

Original Paper

From Algorithmic Filter Bubbles to Socially Mediated Echo Chambers: Explaining the Spread of Pro-Bolsonaro Disinformation on WhatsApp during Brazil's 2018 Presidential Election

Jiarui Lin^{1*}

¹ Keystone Academy, Beijing, China

* Jiarui Lin, Keystone Academy, Beijing, China

Received: August 17, 2026 Accepted: September 10, 2026 Online Published: October 5, 2026
doi:10.22158/csm.v9n1p166 URL: <http://dx.doi.org/10.22158/csm.v9n1p166>

Abstract

The spread of false information and political polarization on social media is often explained through the theories of “filter bubbles” and “echo chambers”. These models are applicable to information channels that rely on algorithmic recommendations, but their applicability is limited for private, encrypted, and user-driven platforms such as WhatsApp.

To examine how information spreads through private groups, social trust, and forwarding networks on digital platforms, this paper adopts a theory-guided case study approach, based on the spread of pro-Bolsonaro disinformation during Brazil's 2018 presidential election. The findings show the extent to which theories such as information disorder, echo chambers, filter bubbles, and populist communication can explain information dissemination on a private, networked interpersonal communication platform.

This paper proposes the concept of socially mediated echo chambers to extend explanations of information dissemination from the traditional emphasis on algorithmic filtering to social trust, private communication structures, identity, and limited correction mechanisms, which may also help explain information dissemination and polarization on closed or semi-closed digital platforms such as Telegram, WeChat, and Discord.

Keywords

disinformation, WhatsApp, information disorder, echo chambers, filter bubbles, populist communication, socially mediated echo chambers

1. Introduction

Digital platforms have increasingly been criticized for intensifying social polarization, which is often attributed to the effects of algorithmic filter bubbles and echo chambers. Yet these theoretical models assume open, feed-based platforms that rely on public recommendation systems, limiting their applicability to real-world cases beyond this scope. One such case is the 2018 Brazilian presidential election, which was engulfed in controversy due to the widespread dissemination of disinformation by supporters of then-presidential candidate Jair Bolsonaro, primarily through WhatsApp, a private messaging and calling app. Therefore, this paper explores the following research question: To what extent can established theories of information disorder, echo chambers, filter bubbles, and populist communication explain the spread of pro-Bolsonaro disinformation on WhatsApp during Brazil's 2018 presidential election? This paper argues that Brazil's 2018 election suggests political polarization on closed messaging platforms is driven not only by algorithmic filtering, but also by private-group forwarding, social trust, and populist anti-elite narratives.

2. Method

2.1 Information Disorder Theory

Since this study seeks to explain the spread of pro-Bolsonaro disinformation on WhatsApp, it first draws on information disorder theory, which begins by distinguishing among fake news, misinformation, and disinformation. The information disorder theory explains what information spreads, how it spreads effectively and persuasively through digital platforms, and the consequences for public opinion and trust (Wardle et al., 2017). The theory describes the modern polluted information ecosystem encompassing Misinformation, Disinformation, and Malformation (MDM) and identifies seven types of information disorder along a spectrum of harm, focusing on the agents (creators), messages (content), interpreters (consumers), phases of creation, production, and distribution, digital impact, societal threat, and solutions. The concepts of misinformation and disinformation are inherent in discussions of fake news. Fake news is defined as fabricated information that mimics the form of news media content but not in organizational process or intent (Wardle et al., 2017). Misinformation is false information created and spread without ill intent, while disinformation is deliberate in its creation and spread to inflict harm on entities which may include but are not limited to individuals, social groups, organizations, or countries (Wardle et al., 2017). Compared with true content, false content spreads “farther, faster, deeper, and more broadly” across all categories of information (Vosoughi et al., 2018). Vosoughi et al.'s (2018) study, “The Spread of True and False News Online”, found that novel information is more likely to be shared (retweeted on Twitter), and that fake news is significantly more novel, both objectively and as perceived by audiences, than true news across all novelty metrics. Their data further show that false rumors exhibit much greater informational uniqueness and elicit replies

expressing greater surprise and disgust. False political news, in particular, travels faster, “more broadly at greater depths”, reaches more people, and is more viral than any other category of false information (Vosoughi et al., 2018). Social technologies enable this spread by allowing rapid information sharing and large-scale information cascades. Together with other studies on social media, these findings suggest that emotional, identity-confirming content spreads particularly quickly.

