aux fake news et mécanismes d'autorégulation conversationnelle* (Doctoral dissertation, Université Paris Cité).

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 US Presidential election online discussion. *First monday*, *21*(11-7).

Bond, R., & Messing, S. (2015). Quantifying social media's political space: Estimating ideology from publicly revealed preferences on Facebook. *American Political Science Review*, *109*(1), 62-78.

Bridgman, A. (2024). *Studies in digital politics: activism, mobilization, and misinformation*. McGill University (Canada).

Cardon, D., Cointet, J. P., Ooghe, B., & Plique, G. (2019). Unfolding the multi-layered structure of the French mediascape.

Cointet, J. P., Morales, P. R., Cardon, D., Froio, C., Mogoutov, A., Ooghe, B., & Plique, G. (2021a). What colours are the yellow vests? an ideological scaling of facebook groups. *Statistique et Société*, *9*(1-2), 79-107.

Cointet, J. P., Cardon, D., Mogoutov, A., Ooghe-Tabanou, B., Plique, G., & Morales, P. (2021b). Uncovering the structure of the French media ecosystem. *arXiv preprint arXiv:2107.12073*.

Colella, S., Dufourt, F., Hildebrand, V. A., & Vivès, R. (2023). Mental health effects of COVID-19 lockdowns: a Twitter-based analysis. *Economics & Human Biology*, *51*, 101307.

Davis, C. A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. (2016, April). Botornot: A system to evaluate social bots. In *Proceedings of the 25th international conference companion on world wide web* (pp. 273-274).

de Oliveira, D. V. B., & Albuquerque, U. P. (2021). Cultural evolution and digital media: Diffusion of fake news about COVID-19 on Twitter. *SN computer science*, *2*(6), 430.

DeVerna, M. R., Aiyappa, R., Pacheco, D., Bryden, J., & Menczer, F. (2022). Identification

and characterization of misinformation superspreaders on social media. *CoRR*.

Duchesne, S., & Scherrer, V. (2003). L'identité politique comme force de conflictualisation et de hiérarchisation des appartenances sociales : justification théorique d'une définition empirique. In *Identité (s)* (pp. 325-336). Presses Universitaires de Rennes

Duverger, M. (1951). Adhérents et électeurs des partis. *Revue française de science politique*, *1*(1/2), 56-75.

Ecker, U. K., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., ... & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, *1*(1), 13-29.

Fong, A., Roozenbeek, J., Goldwert, D., Rathje, S., & Van Der Linden, S. (2021). The language of conspiracy: A psychological analysis of speech used by conspiracy theorists and their followers on Twitter. *Group Processes & Intergroup Relations*, *24*(4), 606-623.

Franklin, E. (2015, September). Some theoretical considerations in off-the-shelf text analysis software. In *Proceedings of the Student Research Workshop* (pp. 8-15).

Garcia, D., & Rimé, B. (2019). Collective emotions and social resilience in the digital traces after a terrorist attack. *Psychological science*, *30*(4), 617-628.

Garrett, R. K. (2009). Echo chambers online?: Politically motivated selective exposure among Internet news users. *Journal of computer-mediated communication*, *14*(2), 265-285.

Gaumont, N., Panahi, M., & Chavalarias, D. (2018). Reconstruction of the socio-semantic dynamics of political activist Twitter networks—Method and application to the 2017 French presidential election. *PloS one*, *13*(9), e0201879.

Gaxie, D. (1978). *Le cens caché: inégalités culturelles et ségrégation politique*. Éditions du Seuil.

Gerring, J. (2012). Mere description. *British Journal of Political Science*, *42*(4), 721-746.

Gibson, R., Greffet, F., & Cantijoch, M. (2018). Friend or foe? Digital technologies and the changing nature of party membership. In *Digital Politics: Mobilization, Engagement and Participation* (pp. 89-111). Routledge.

Greene, S. (2004). Social identity theory and party identification. *Social science quarterly*, *85*(1), 136-153.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science*, *363*(6425), 374-378.

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science advances*, *5*(1), eaau4586.

Han, J., Lee, S. E., & Cha, M. (2023). The secret to successful evocative messages: Anger takes the lead in information sharing over anxiety. *Communication Monographs*, *90*(4), 545-565.

Hargittai, E. (2015). Is bigger always better? Potential biases of big data derived from social network sites. *The ANNALS of the American Academy of Political and Social Science*, *659*(1), 63-76.

He, L., He, C., Reynolds, T. L., Bai, Q., Huang, Y., Li, C., ... & Chen, Y. (2021). Why do people oppose mask wearing? A comprehensive analysis of US tweets during the COVID-19 pandemic. *Journal of the American Medical Informatics Association*, *28*(7), 1564-1573.

Hix, S., Noury, A., & Roland, G. (2006). Dimensions of politics in the European Parliament. *American Journal of Political Science*, *50*(2), 494-520.

Honey, C., & Herring, S. C. (2009, January). Beyond microblogging: Conversation and collaboration via Twitter. In *2009 42nd Hawaii International Conference on System Sciences* (pp. 1-10). Ieee.

Hopp, T., Ferrucci, P., & Vargo, C. J. (2020). Why do people share ideologically extreme, false, and misleading content on social media? A self-report and trace data–based analysis of countermedia content dissemination on Facebook and Twitter. *Human Communication Research*, *46*(4), 357-384.

Jiang, S., & Wilson, C. (2018). Linguistic signals under misinformation and fact-checking: Evidence from user comments on social media. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 1-23.

Keller, T. R., & Klinger, U. (2019). Social bots in election campaigns: Theoretical, empirical, and methodological implications. *Political Communication*, *36*(1), 171-189.

Kennedy, I., Wack, M., Beers, A., Schafer, J. S., Garcia-Camargo, I., Spiro, E. S., & Starbird, K. (2022). Repeat spreaders and election delegitimization: A comprehensive dataset of misinformation tweets from the 2020 US election. *Journal of Quantitative Description: Digital Media*, *2*.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A. L., Brewer, D., ... & Van Alstyne, M. (2009). Computational social science. *Science*, *323*(5915), 721-723.

Lopez, I., Quillivic, R., Evans, H., & Arriaga, R. I. (2019, August). Denouncing sexual violence: A cross-language and cross-cultural analysis of# MeToo and# BalanceTonPorc. In *IFIP Conference on Human-Computer Interaction* (pp. 733-743). Cham: Springer International Publishing.

Messing, S., DeGregorio, C., Hillenbrand, B., King, G., Mahanti, S., Mukerjee, Z., ... & Wilkins, A. (2020). Facebook privacy-protected full URLs data set.

Metzger, M. J., Flanagin, A. J., Mena, P., Jiang, S., & Wilson, C. (2021). From dark to light: The many shades of sharing misinformation online. *Media and Communication*, *9*(1), 134-143.

Mineur, D. (2010). *Archéologie de la représentation politique*. Presses de Sciences Po.

Morales, P. R., Cointet, J. P., Benbouzid, B., Cardon, D., Froio, C., Metin, O. F., ... & Plique, G. (2021). Atlas multi-plateforme d'un mouvement social: le cas des Gilets jaunes. *Statistique et société*, *9*, 39-77.

Morosoli, S., Van Aelst, P., Humprecht, E., Staender, A., & Esser, F. (2025). Identifying the drivers behind the dissemination of online misinformation: A study on political attitudes and individual characteristics in the context of engaging with misinformation on social media. *American Behavioral Scientist*, *69*(2), 148-167.

Mosleh, M., Pennycook, G., Arechar, A. A., & Rand, D. G. (2021). Cognitive reflection correlates with behavior on Twitter. *Nature communications*, *12*(1), 921.

