Online Journal of Communication and Media Technologies, 2026, **16**(2), e202612
e-ISSN: 1986-3497

OPEN ACCESS

Review Article

Social media, fragmentation, and polarization in the *platformized* public sphere: Making the case for WhatsApp research

Javier Martín Merchán [1*]
0009-0003-7674-7215

Carlos Rico Motos [1]
0000-0001-6808-8262

[1] Department of International Relations, Universidad Pontificia Comillas, Madrid, SPAIN
[*] Corresponding author: jmmerchan@comillas.edu

Citation: Martín Merchán, J., & Rico Motos, C. (2026). Social media, fragmentation, and polarization in the *platformized* public sphere: Making the case for WhatsApp research. *Online Journal of Communication and Media Technologies, 16*(2), Article e202612. https://doi.org/10.30935/ojcmt/18142

ARTICLE INFO

Received: 3 Dec 2024
Accepted: 30 Nov 2025

ABSTRACT

As digital platforms increasingly mediate public communication and access to information, they do not only redefine the structural makeup of the public sphere but also impact the practices and outcomes of political deliberation, often giving room to concerns about heightened fragmentation and polarization. This paper critically assesses the link of social media with these two phenomena, juxtaposing traditional social networks with the rising prominence of private social media. Drawing from an extensive literature review, we argue that existing research is primarily focused on two algorithm-led public social media, Twitter and Facebook, which possess pro-diversity, weak-tie, heterophilic affordances sometimes argued to counter segregation and polarization. Non-algorithm-led private social media, such as WhatsApp, rather have a strong-tie, privacy-securing, homophilic architecture that connects with individuals' tribalist drives and could better facilitate sectarian interactions and polarizing outcomes. Moreover, they rank among the most utilized platforms worldwide and incarnate prominent spaces for engaging with politics. However, their relationship with political talk fragmentation and polarization is comparatively under-researched. Focusing on the paradigmatic WhatsApp, we re-scrutinize the scarce corresponding literature, identify certain issues of concern, and call for the strengthening of a nascent research agenda around this platform. Our reviewing exercise emphasizes the importance of diversifying research to encompass the entire spectrum of social media's influence on democratic processes.

Keywords: fragmentation, *platformization*, polarization, social media, WhatsApp

INTRODUCTION

At least normatively, democratic politics relies on the public sphere, a robust space for society-state negotiation and the continuous formation of public opinion within an associated media landscape (Rosa, 2022, pp. 17-18). This space has undergone various transformations (Habermas, 1962, 2006). Today, at a time when public communication and access to information are mediated by digital platforms driven by profit-maximizing companies, it seems reasonable to argue that we traverse a new –digitally rooted– paradigm (De Blasio et al., 2020). This is what some scholars have called the "new structural transformation" (Seeliger & Sevignani, 2022; Stewart & Hartmann, 2020) or the *platformization* (Smyrnaios & Baisnée, 2023) of the public sphere.

A preliminary version of this article was presented at the "Politics and Emotions" panel of the 17th Congress of Spain's Political Science Association.

Beyond formal conceptualizations, *platformization* is ultimately a transformation of what Dahlgren (2005) terms the "structure" of the public sphere. An evolution from a situation in which mass media channeled public communication flows to another one in which those flows are inevitably mediated by platforms that encourage the unconstrained sharing of many themes among many users. The centrality of platforms, which do not perform the journalistic tasks of content selection and treatment but excel as info-mediation tools (Smyrnaios, 2015), has fragmented publics along numerous networks that now compete with mass media for attention while eroding the latter's monopoly in shaping the limits of public discourse (Gillespie, 2018). Simultaneously, this changing scenario conditions the manners in which we access/disseminate political information and interact with others' views (Smyrnaios & Baisnée, 2023). Thus, the institutional (re)organization of the public sphere does not only provide a new technical framework for the development of public communication, but also influences deliberation practices and outcomes (Habermas, 2022, p. 157).