2.2 Echo Chamber and Filter Bubble

The echo chamber and filter bubble theories are also important theoretical frameworks to analyze social media misinformation. Social media creates echo chambers, environments in which individuals are exposed primarily to information from like-minded people who share their views, and filter bubbles, personalized online information ecosystems created by algorithms that curate content based on individual user data, such as past behavior on the platform (Kitchens et al., 2020). Both echo chambers and filter bubbles reinforce prior beliefs while shielding the audience from opposing viewpoints, heavily driving polarization. Polarization here then becomes a self-reinforcing positive feedback loop through both algorithmic and users’ active choice in selective exposure, where reinforced prior beliefs lead to more radical stances, which become increasingly intensified from further reinforcement and lack of opposition (Bakshy et al., 2015).

2.3 Populist Communication

Populist communication is also critical to this paper’s analysis, as Bolsonaro was not merely a candidate in a polarized environment but also a practitioner of a right-wing populist communication style. Populist communication is a political strategy that frames politics as a moral struggle, dividing society between the “ordinary people” and the “culprit others” (Hameleers, 2018). It relies on creating a “us versus them” narrative to appeal to public sentiment. In the 2018 Brazilian case, as in many others, the “culprit others” took the form of the “corrupt elites” and the “corrupt system”. One common idea in populist communication is opposition towards the conventional, well-established media institutions, where anti-media narratives are constructed (Hameleers, 2018). This encourages people to turn towards social media and like-minded, and thereby “trustworthy”, people, such as right-wing politicians, for news and information (Hameleers, 2022). By appealing to people’s lived experiences and existing worldviews, and by invoking conspiracy theories and a sense of social or political crisis, populist communicators can present expert knowledge and empirical evidence as detached, biased, or untrustworthy. In doing so, they not only validate public distrust of established media but also extend that distrust to political, scientific, and other expert institutions more broadly.

2.4 Case Study of Brazil’s 2018 Presidential Election

Taken together, these existing frameworks (information disorder, echo chambers & filter bubbles, and populist communication) predict that algorithmically curated information environments reinforce ideological divisions and facilitate the spread of disinformation. Brazil’s 2018 presidential election

serves as a useful case through which to examine these dynamics, as its highly polarized political environment and surprising electoral outcome provide an opportunity to evaluate how well this theoretical model explains real-world phenomena. Following the end of Brazil's 21-year military dictatorship in 1985, two major parties dominated the country's political landscape, holding the presidency until 2018. Bolsonaro was an outsider - a congressman from a party with only a small political presence, who had comparatively limited public campaign funding and television airtime (Evangelista & Bruno, 2019). Nevertheless, as a populist figure with anti-establishment rhetoric, Bolsonaro won 55.1% of votes in the final round through a grassroots campaign and a massive social media presence (Welle, 2018).

WhatsApp has been the dominant communication platform in Brazil for a decade (Statista, 2022). It is a private, encrypted, group-based messaging and calling app. Studies have found evidence of the mass forwarding of misinformation involving presidential candidates through the app: in the months leading up to and throughout the election, the platform was engulfed in the circulation of misleading content (Dourado & Salgado, 2021). Before the election, public distrust of mainstream media was increasing, particularly among right-wing voters who viewed traditional outlets as aligned against their interests. This distrust was further intensified by Bolsonaro's campaign, which expressed the same view as its supporters. The campaign's disinformation was sensational and provocative, containing shocking or exciting claims that generated intense negative emotions among audiences, including anger and disgust. Yet these emotional responses were precisely what encouraged the disinformation to spread. This case therefore aligns well with the predictions of Vosoughi et al.'s (2018) study and Wardle and Derakhshan's information disorder theory that emotional, identity-confirming content spreads far faster and reaches more people.

In the context of this election, its disinformation may be analyzed by understanding the "for" or "against" the Workers' Party (PT) divide. Pro-PT content, centered on Luiz Inácio Lula da Silva and Dilma Rousseff, supports social welfare and economic redistribution, is progressive on social issues such as LGBTQ+ rights and abortion, and advocates a reformist approach to public security (Rennó, 2020). Its support base is concentrated in the Northeast and among lower-income voters and Catholics. Meanwhile, anti-PT content, led by Jair Bolsonaro and allied right-wing movements, supports privatization and economic liberalism, is conservative on social issues, including opposition to abortion and support for the "traditional family", and advocates tough-on-crime policies, gun rights, and the death penalty. Its support base is concentrated in the South and Southeast and among evangelical Christians and affluent voters. Anti-PT supporters view the PT as corrupt, ideologically radical, and responsible for the country's economic crises.