Mosleh, M., & Rand, D. G. (2022). Measuring exposure to misinformation from political elites on Twitter. *Nature Communications*, *13*(1), 7144.

Mudde, C. (2004). The populist zeitgeist. *Government and opposition*, *39*(4), 541-563.

Munger, K., Guess, A. M., & Hargittai, E. (2021). Quantitative description of digital media: A modest proposal to disrupt academic publishing. *Journal of Quantitative Description*, (1), 1-13.

Odabaş, M. (2022, March 16). 5 facts about Twitter 'lurkers'. *Pew Research Center*.

Ooghe-Tabanou, B., Mazoyer, B., Farjas, J., & Plique, G. (2022). *Gazouilloire: A command-line tool for long-term collections of tweets* [Computer software].

Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2021). Partisan polarization is the primary psychological motivation behind political fake news sharing on Twitter. *American political science review*, *115*(3), 999-1015.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of LIWC2015.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2001). Linguistic inquiry and word count: LIWC [Computer software]. *Austin, TX: liwc. net*, *135*.

Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in cognitive sciences*, *25*(5), 388-402.

Petersen, M. B., Osmundsen, M., & Arceneaux, K. (2023). The "need for chaos" and motivations to share hostile political rumors. *American Political Science Review*, *117*(4), 1486-1505.

Piolat, A., Booth, R. J., Chung, C. K., Davids, M., & Pennebaker, J. W. (2011). The French dictionary for LIWC: Modalities of construction and examples of use. *Psychologie française*, *56*(3), 145-159.

Plique, G., Breteau, P., Farjas, J., Théro, H., Descamps, J., Pellé, A., & Miguel, L. (2019). Minet, a webmining CLI tool & library for python.

Ramaciotti Morales, P., Cointet, J. P., & Zolotoochin, G. M. (2021, November). Unfolding the dimensionality structure of social networks in ideological embeddings. In *Proceedings of the 2021 IEEE/ACM international conference on advances in social networks analysis and mining* (pp. 333-338).

Ramaciotti Morales, P., & Muñoz Zolotoochin, G. (2022, February). Measuring the accuracy of social network ideological embeddings using language models. In *International Conference on Information Technology & Systems* (pp. 267-276). Cham: Springer International Publishing.

Ramaciotti Morales, P., Berriche, M., & Cointet, J. P. (2023, June). The geometry of misinformation: embedding twitter networks of users who spread fake news in geometrical opinion spaces. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 17, pp. 730-741).

Rathje, S., He, J. K., Roozenbeek, J., Van Bavel, J. J., & Van Der Linden, S. (2022). Social media behavior is associated with vaccine hesitancy. *PNAS nexus*, *1*(4), pgac207.

Rauchfleisch, A., & Kaiser, J. (2020). The false positive problem of automatic bot detection in social science research. *PloS one*, *15*(10), e0241045.

Rogers, N., & Jones, J. J. (2021). Using twitter bios to measure changes in self-identity: Are Americans defining themselves more politically over time?. *Journal of Social Computing*, *2*(1), 1-13.

Ross Arguedas, A., Robertson, C., Fletcher, R., & Nielsen, R. (2022). Echo chambers, filter bubbles, and polarisation: A literature review.

Rubin, D. B. (1973). Matching to remove bias in observational studies. *Biometrics*, 159-183.

Samory, M., & Mitra, T. (2018, June). Conspiracies online: User discussions in a conspiracy community following dramatic events. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 12, No. 1).

Scarrow, S. (2014). *Beyond party members: Changing approaches to partisan mobilization*. OUP Oxford.

Schwartz, H. A., & Ungar, L. H. (2015). Data-driven content analysis of social media: A systematic overview of automated methods. *The ANNALS of the American Academy of Political and Social Science*, *659*(1), 78-94.

Tajfel, H., & Turner, J. C. (1986). *The social identity theory of intergroup behavior*. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Nelson-Hall.

Temporão, M. (2026). "Infoxicated" feelings? How affective polarization and misinformation reinforce each other in the US and France. *French Politics*, *24*(1), 151-173.

Tufekci, Z. (2014, May). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. In *Proceedings of the international AAAI conference on web and social media* (Vol. 8, No. 1, pp. 505-514).

Valenzuela, S., Halpern, D., Katz, J. E., & Miranda, J. P. (2019). The paradox of participation versus misinformation: Social media, political engagement, and the spread of misinformation. *Digital Journalism*, *7*(6), 802-823.

Varol, O., Ferrara, E., Davis, C., Menczer, F., & Flammini, A. (2017, May). Online human-bot interactions: Detection, estimation, and characterization. In *Proceedings of the international AAAI conference on web and social media* (Vol. 11, No. 1, pp. 280-289).

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *science*, *359*(6380), 1146-1151.

Wagner, M. C., & Boczkowski, P. J. (2019). The reception of fake news: The interpretations and practices that shape the consumption of perceived misinformation. *Digital journalism*, *7*(7), 870-885.

Weeks, B. E., Ksiazek, T. B., & Holbert, R. L. (2016). Partisan enclaves or shared media experiences? A network approach to understanding citizens' political news environments. *Journal of Broadcasting & Electronic Media*, *60*(2), 248-268.

Widjaya, R. (2022). Jobs, hobbies top the list of things US adults put in their Twitter profiles; references to politics relatively rare. *Pew Research Center*.

Wojcieszak, M., Casas, A., Yu, X., Nagler, J., & Tucker, J. A. (2022). Most users do not follow political elites on Twitter; those who do show overwhelming preferences for ideological congruity. *Science advances*, *8*(39), eabn9418.

Wojcik, S., Messing, S., Smith, A., Rainie, L., & Hitlin, P. (2018). Bots in the Twittersphere: An estimated two-thirds of tweeted links to popular websites are posted by automated accounts – not human beings. *Pew Research Center*.

Wojcik, S., & Hughes, A. (2019, April 24). Sizing up Twitter users. *Pew Research Center*.

Zhang, Y., Shah, D., Foley, J., Abhishek, A., Lukito, J., Suk, J., ... & Garlough, C. (2019). Whose lives matter? Mass shootings and social media discourses of sympathy and policy, 2012–2014. *Journal of Computer-Mediated Communication*, *24*(4), 182-202.

Supplementary Materials

Table of contents

1. Overview of Academic Literature on Misinformation Sharers on Twitter
2. Analyses of Misinformation Content
3. Verifying the Presence of Bots among Misinformation Sharers on Twitter
4. Robustness Checks
5. Political Bio Classifier
6. Pseudonym Classifier
7. Topic-modeling Analysis of Circulated Tweets

1. Overview of Academic Literature on Misinformation Sharers on Twitter

Supplementary Table 1. Overview of Academic Literature on Misinformation Sharers on Twitter.

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
Gaumont et al., 2018	2.4 M	60 M	France	August 1, 2016 to May 8, 2017	179 *fake news* stories	Analysis of digital traces Political community detection	(1) Fake news accounted for just 0.081% of all tweet shares. (2) Right-wing and extreme-right wing communities were the most active with respectively 50.75% and 22.21% of the total number of fake links shared.

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
Grinberg et al., 2019	16,442	N/A	US	August 1 to December 6, 2016	171 black, 64 red, and 65 orange fake news sources	Link U.S. voter registration records to their Twitter accounts	(1) 1% of users were exposed to 80% of fake news. (2) 0.1% of users were responsible for sharing 80% of fake news. (3) While fewer than 5% of people on the left or in the center ever shared any fake news content, 11% and 21% of people on the right and extreme right did. (4) Age and low ratio of followers to followees were also positively associated with sharing fake news sources.