During Bolsonaro's campaign, there were several types of information disorder (in this case, all disinformation circulating on WhatsApp): fabricated claims and rumors including how Lula da Silva

“had a cement safe-deposit box larger than a swimming pool” (Dourado & Salgado, 2021); fabricated images including a graph showing that families’ consumption decreased while Brazilian government’s expenditure increased over the last five years under PT government; imposter content including an alleged criminal record of ex-president Dilma Rousseff during times of military dictatorship in which she would be accused of being a terrorist and bank robber (Evangelista & Bruno, 2019); false context including the image of two men making out shown with the sentence “Globo who do not support Bolsonaro!!!” which was actually a record of the Burning Man music festival that took place in the Desert of Nevada in the US in 2011 (Evangelista & Bruno, 2019).

Through directing disinformation toward people whose ideologies are already aligned—for example, conservative anti-abortion and traditional family-related content toward Anti-PT supporters—the result is receptive audiences whose existing beliefs are continuously reinforced and thereby intensified, becoming more extreme. Moreover, the growing distrust of traditional news media among Anti-PT supporters leads people to seek out and rely on alternative information ecosystems on unregulated, yet faster-moving, social media platforms, where levels of disinformation and misinformation are significantly higher. In turn, the populist messaging of the “corrupt elites” and “fake mainstream media” promoted through Bolsonaro’s extensive social media campaign strongly reinforces, and in some cases establishes, people’s in-group versus out-group thinking. Therefore, this case appears to confirm the predictions of echo chamber and populist communication models.

However, Bolsonaro’s election presents a case that does not fully align with the algorithmic filter bubble model, which has long been used to explain the circulation of disinformation on social media. WhatsApp is not a typical algorithmic feed platform. The app has no public ranking feed like those of Facebook or Twitter. Instead, content spreads through forwarding chains and group chats, making visibility socially mediated rather than algorithmically curated. Current research focuses on polarization driven by algorithmic filtering, yet polarization in this case occurs without such algorithmic curation.

Thus, the current theoretical models may not sufficiently account for some mechanisms. The circulation of disinformation on WhatsApp is a very successful polarization process, for several reasons. The first is social trust, meaning that information passed on by friends, family, and even acquaintances is trusted, or at least more trusted than information from other sources. The second is WhatsApp’s closed network model, which is a private, encrypted messaging and calling application. There is no built-in mechanism to fact-check or counter misinformation, unlike X’s Community Notes, which provide context for potentially misleading posts. Moreover, Brazilians are further disincentivized from fact-checking because of the country’s “zero-rating” mobile network system, under which users can access certain messaging and social media apps, including WhatsApp, without consuming their mobile data allowance, while access to the broader internet still requires data (Evangelista & Bruno, 2019). As

a result, many Brazilians rely primarily on these platforms, where unreliable information circulates with little opportunity for contemporaneous fact-checking (Deen, 2025). Third is the private virality of information on WhatsApp: information spreads largely outside the view of the broader public and government authorities, thereby limiting opportunities for public correction. Fourth is the coordinated mass messaging and forwarding of provocative, emotionally charged, and politically aligned content (Evangelista & Bruno, 2019). This large-scale dissemination structure further reinforces people's existing views, intensifying ideological polarization.

3. Result: Revised Theoretical Frameworks

Thus, the aim of this paper is to complement the current academic literature by suggesting the modification of the widely used algorithmic filter bubble model and extending it to a larger model of "socially mediated echo chambers". This model aims to be a theoretical extension of the description and explanation of algorithmic filtering by also incorporating aspects of interpersonal trust networks, private communication networks, and group identity dynamics.

My new model can be divided into four parts, each of which is dependent upon the earlier. The first is the current ideology alignment theory, which is created and strengthened by a combination of the mechanisms outlined in information disorder theory, echo chamber theory, filter bubble theory, and populist communication. The second depends on the nature of the platform itself, specifically how platform-specific features and design shape the spread of information. In the case study explored in this paper, the users can communicate through private group chats, transfer messages to a lot of people at a fast pace, and utilize end-to-end encryption, which makes it difficult for outsiders to see and moderate the discussion. These affordances make it harder to trace back to the origin and routes of misinformation, allowing it to propagate more quickly and without as much public oversight as on other platforms that are more public, like X (formerly Twitter).

Within these platforms like WhatsApp, the third component is social trust amplification from a psychological perspective: peer-to-peer sharing of information on a platform perceived as personal and private naturally generates credibility. Finally, the fourth component consists of reduced correction mechanisms on private messaging platforms, resulting in a continuous cycle of limited exposure to opposing viewpoints and, in some cases, such as Brazil's 2018 presidential election, to factual information itself. Ultimately, this proposed model of socially mediated echo chambers aims to explain the circulation of disinformation on WhatsApp more effectively than the classic algorithmic filter bubble model.