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
Hopp et al., 2020	543	N/A	US	March 7 to June 6, 2017	106 countermedia websites	Digital traces linked to survey data	(1) 95.0% of respondents did not share any countermedia content. (2) Countermedia content sharing is negatively associated with social trust. (3) No significant relationship was observed between ideological extremity, trust in mainstream news media and the likelihood of sharing countermedia content (4) None of the key predictors (ideological extremity, social trust, and trust in mainstream news media) were

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
							significantly associated with the sharing of mainstream news
Osmundsen et al., 2021	2,337	2,709,052	US	December 2018 to January 2019	608 fake news sources	Digital traces linked to survey data	(1) 4% of the tweeted news links come from websites with a history of publishing fake news. (2) 11% of participants shared stories from fake news sources (3) 1% of the panelists are responsible for sharing about 75% of all links to fake news publishers. (4) Measures of ignorance – cognitive

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
							reflection, political knowledge, and political digital media literacy – were not associated with fake news sharing. (5) Older panelists are much more likely to share content from fake news sources. (6) Partisan motivations, notably animus towards political opponents, are key to understanding fake news sharing.
Fong et al., 2021	10 influencers + 1,656 follower	16,290 influencer tweets 160,949 follower	UK/US	July 21 to 25, 2017	8,112 tweets from the timelines of 5 influencers conspiracy	Psycholinguistic analysis (LIWC)	(1) Conspiracy influencers scored significantly higher in the use of negative emotion (e.g.anger,

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
	accounts	tweets			theorists 85,071 tweets from users exposed to conspiratorial messages		anxiety) compared to science influencers. (2) A focus on topics such as death, religion, and power was found among conspiracy theorists. (3) No significant differences for outgroup language, nor for cognitive processes.
Mosleh et al., 2021	1,901	1,871,963	US	August 18, 2018 and April 12, 2020	tweet links to 60 news websites whose trustworthiness was rated by fact-checkers	Psycholinguistic analysis (LIWC) Digital traces linked to survey data Network analysis	(1) Users who scored higher on the CRT are significantly more likely to tweet links to news websites, suggesting that users who scored higher on the CRT are more likely to tweet about weightier subjects

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
							namely the news. (2) People who score higher on the Cognitive Reflection Test were more discerning in their social media use, as evidenced by the types and number of accounts followed, and by the reliability of the news sources they shared.
Kennedy et al., 2022	2 .8 M	32.4 M	US	September 1 to December 15, 2020	456 distinct misinformation stories	Network analysis	(1) Asymmetry of misinformation spread: accounts associated with support for President Biden shared far less fake news than accounts supporting his opponent (e.g. among the top 100 "repeat spreaders," all but two are associated with the

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
							Trump camp ; and approximately 79% of all misinformation posts come from accounts identified as right-wing). (2) Stories marked as sowing doubt in the election spread more within pro-trump communities than pro-biden communities.
de Oliveira & Albuquerque, 2021	N/A	259,176	Brazil	April to June 2020	3,100 entries classified as fake new	Machine learning	(1) Fake news was not less shared and less adopted than non-fake news. (2) Social prestige (i.e. influence or celebrity status) was not found to be a determining factor in the dissemination of misinformation. This suggests that even

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
							ordinary users contribute significantly to the spread of fake news.
DeVerna et al., 2022	Study 1 *N* = 487,156 Study 2 N/A	Study 1 N/A Study 2 33 M	US	Study 1 April 15 to April 25, 2013 Study 2 July 9 to December 31, 2019	Study 1 96,903 tweets related to two rumors about the Boston Marathon bombings Study 2 411,423 tweets related to rumors about Epstein's death	Supervised machine-learning techniques Ideological inference	(1) Rumors related to the Boston Marathon bombings were spread more than twice as often by conservatives as liberals before the correction period, and they were spread nearly 10 times as often by conservatives after the attempt at correction was made. (2) Rumors related to the death of Jeffrey Epstein were spread twice as often by conservatives as

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
							liberals before the correction period, and they were spread nearly eight times as often by conservatives after the attempt at correction was made. (3) With respect to ideologically congenial rumors, conservatives circulated the rumor that the Clinton family was involved in Epstein's death 18.6 times more often than liberals circulated the rumor that the Trump family was involved. (4) More than 96% of all fake news domains were shared by conservative Twitter users.

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
Rathje et al, 2022	Study 1 $N = 464$ Study 2 N=1,600	Study 1 N/A Study 2 1,831,308	US & UK	April to October 2021	1,030 hyper-partisan websites	Digital traces linked to survey data Network analysis	(1) In the US, following Republican politicians and hyper-partisan news sources is significantly associated with lower vaccine confidence — even when controlling for key demographics such as self-reported political ideology and education. In the UK, however, these associations are not significant (2) Among US influencers, right-wing figures are linked to the lowest levels of vaccine confidence among their followers. (3) US participants segregate into two

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
							distinct communities: (1) Community A (more liberal) shows lower conservatism and higher vaccine confidence ; (2) Community B (more conservative) exhibits higher political conservatism and lower vaccine confidence. In the UK, while a similar community structure exists, the relationship between political ideology and vaccine confidence is less straightforward. (4) In the US, higher connectedness within the conservative community is associated with lower

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
							vaccine confidence. No significant centrality effects were found in the UK. (5) Individuals with lower self-reported vaccine likelihood are more likely to share and interact with low-quality, hyper-partisan news on Twitter. (6) Self-reported political conservatism is negatively associated with vaccine confidence, particularly in the US.

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
Mosleh and Rand, 2022	5,000	N/A	US	last 3,200 tweets retrieved on July 23, 2021	20 fake news and 20 hyper-partisan websites	Analysis of digital traces Ideological inference Network analysis	(1) Users' misinformation-exposure scores are negatively correlated with the quality of news they share, and positively correlated with estimated conservative ideology. (2) Misinformation-exposure scores were positively correlated with language toxicity and moral-outrage expressions. (3) Users in the politically moderate cluster used less toxic and outraged language compared to the liberal and conservative clusters.

Study	Sample of users	Corpus of tweets	Country	Time period	Corpus of Misinformation	Methods	Main findings
Baribi-Bartov et al., 2024	664,391 registered U.S. voters active on Twitter; 2,107 identified as supersharers	N/A	US	August to November 2020	Fake-news links based on Grinberg et al.'s fake-news source list and updated NewsGuard ratings	Digital traces linked to voter registration records Comparison with reference groups Network/reach analysis	(1) 2,107 supersharers accounted for 80% of fake-news content shared on Twitter during the 2020 U.S. presidential election. (2) Supersharers reached 5.2% of registered voters on the platform. (3) They were overrepresented among women, older adults, and registered Republicans. (4) Their activity did not appear automated but was generated through manual and persistent retweeting.

Note. The literature review table above provides a comprehensive overview of studies analyzing the dissemination of various types of misinformation on Twitter and examining the characteristics of users who share such content. For each study, the table details the total number of tweets and users analyzed, the specific misinformation corpus utilized, the methodologies employed, and the main findings. This structured presentation allows for a clear comparison of research approaches and outcomes, facilitating a deeper understanding of how misinformation spreads on social media platforms and the user traits associated with its propagation.

2. Analyses of Misinformation Content

2.1. Themes

Supplementary Table 2.1. Annotation grid used to classify the themes of misinformation items shared on Twitter.

Theme	Description	Example
Politics	Content that mentions political figures or parties, or refers to events related to political news (e.g., law adoption, protests, scandal).	"Macron plans to cut hospital payrolls by €1.2 billion, which means the dismissal of 30,000 nurses."
Society	Content that focuses on topics of general interest by addressing social issues or societal challenges without explicitly referencing political figures or parties.	"France: It's confirmed that Arabic will be taught from first grade."
Health	Content that refers to diseases, medications, or vaccines.	"Vaping would increase the risk of cancer and heart disease."
Ecology	Content concerning environmental issues, climate change, or nature conservation.	"It is arrogant to believe that in 150 years of industrialization we have changed the climate."
Miscellaneous	Content that reports an unusual or anecdotal incident (e.g., accident, disappearance, strange phenomenon) and is not related to a public interest topic.	"A cosmic event that only happens every 35,000 years: On July 27, 2018, Mars will be as large as the Moon."

2.2. Topics and Rhetoric

Supplementary Table 2.2. Annotation grid used to classify the rhetorical features of misinformation items shared on Twitter.

Rhetoric	Description	Example
Anti-elite	Content that opposes the discourses and actions of individuals or institutions holding significant political, economic, social, or cultural power in society.	"€120,000 misappropriated by Macron: The internet exposes his dirty laundry because the media won't."
Identity-based	Content that defends the values, culture, or religion of the French or European population while rejecting those of other groups.	"Yasmine, who wears a hijab, cannot receive the Young European Award; she cannot be European."
Vaccine-skeptical	Content that expresses one or more criticisms regarding various issues related to vaccination.	"Hepatitis B vaccine: The European Court recognizes the link with multiple sclerosis."
Chemophobia/Technophobia	Content that denounces the use of chemical products or new technologies.	"Dr. Henri Joyeux, oncologist, stated: '85% of chemotherapies are questionable, even useless.'"

Rhetoric	Description	Example
Climate-skeptical	Content that questions the existence of global warming or denies human responsibility for the phenomenon.	“A Finnish scientific study concludes that the IPCC is mistaken: there is no evidence that climate change is man-made.”
Sensationalist	Content that does not aim to influence public opinion but rather to surprise or shock the audience.	“This dead woman rises from her grave and walks 3 years after her death.”
Security-focused	Content that emphasizes maintaining order and denounces acts of violence or delinquency.	“Notre-Dame: The criminal investigation should be prioritized; what happened was impossible.”
Other	Content that does not fit into any of the above categories.	“Le journal de 18h: Increase in road accidents at 80 km/h.”

3. Verifying the presence of bots among misinformation sharers on Twitter

Because misinformation sharers displayed high levels of activity, we assessed whether this pattern could be driven by automated accounts using Botometer, a well-established bot detection tool in academic literature. This tool employs various metrics to determine the likelihood that a Twitter account is managed by a bot rather than a human (Davis et al., 2016; Varol et al., 2017; Bessi & Ferrara, 2016; Wojcik et al., 2018; Keller & Klinger, 2019). These metrics include the account's posting activity, tweet content, and network characteristics such as the follower-to-friend ratio.

Botometer assigns a score between 0 and 1, with higher scores indicating a greater probability that the account is a bot. Researchers typically define a manual threshold ranging from 0.25 (Zhang et al., 2019) to 0.76 (Keller & Klinger, 2019) to classify accounts in a binary manner. While there is no consensus in the scientific literature on the optimal threshold, most studies select a value between 0.4 and 0.5 (Wojcik et al., 2018).

To validate Botometer's performance, we manually reviewed all user profiles with a score above 0.4 (n=272). Among these users, only 11 were suspected of being bots. We did not exclude these accounts from the main sample, both because they represented a very small number of cases and because Botometer can generate false positives (Rauchfleisch & Kaiser, 2020; He et al., 2021).

4. Robustness checks

To assess whether our findings depend on specific features of the main ideology-matched sample, we conducted three robustness checks: one restricted to one-time misinformation sharers, one matching misinformation sharers and control users on both ideological position and daily Twitter activity, and one based on misinformation and news sharers who could not be located in the ideological space.

4.1. Construction of the samples

4.1.1. One-time misinformation sharers

To ensure that our results are not driven solely by users who repeatedly shared misinformation, we constructed a robustness sample restricted to one-time misinformation sharers. We identified these accounts at the profile level by selecting misinformation sharers with only one misinformation share, and then retained all tweets available in their collected timelines for analysis. This procedure yielded 302 one-time misinformation sharers, corresponding to 412,934 tweets.

4.1.2. Ideology and Activity Matched Samples

To address the possibility that the observed differences merely reflect baseline differences in Twitter activity, we constructed an activity-matched robustness sample restricted to users for whom a comparable counterpart could be identified. Misinformation sharers were matched one-to-one, without replacement, to control users located in the same cell of the ideological space and displaying comparable observed daily tweeting activity. Daily activity was measured over each user's collected timeline as the number of observed tweets, including retweets, divided by the number of days between the first and last collected tweet. Because posting activity is highly skewed, matching was performed on log-transformed daily activity.

A sensitivity analysis comparing alternative matching thresholds led us to retain a maximum absolute distance of 0.20 on log-transformed daily activity, which offered the best compromise between sample retention and matching quality (Supplementary Table 3). This threshold retained 454 matched pairs, corresponding to 908 users and 813,440 tweets, while covering all 16 ideological cells. Balance diagnostics indicate that the procedure substantially reduced pre-existing differences in activity: before matching, misinformation sharers had a median daily activity of 21.97 tweets per day, compared with 0.91 among control users; after matching, the two groups had nearly identical median activity levels, 3.17 and 3.12 tweets per day respectively. The standardized mean difference for daily activity decreased from 0.949 before matching to 0.003 after matching (Supplementary Table 4). Pairwise distances were also very small, with a median log-activity distance of 0.008.

This robustness sample is not intended to represent all misinformation sharers, who are substantially more active in the full dataset. Instead, it restricts the comparison to the subset of users who can be meaningfully compared on both ideological position and observed daily activity. It therefore allows us to assess whether group differences persist when daily tweeting intensity is held approximately constant.

Supplementary Table 3. Sensitivity of the daily-activity matching to alternative calipers

Caliper	Matched pairs	Misinformation sharers retained (%)	Control retained (%)	Mean distance	Max distance	SMD tweets/day	SMD log tweets/day
0.05	395	20.7	41.2	0.011	0.049	-0.000	0.001
0.10	431	22.6	45.0	0.017	0.095	-0.002	0.001
0.15	447	23.4	46.7	0.020	0.149	0.002	0.003
0.20	**454**	**23.8**	**47.4**	**0.023**	**0.189**	**0.003**	**0.006**
0.25	457	24.0	47.7	0.024	0.237	0.003	0.006
0.30	460	24.1	48.0	0.026	0.296	0.004	0.008
0.35	462	24.2	48.2	0.027	0.350	0.004	0.010
0.40	462	24.2	48.2	0.027	0.350	0.004	0.010
0.50	468	24.5	48.9	0.033	0.500	0.005	0.015

Supplementary Table 4. Balance before and after activity matching

Variable	Median misinformation before	Median control before	SMD before	Median misinformation after	Median control after	SMD after
Tweets per day	21.97	0.91	0.949	3.17	3.12	0.003
Log daily activity	3.13	0.65	1.859	1.43	1.42	0.006
Observed tweets	2,216.00	139.50	1.660	559.50	549.50	-0.013

4.1.3. Ideologically unclassified samples

To assess whether our results depend on restricting the analysis to users who could be located in the ideological space, we constructed an additional robustness sample of ideologically unclassified users. This sample was drawn from the broader French media-sharing population, defined as users who had shared at least one URL from one of the 420 media outlets included in the study. We identified misinformation sharers as users who had shared at least one fact-checked misinformation URL but had not received an ideological label, and compared them with news sharers drawn from the same population after excluding identified misinformation sharers and ideologically classified users. The full sample included 3,764 ideologically unclassified misinformation sharers and 4,889 ideologically unclassified news sharers. Profile-level analyses were conducted on this full sample, whereas computationally intensive tweet-level analyses were conducted on a random half-sample of 1,882 misinformation sharers and 2,444 news sharers. Finally, analyses of daily activity and high-intensity posting were restricted to the common observation window from March 31 to October 1, 2020, yielding 747 misinformation sharers who posted 328,688 tweets and 777 news sharers who posted 301,250 tweets.