4. Discussion

Although the proposed socially mediated echo chamber model was developed through extensive research and analysis of Brazil's 2018 presidential election, it is not limited to this single case. Rather,

the model can be applied across a range of contexts, extending to Telegram groups, WeChat circles, and private Discord communities, all of which are private communication platforms similar to WhatsApp. However, the mechanisms through which polarization develops may still differ across platforms. For example, while WhatsApp limits group chats to up to 1,024 members, Discord's servers are designed to support worldwide communities, with a default limit of 25 million (WhatsApp, n.d.; Discord, 2026). By contrast, WeChat groups are limited to 500 members and impose additional restrictions, including requiring confirmation before inviting new members once a group exceeds 40 people, requiring identity verification once a group exceeds 100 members, and limiting group QR codes to a validity period of seven days (Tencent, 2025). Overall, polarization in increasingly digital environments may depend on hybrid algorithmic-social systems rather than algorithmic mechanisms alone.

To conclude, this paper serves as a case study of the spread of pro-Bolsonaro disinformation on WhatsApp during Brazil's 2018 presidential election. It finds that existing theories, including information disorder, echo chambers, filter bubbles, and populist communication, are useful but incomplete, thereby emphasizing the importance of accounting for both platform structures and social dynamics. As a result, this paper proposes a revised model: the socially mediated echo chamber model. Future research should continue to integrate technological and sociological perspectives, adopting an interdisciplinary approach that has become increasingly important in understanding contemporary information environments.

Acknowledgement

The author thanks Dr. Melissa Li of Columbia University for her mentorship and guidance on this paper. The author hopes to continue researching the impact of social media on international politics and economics.

References

- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130-1132. <https://doi.org/10.1126/science.aaa1160>
- Deen, A. (2025, June 2). *The High Cost of "Free" Data: Zero-Rating and its Impacts on Disinformation in Brazil*. Data-Pop Alliance. Retrieved from <https://datapopalliance.org/the-high-cost-of-free-data-zero-rating-and-its-impacts-on-disinformation-in-brazil/>
- Discord. (2026, May). *Discord account caps, server caps, and more*. Retrieved from <https://support.discord.com/hc/en-us/articles/33694251638295-Discord-Account-Caps-Server-Caps-and-More>

- Dourado, T., & Salgado, S. (2021). Disinformation in the Brazilian pre-election context: Probing the content, spread and implications of fake news about Lula da Silva. *The Communication Review*, 24(4), 297-319. <https://doi.org/10.1080/10714421.2021.1981705>
- Evangelista, R., & Bruno, F. (2019). *WhatsApp and political instability in Brazil: targeted messages and political radicalisation*. Internet Policy Review. Retrieved from <https://policyreview.info/articles/analysis/whatsapp-and-political-instability-brazil-targeted-messages-and-political>
- Hameleers, M. (2018, May 14). *A typology of populism: Toward a revised theoretical framework on the sender side and receiver side of communication*. Retrieved from <https://ijoc.org/index.php/ijoc/article/view/7456>
- Hameleers, M. (2022). Empowering the people's truth through social media? (De)Legitimizing truth claims of populist politicians and citizens. *Politics and Governance*, 10(1), 210–219. <https://doi.org/10.17645/pag.v10i1.4726>
- Kitchens, B., Jr., Johnson, S., Gray, P., & University of Virginia McIntire School of Commerce. (2020). Understanding Echo Chambers and Filter Bubbles: The Impact of Social Media on Diversification and Partisan Shifts in News Consumption. *MIS Quarterly*, 44(4), 1619-1649. <https://doi.org/10.25300/MISQ/2020/16371>
- Rennó L. R. (2020). The Bolsonaro Voter: Issue positions and vote choice in the 2018 Brazilian presidential elections. *Latin American Politics and Society*, 62(4), 1-23. <https://doi.org/10.1017/lap.2020.13>
- Statista. (2022). Social media usage in Brazil. In *Statista*. Retrieved from <https://www.statista.com/study/68696/social-media-usage-in-brazil/>
- Tencent. (2025, February 10). *Add friends into a group chat*. WeChat Help Center. Retrieved from <https://cs.help.wechat.com/hc/en-us/articles/11917722718351-Add-friends-into-a-group-chat>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. <https://doi.org/10.1126/science.aap9559>
- Wardle, C., Derakhshan, H., Council of Europe, Burns, A., & Dias, N. (2017). INFORMATION DISORDER : Toward an interdisciplinary framework for research and policy making. In *Council of Europe Report*.
- Welle, D. (2018, October 28). Bolsonaro wins Brazil's presidential election. *DW*. Retrieved from <https://www.dw.com/en/far-right-jair-bolsonaro-wins-brazils-presidential-election/a-46065594>
- WhatsApp. (n.d.). *How to join a group in a community*. WhatsApp Help Center. Retrieved from https://faq.whatsapp.com/967457667545238/?cms_platform=web