4.2. Results

4.2.1. Online identity

The association between misinformation sharing and political self-presentation is robust across checks. In the main ideology-matched sample, misinformation sharers were more than twice as likely as control users to display a political bio (42.1% vs. 19.4%; $\chi^2(1, N = 2{,}865) = 145.27$, p_adj < .001, $\varphi = 0.225$). The pattern remains among one-time misinformation sharers (39.4% vs. 19.4%; $\chi^2(1, N = 1{,}260) = 50.00$, p_adj < .001, $\varphi = 0.199$), and in the ideology-and-activity matched sample, where misinformation sharers are still more likely than their matched controls to foreground politics in their bios (43.2% vs. 20.7%; McNemar exact test, p_adj < .001; paired OR = 3.24).[10] For the ideologically unclassified sample, political bios were identified using a supervised classifier trained on manually annotated bios (see Supplementary Section 5). In this broader sample, the overall propensity to foreground political identity is lower, but misinformation sharers remain more likely than ideologically unclassified news sharers to have a political bio (20.1% vs. 15.1%; $\chi^2(1, N = 8{,}653) = 37.34$, p_adj < .001, $\varphi = 0.066$). Overall, these results indicate that, although political self-presentation is partly shaped by the politically engaged nature of the populations under study, it remains a distinctive and robust correlate of misinformation sharing.

4.2.2. Informational practices

The robustness checks confirm the informational-practice pattern observed in the main ideology-matched sample, while clarifying its scope. Among one-time misinformation sharers, media URLs account for 19% of tweets, compared with 14% among control users, a smaller but still significant difference than in the main sample,

[10] We use McNemar's exact test for the ideology-and-activity matched sample because users are compared within one-to-one matched pairs rather than as two independent groups.

where the corresponding proportions are 23% and 14% (Table 1; Supplementary Table 5). This suggests that the higher orientation toward media URLs is not driven only by recurrent misinformation sharers. However, URLs not matched to the media database do not differ significantly after correction, and the composition of media sources is closer to that of control users than in the other misinformation-sharer samples. This is particularly visible for partisan and counter-informational media, whose descriptive shares are almost identical after rounding. By contrast, one-time misinformation sharers stand out more clearly in their use of tabloid/local press: this category accounts for 17% of their media URLs, compared with 14% among control users, 14% among misinformation sharers in the main sample, and 13% in the ideology-and-activity matched sample. Their media repertoire therefore appears less strongly differentiated by partisan or counter-informational sources than that of recurrent misinformation sharers, and relatively more marked by tabloid/local press.

The ideology-and-activity matched sample provides a stronger test of the same pattern (Supplementary Table 6). Even when misinformation sharers are matched to control users on both ideological position and daily activity, they share more media URLs than their matched controls (22% vs. 14%) and fewer tweets without URLs (52% vs. 58%). Their media repertoire also remains less centered on mainstream media (53% vs. 62%) and more oriented toward partisan media (20% vs. 16%) and counter-informational media (14% vs. 8%).

The ideologically unclassified sample shows both continuity and a different baseline. Overall media sharing is much lower among ideologically unclassified users than among ideologically classifiable users: media URLs account for only 10% of tweets among ideologically unclassified misinformation sharers and 3% among ideologically unclassified news sharers, compared with 23% and 14% in the main ideology-matched sample (Table 1; Supplementary Table 7). Yet the gap between misinformation sharers and ideologically unclassified news sharers remains substantial. Ideologically unclassified misinformation sharers also share more URLs not matched to the media database (35% vs. 22%) and far fewer tweets without URLs (55% vs. 76%). Among media URLs, the pattern closely

resembles the main sample: ideologically unclassified misinformation sharers share fewer mainstream sources (52% vs. 66%) and more partisan (16% vs. 10%) and counter-informational sources (17% vs. 8%), while tabloid/local press does not differ significantly.

Overall, the robustness checks indicate that misinformation sharing is consistently associated with a more media-oriented and less mainstream-centered informational repertoire, although this pattern is strongest among ideologically classifiable users

Supplementary Table 5. Distribution of content types and media sources among one-time misinformation sharers

Measures	Variables	One-time misinformation sharers	Control users	U	p_adj	Effect size
Type of content	media URL	19%	14%	172,823.0	< .001	0.195
	other URL	26%	28%	152,100.0	.177	0.051
	no URL	56%	58%	128,655.5	.004	-0.111
Type of media sources	mainstream media	60%	62%	118,314.5	< .001	-0.182
	partisan media	16%	16%	185,579.5	< .001	0.283
	counter-informational media	8%	8%	187,667.0	< .001	0.297

Measures	Variables	One-time misinformation sharers	Control users	U	p_adj	Effect size
	tabloid/local press	17%	14%	173,907.0	< .001	0.202

Note. Statistical tests are conducted at the user level using Mann–Whitney U tests. P-values are adjusted for multiple comparisons using the Benjamini-Hochberg procedure. Effect size is reported as rank-biserial correlation.

Supplementary Table 6. Distribution of content types and media sources in the ideology-and-activity matched sample

Measures	Variables	Misinformation sharers	Control users	W	p_adj	Effect size
Type of content	media URL	22%	14%	25,781.0	< .001	0.499
	other URL	27%	29%	50,140.0	.755	-0.025
	no URL	52%	58%	38,970.5	< .001	-0.242
Type of	mainstream media	53%	62%	30,267.0	< .001	-0.393
	partisan media	20%	16%	26,741.5	< .001	0.409

Measures	Variables	Misinformation sharers	Control users	W	p_adj	Effect size
media sources	counter-informational media	14%	8%	15,625.5	< .001	0.542
	tabloid/local press	13%	14%	46,732.0	.859	-0.010

Note. Statistical tests are conducted at the user level using Wilcoxon signed-rank tests because users are compared within one-to-one matched pairs. P-values are adjusted for multiple comparisons using the Benjamini-Hochberg procedure. Effect size is reported as paired rank-biserial correlation.

Supplementary Table 7. Distribution of content types and media sources in the ideologically unclassified sample

Measures	Variables	Ideologically unclassified misinformation sharers	Ideologically unclassified news sharers	U	p_adj	Effect size
Type of content	media URL	10%	3%	3,425,631.0	< .001	0.490
	other URL	35%	22%	3,385,814.0	< .001	0.472
	no URL	55%	76%	932,247.5	< .001	-0.595
	mainstream media	52%	66%	927,292.5	< .001	-0.399

Measures	Variables	Ideologically unclassified misinformation sharers	Ideologically unclassified news sharers	U	p_adj	Effect size
Type of media sources	partisan media	16%	10%	2,213,367.0	< .001	0.434
	counter-informational media	17%	8%	2,333,548.5	< .001	0.512
	tabloid/local press	15%	17%	1,545,143.5	.964	0.001

Note. Statistical tests are conducted at the user level using Mann–Whitney U tests. P-values are adjusted for multiple comparisons using the Benjamini-Hochberg procedure. Effect size is reported as rank-biserial correlation.

4.2.3. Behavior on Twitter

The robustness checks confirm that differences in baseline Twitter activity are not limited to recurrent misinformation sharers. One-time misinformation sharers are also more active than control users, although less so than misinformation sharers in the main sample. They post a median of 5.74 tweets per day, compared with 0.91 among control users, and the difference is significant after correction for multiple comparisons. By contrast, the ideology-and-activity matched sample is well balanced by construction: misinformation sharers and matched control users post almost identical median numbers of tweets per day, 3.17 and 3.12 respectively. Finally, in the ideologically unclassified sample, daily activity is also virtually identical across groups: misinformation sharers post a median of 1.30 tweets per day, compared with 1.41 among ideologically unclassified news sharers. Together, these robustness checks show that the strong activity gap observed

in the main ideology-matched sample is not uniformly reproduced across all alternative comparison samples.

Supplementary Table 8. Daily Twitter activity across robustness samples

Sample	Median tweets/day, misinformation group	Median tweets/day, control users	U/W	p_adj	Effect size
Main ideology-matched sample	21.97	0.91	1,637,278.0	< .001	0.792
One-time misinformation sharers	5.74	0.91	221,842.0	< .001	0.534
Ideology-and-activity matched sample	3.17	3.12	47,755.0	.252	0.071
Ideologically unclassified sample	1.30	1.41	290,082.0	.988	0.000

Note. The statistic reported is Mann–Whitney U for independent samples and Wilcoxon W for the ideology-and-activity matched sample. Effect size is reported as rank-biserial correlation for Mann–Whitney U tests and paired rank-biserial correlation for the Wilcoxon signed-rank test. P-values are adjusted using the Benjamini-Hochberg procedure.

The robustness checks confirm that high-intensity posting is not limited to recurrent misinformation sharers. Among one-time misinformation sharers, 74.5% posted at least 10 tweets within one hour, compared with 34.4% of control users. The effect remains significant across all alternative cutoffs from 5 to 30 tweets within one hour (Supplementary table 9). In the ideology-and-activity matched sample, misinformation sharers are still more likely than matched controls to meet the 10-tweet bombardment criterion (70.3% vs. 58.4%). This indicates that the bombardment pattern is not reducible to overall activity differences (Supplementary table 10). We also replicated the analysis in

the ideologically unclassified sample. Although daily activity was virtually identical across groups in this sample, misinformation sharers were slightly more likely than ideologically unclassified news sharers to meet the 10-tweet bombardment criterion, 42.3% compared with 37.1%. The difference is not significant at the lowest cutoff of 5 tweets within one hour, but it becomes significant from the 10-tweet threshold onward and increases at stricter cutoffs (Supplementary table 11). These results suggest that high-intensity posting is not only a by-product of overall daily activity: even when average activity is comparable, misinformation sharers are somewhat more likely to concentrate their posting activity in short bursts.

Supplementary Table 9. Bombardment among one-time misinformation sharers

Cutoff: k tweets within 1h	One-time misinformation sharers	Control users	Difference pp	χ^2	p_adj	φ	OR
5	90.4%	62.2%	28.2	85.22	< .001	0.260	5.63
10	74.5%	34.4%	40.1	149.49	< .001	0.344	5.53
15	57.0%	21.8%	35.1	134.39	< .001	0.327	4.73
20	43.4%	14.7%	28.7	111.42	< .001	0.297	4.43
25	34.8%	9.4%	25.4	113.01	< .001	0.300	5.13
30	28.1%	6.9%	21.3	98.36	< .001	0.279	5.28

Note. Statistical tests are chi-square tests. P-values are adjusted using the Benjamini-Hochberg procedure. Effect size is reported as φ.

Supplementary Table 10. Bombardment in the ideology-and-activity matched sample

Cutoff: k tweets within 1h	Misinformation sharers	Control users	Difference pp	Discordant pairs	p_adj	Paired OR
5	89.2%	84.6%	4.6	49	.004	2.50
10	70.3%	58.4%	11.9	118	< .001	2.69
15	49.8%	40.7%	9.0	133	.001	1.89
20	35.0%	27.8%	7.3	115	.003	1.81
25	25.3%	18.3%	7.0	96	.002	2.00
30	20.3%	13.9%	6.4	79	.002	2.16

Note. Statistical tests are exact McNemar tests because users are compared within one-to-one matched pairs. P-values are adjusted using the Benjamini-Hochberg procedure. Effect size is reported as paired odds ratio.

Supplementary Table 11. Bombardment in the ideologically unclassified sample

Cutoff: k tweets within 1h	Ideologically unclassified misinformation sharers	Ideologically unclassified news sharers	Difference pp	χ^2	p_adj	φ	OR
5	70.0%	69.9%	0.1	0.00	.956	0.001	1.01
10	42.3%	37.1%	5.2	4.37	.044	0.054	1.24

Cutoff: k tweets within 1h	Ideologically unclassified misinformation sharers	Ideologically unclassified news sharers	Difference pp	χ^2	p_adj	φ	OR
15	25.2%	19.4%	5.7	7.24	.011	0.069	1.39
20	15.3%	9.9%	5.4	9.95	.004	0.081	1.64
25	9.9%	5.7%	4.2	9.60	.004	0.079	1.83
30	6.6%	2.8%	3.7	11.92	.003	0.088	2.41

Note. Statistical tests are chi-square tests. P-values are adjusted using the Benjamini-Hochberg procedure. Effect size is reported as φ. Percentages indicate the share of users who posted at least k tweets within any one-hour window during the observation period.

The robustness checks confirm that the retweet-oriented mode of participation is strongest among recurrent misinformation sharers, but is not reducible to overall activity. Among one-time misinformation sharers, the descriptive pattern is attenuated: retweets account for 59% of their posts, compared with 52% among control users, while original tweets account for 35% and 41% (Supplementary Table 12), respectively. However, these differences are no longer significant after correction for multiple comparisons. By contrast, in the ideology-and-activity matched sample, misinformation sharers remain substantially more retweet-oriented than matched control users: retweets account for 68% of their posts, compared with 52% among matched controls, while original tweets account for 26% and 41%, respectively. These differences remain significant in paired Wilcoxon tests (Supplementary Table 13). The same broad pattern also appears in the ideologically unclassified sample, where ideologically unclassified misinformation sharers post fewer original tweets, more retweets, fewer replies, and more hashtags than ideologically unclassified news sharers (Supplementary Table 14).

Supplementary Table 12. Twitter usage among one-time misinformation sharers

Measures	Variables	One-time misinformation sharers	Control users	U	p_adj	Effect size
Tweet types	original tweet	35%	41%	131,832.0	.070	−0.089
	retweet	59%	52%	158,729.5	.070	0.097
	quote	5%	7%	150,729.0	.366	0.042
Tweet features	hashtag	31%	31%	143,442.0	.959	−0.008
	multimedia	36%	37%	134,853.0	.176	−0.068
Interactivity	reply	27%	28%	144,945.0	.959	0.002
	mention	27%	30%	106,939.0	.366	−0.047

Note. Statistical tests are Mann–Whitney U tests conducted at the user level. Effect sizes are rank-biserial correlations. P-values are adjusted using the Benjamini-Hochberg procedure. Mentions are measured among pure original tweets only.

Supplementary Table 13. Twitter usage in the ideology-and-activity matched sample

Measures	Variables	Misinformation sharers	Control users	W	p_adj	Effect size
Tweet types	original tweet	26%	41%	31,393.5	< .001	−0.384
	retweet	68%	52%	30,321.5	< .001	0.405
	quote	6%	8%	33,125.5	.004	−0.174
Tweet features	hashtag	36%	31%	41,976.5	.001	0.187
	multimedia	39%	37%	48,285.0	.268	0.065
Interactivity	reply	17%	28%	31,126.0	< .001	−0.376
	mention	32%	29%	31,389.5	.640	0.028

Note. Statistical tests are Wilcoxon signed-rank tests conducted within matched pairs. Effect sizes are paired rank-biserial correlations. P-values are adjusted using the Benjamini-Hochberg procedure. Mentions are measured among pure original tweets only.

Supplementary Table 14. Twitter usage in the ideologically unclassified sample

Measures	Variables	Ideologically unclassified misinformation sharers	Ideologically unclassified news sharers	U	p_adj	Effect size
Tweet types	original tweet	39%	68%	1,213,629.0	< .001	−0.472
	retweet	57%	25%	3,428,165.0	< .001	0.491
	quote	4%	7%	1,541,627.5	< .001	−0.330
Tweet features	hashtag	24%	14%	3,225,841.5	< .001	0.403
	multimedia	34%	27%	2,865,047.0	< .001	0.246
Interactivity	reply	19%	47%	997,739.5	< .001	−0.566
	mention	15%	14%	2,562,904.5	< .001	0.169

Note. Statistical tests are Mann–Whitney U tests conducted at the user level. Effect sizes are rank-biserial correlations. P-values are adjusted using the Benjamini-Hochberg procedure. Mentions are measured among pure original tweets only.

4.2.4. Online visibility and engagement

The robustness checks indicate that the online visibility and engagement pattern is not driven only by recurrent misinformation sharers. One-time misinformation sharers also have substantially higher median follower counts than control users (546 vs. 94.5), follow more accounts (980 vs. 358.5), and have a higher follower/following ratio (0.618 vs. 0.253). These differences remain significant after correction, although effect sizes are smaller than in the main sample. In the ideology-and-activity matched sample, the same pattern persists: activity-matched misinformation sharers have a much higher median follower count than matched control users (752 vs. 103.5), follow more accounts (1,117 vs. 361), and have a higher median follower/following ratio (0.731 vs. 0.262).

Engagement indicators are more mixed. One-time misinformation sharers do not differ significantly from control users in average likes per tweet or in the proportion of tweets receiving at least one like after correction. In the activity-matched sample, however, misinformation sharers receive fewer likes than matched control users, both in average likes per tweet and in the proportion of tweets receiving at least one like. Retweet-based engagement is weaker as a robustness result: while one-time misinformation sharers show a small difference in retweets received, differences in retweet-based engagement are no longer significant in the ideology-and-activity matched sample.

Ideologically unclassified misinformation sharers have only slightly higher median follower and following counts than ideologically unclassified news sharers, lower follower/following ratios, and substantially lower like-based engagement.

Supplementary Table 15. Online visibility and engagement among one-time misinformation sharers

Measures	Variables	One-time misinformation sharers median	Control users median	U	p_adj	Effect size
Online success	followers count	546.0	94.5	223,845.0	< .001	0.547
	following count	980.0	358.5	211,237.0	< .001	0.460
	follower/following ratio	0.618	0.253	204,476.0	< .001	0.414
Per-tweet engagement	average likes per tweet	0.247	0.310	139,010.0	.382	−0.039
	proportion of tweets with at least one like	11.6%	14.3%	138,987.5	.382	−0.039
	average retweets received per original tweet	0.148	0.120	146,114.0	.018	0.098
	proportion of original tweets with at least one retweet	9.1%	7.8%	144,787.5	.033	0.088

Note. Statistics are Mann–Whitney U tests conducted at the user level. Effect sizes are rank-biserial correlations. P-values are adjusted using the Benjamini-Hochberg procedure.

Supplementary Table 16. Online visibility and engagement in the ideology-and-activity matched sample

Measures	Variables	Misinformation sharers median	Control users median	W	p_adj	Effect size
Online success	followers count	752.0	103.5	17,662.5	< .001	0.658
	following count	1,117.0	361.0	19,239.5	< .001	0.627
	follower/following ratio	0.731	0.262	23,958.0	< .001	0.536
Per-tweet engagement	average likes per tweet	0.156	0.417	30,018.0	< .001	−0.400
	proportion of tweets with at least one like	6.8%	17.3%	27,765.0	< .001	−0.445
	average retweets received per original tweet	0.178	0.169	38,395.0	.210	−0.084
	proportion of original tweets with at least one retweet	10.2%	9.4%	40,784.0	.802	−0.022

Note. Statistics are Wilcoxon signed-rank tests conducted within matched pairs. Effect sizes are paired rank-biserial correlations. P-values are adjusted using the Benjamini-Hochberg procedure.

Supplementary Table 17. Online visibility and engagement in the ideologically unclassified sample

Measures	Variables	Ideologically unclassified misinformation sharers median	Ideologically unclassified news sharers median	U	p_adj	Effect size
Online success	followers count	268.0	221.0	9,722,209.5	< .001	0.057
	following count	371.5	279.0	10,510,951.5	< .001	0.142
	follower/following ratio	0.719	0.807	8,783,239.5	< .001	−0.045
Per-tweet engagement	average likes per tweet	0.167	0.853	3,968,253.0	< .001	−0.569
	proportion of tweets with at least one like	7.1%	30.3%	3,468,142.5	< .001	−0.623
	average retweets received per non-retweet	0.117	0.103	9,181,934.0	.485	0.009
	proportion of original tweets with at least one retweet	6.8%	5.1%	9,862,494.5	< .001	0.084

Note. Statistics are Mann–Whitney U tests conducted at the user level. Effect sizes are rank-biserial correlations. P-values are adjusted using the Benjamini-Hochberg procedure.

4.2.5. Use of pseudonyms

In the main ideology-matched sample, 65.0% of misinformation sharers used a pseudonym, compared with 45.1% of control users ($\chi^2(1, N = 2{,}865) = 104.22$, p_adj < .001, $\varphi = 0.191$; OR = 2.26, 95% CI [1.93, 2.65]). The pattern remains among one-time misinformation sharers (65.9% vs. 45.1%; $\chi^2(1, N = 1{,}260) = 39.74$, p_adj < .001, $\varphi = 0.178$; OR = 2.35, 95% CI [1.80, 3.08]), and in the ideology-and-activity matched sample, where misinformation sharers remain more likely than their matched controls to use a pseudonym (65.0% vs. 49.6%; McNemar exact test, p_adj < .001; paired OR = 1.95). For the ideologically unclassified sample, pseudonym use was identified using a supervised classifier trained on manually annotated accounts (see Supplementary Section 6). In this broader sample, the association is weaker and reverses direction: ideologically unclassified misinformation sharers are less likely than ideologically unclassified news sharers to be classified as pseudonymous (57.8% vs. 62.4%; $\chi^2(1, N = 8{,}653) = 18.43$, p_adj < .001, $\varphi = 0.046$; OR = 0.83, 95% CI [0.76, 0.90]). Overall, these results suggest that pseudonymous self-presentation is strongly associated with misinformation sharing among ideologically classifiable users, but that this relationship does not extend to the broader ideologically unclassified population.

4.2.6. Psycholinguistic traits

The LIWC robustness checks confirm that affective differences are more consistent than cognitive-process differences. Among one-time misinformation sharers, overall cognitive-process scores, insight, causation, and certainty do not differ significantly from those of control users after correction for multiple comparisons, while use of words in the inhibition category is somewhat higher. By contrast, one-time misinformation sharers use fewer positive-emotion words and more negative-emotion, anxiety, anger, and sadness words (Supplementary Table 18).

The same affective pattern persists in the ideology-and-activity matched sample: misinformation sharers use fewer positive-emotion words and more negative-emotion, anxiety, anger, and sadness words than matched control users. Cognitive-process results remain mixed: certainty is lower and inhibition higher, whereas overall cognitive processes, insight, and causation do not differ significantly (Supplementary Table 19).

Overall, these robustness checks indicate that differences in affective language persist among one-time misinformation sharers and when daily activity is held approximately constant. Cognitive-process differences, by contrast, are less consistent and generally modest.

Supplementary Table 18. LIWC scores among one-time misinformation sharers

Measures	Variables	One-time misinformation sharers	Control users	U	p_adj	Effect size
Cognitive style	cognitive processes	9.44	9.82	127,524.5	.332	−0.042
	insight	1.40	1.34	141,006.5	.171	0.060
	causation	1.51	1.49	134,265.5	.875	0.009
	certainty	1.37	1.40	133,456.5	.936	0.003
	inhibition	0.66	0.565	151,791.5	< .001	0.141
	analytic thinking	30.70	30.00	142,159.0	.120	0.069

Measures	Variables	One-time misinformation sharers	Control users	U	p_adj	Effect size
	agency	36.54	38.60	121,324.0	.040	−0.088
Affect and emotion	emotional tone	0.30	0.63	109,691.0	< .001	−0.176
	affect	4.06	4.14	126,531.5	.263	−0.049
	positive emotion	2.14	2.36	117,800.5	.006	−0.115
	negative emotion	1.75	1.565	149,841.0	.003	0.126
	anxiety	0.19	0.15	152,417.0	< .001	0.146
	anger	0.58	0.46	154,951.0	< .001	0.165
	sadness	0.52	0.42	158,252.5	< .001	0.190

Supplementary Table 19. LIWC scores in the ideology-and-activity matched sample

Measures	Variables	Misinformation sharers	Matched control users	W	p_adj	Effect size
Cognitive style	cognitive processes	9.40	10.05	40,293.5	.148	−0.089
	insight	1.34	1.42	41,155.5	.371	−0.056
	causation	1.49	1.61	41,047.5	.371	−0.054
	certainty	1.34	1.46	37,040.0	.032	−0.129
	inhibition	0.69	0.605	33,477.0	.002	0.198
	analytic thinking	31.35	29.33	35,493.0	.002	0.197
	agency	36.34	39.41	36,528.0	.004	−0.174
Affect and emotion	emotional tone	0.115	0.615	29,344.5	< .001	−0.330
	affect	4.14	4.22	41,689.5	.396	−0.048
	positive emotion	2.14	2.41	35,173.0	.002	−0.193

Measures	Variables	Misinformation sharers	Matched control users	W	p_adj	Effect size
	negative emotion	1.89	1.66	34,747.5	.002	0.203
	anxiety	0.21	0.18	31,977.0	.018	0.148
	anger	0.61	0.54	34,735.5	.004	0.180
	sadness	0.52	0.465	35,427.0	.009	0.159

5. Political bio classifier

To identify political self-presentation in the ideologically unclassified samples, we trained a supervised text classifier on the manually annotated user bios from the main sample. The model used character-level TF-IDF features, based on 3- to 5-character n-grams, and a class-balanced logistic regression. Empty bios were treated as non-political, following the coding rule used in the manual annotation. On a stratified 80/20 train-test split, the classifier achieved an accuracy of 0.87 and a macro-F1 score of 0.86. Performance was strong for non-political bios, with precision of 0.94, recall of 0.86, and F1-score of 0.89. Performance was also satisfactory for political bios, with precision of 0.77, recall of 0.89, and F1-score of 0.82. Five-fold cross-validation confirmed the stability of the model, with mean accuracy of 0.84, macro-F1 of 0.83, and ROC-AUC of 0.94. We therefore used this classifier to predict whether user bios in the ideologically unclassified samples contained political self-presentation.

6. Pseudonym classifier

To identify pseudonymous self-presentation in the ideologically unclassified samples, we trained a supervised text classifier on the manually annotated accounts from the main sample. The model used character-level TF-IDF features, based on 2- to 5-character n-grams, and a class-balanced logistic regression. On a stratified 80/20 train-test split, the classifier achieved an accuracy of 0.71 and a macro-F1 score of 0.70. Performance was moderate for non-pseudonymous accounts, with precision of 0.65, recall of 0.64, and F1-score of 0.64. It was stronger for pseudonymous accounts, with precision of 0.74, recall of 0.75, and F1-score of 0.75. Five-fold cross-validation produced similar results, with mean accuracy of 0.71, macro-F1 of 0.70, and ROC-AUC of 0.77. We therefore used this classifier to predict whether accounts in the ideologically unclassified samples relied on pseudonymous rather than non-pseudonymous self-presentation.

7. Topic-modeling analysis of circulated tweets

7.1. Corpus construction and preprocessing

To complement the URL-level analysis, we conducted an exploratory topic-modeling analysis of the tweets posted or reshared by misinformation sharers and ideologically matched control users. The purpose was to examine whether the two groups circulated different thematic repertoires beyond the URLs they shared.

The source corpus contained 3,644,421 tweets from misinformation sharers and 459,422 tweets from control users. Because this imbalance could mechanically influence topic identification, we constructed a balanced analytical corpus by randomly sampling 10,000 tweets from each group, using random seed 42 for reproducibility. Before sampling, we retained tweets containing non-empty textual content of more than 20 characters.

We then removed retweet headers, URLs, user mentions, and excessive whitespace. Tweets containing fewer than 15 characters after cleaning were excluded. The final analytical corpus comprised 18,998 tweets.

7.2. Topic-model specification and interpretation

Topics were identified using BERTopic. Tweet embeddings were generated with the multilingual sentence-transformer model paraphrase-multilingual-MiniLM-L12-v2, which produces 384-dimensional semantic representations and supports multiple languages, including French.

Dimensionality reduction was performed using UMAP with the following parameters: n_neighbors = 15, n_components = 5, min_dist = 0.0, metric = cosine, and random_state = 42. The resulting embeddings were then clustered using HDBSCAN with min_cluster_size = 150, min_samples = 50, metric = euclidean, cluster_selection_method = eom, and prediction_data = True.

Topic representations were generated using BERTopic's class-based TF-IDF procedure. The count vectorizer excluded French and English stopwords, retained terms appearing in at least five documents, and included both unigrams and bigrams.

This specification produced 17 substantive topics. An additional Topic −1 category contained 6,231 tweets, corresponding to 32.8% of the analytical corpus. These tweets were treated as outliers and excluded from the substantive interpretation. The remaining 12,767 tweets, or 67.2% of the corpus, were assigned to one of the 17 topics.

Topic labels were assigned by the first author through manual inspection of the most highly weighted c-TF-IDF terms and a selection of representative tweets associated with each cluster (see Supplementary Figure 1). This procedure provided an interpretive check on the topic labels.

7.3. Comparing topic prevalence across groups

For each substantive topic, we counted the number of tweets assigned to that topic in each group and calculated a smoothed log prevalence ratio. Positive values indicate that

a topic was proportionally more prevalent among misinformation sharers, whereas negative values indicate greater prevalence among control users. The magnitude of the ratio indicates the strength of the difference between the two groups.

For each topic, statistical significance was assessed using a chi-square test comparing the number of tweets assigned to that topic with the number assigned to all other topics across the two groups. P-values were adjusted across the 17 topic-level comparisons using the Benjamini–Hochberg procedure.

Supplementary Figure 1. Highest-weighted terms for BERTopic topics