In this sense, numerous scholars (Persily & Tucker, 2020; Sunstein, 2018; Terren & Borge-Bravo, 2021) have alerted about today's *social-media-led* homophilic tendency by which people constrain themselves to like-minded communicative channels (e.g., "echo chambers"), which set the basis for more polarized polities. The existence of separate niches for separate groups is not new, but, as Rosa (2022) reasons, digital spaces like echo chambers enjoy a unique capacity to hinder reflection upon experiences occurring within a shared lifeworld. "Post-truth politics" is arguably founded on this panorama: if the diverging lifeworld's daily inhabited by citizens do not overlap sufficiently, each of those universes logically develops its own reality (McKay & Tenove, 2021; van Dyk, 2022). Thus, homophilic fragmentation and political polarization have ultimately emerged as two characteristic biases of the process of public will-formation in a platformized public sphere.

Scholars have largely embarked on the task of empirically examining whether and to what extent social media usage affects the consolidation of fragmented spaces and increases polarization. This paper provides a thoughtful review of this state of the art, reflecting on some of its inconsistencies, introducing new considerations for a more refined understanding of it, and eventually suggesting new routes of inquiry. There are some systematic reviews on this matter (Kubin & von Sikorski, 2021; Lorenz-Spreen et al., 2023; Terren & Borge-Bravo, 2021). Yet, this piece is not a formal meta-analysis; rather, we approach this reviewing exercise from a novel angle, deepening on the burgeoning research path that highlights the centrality of non-algorithm-led private social media. To do so, we take WhatsApp as paradigmatic case.

Specifically, we argue that current research on the relationship between social media and political fragmentation/polarization mainly relies on two platforms, Twitter (renamed "X") and Facebook. These are led by an algorithmic rationale that might ultimately counter humans' tendency towards sectarianism, given such platforms' pro-diversity weak-tie and heterophilic affordances. However, as cross-platform research (Boulianne et al., 2024; Yarchi et al., 2021) has anticipated, findings from these public social media are hardly extrapolatable to private social media or instant messaging apps, such as WhatsApp. The latter lacks that algorithmic rationale, and their strong-tie, privacy-securing, and homophilic-friendly affordances seem to connect optimally with individuals' tribalist drives, hence facilitating their segregation and ingroup reaffirmation. WhatsApp plays a central role within the wider media ecosystem, as it ranks among the most utilized platforms worldwide and incarnates a relevant space for engaging with politics (Masip et al., 2021; Valeriani & Vaccari, 2018). However, we still know little about how political talk arises in WhatsApp, and how it relates to fragmentation and polarization. In this light, this paper calls for the strengthening of a nascent research agenda.

Before proceeding, it is important to make some clarifications to define the scope of this work. The term "social media" is quite disputed and potentially encompasses a wide spectrum of platforms. We utilize here McCay-Peet and Quan-Haase's (2017, p. 46) broad definition of social media as Internet-based "services that allow people, communities, and organizations to collaborate, connect, interact, and build a community by enabling them to create, co-create, modify, share, and engage with user-generated content," which turns these sites into spaces for addressing politics (Ruess et al., 2023). This formulation accommodates a variety of platforms, such as Twitter, Facebook, Instagram, WhatsApp, Telegram, TikTok, Reddit, Snapchat, or YouTube, just to mention some examples. WhatsApp thus fits within the definition and has already been considered as a social media in various works (Boulianne et al., 2024; Chadwick et al., 2021; Rossini et al., 2021b; Yarchi et al., 2021). Yet, there may be critics still questioning the extent to which WhatsApp qualifies as a social media *stricto sensu*, stating that its underrepresentation in this literature relates to its problematic

accommodation within the "social media" category. We do not engage in this objection nor in further conceptualization debates, as this goes beyond our purpose and the key points we aim to convey. We simply posit that, be platforms like WhatsApp called "(private) social media," "messaging apps" or else, they are spaces enabling political encounters amidst everyday interactions, diverge from public social media in various dimensions, and have however received much less attention in this academic field.

Having said this, the article proceeds as follows. First, it highlights the issues of the existing literature on the link between social media and fragmentation/polarization from a twofold perspective: the overreliance on Twitter/Facebook and the burdens associated with algorithmic-centric approaches. Second, it introduces WhatsApp –and private social media– and points to the distinct, non-algorithm-led interplay of its affordances with humans' tribal nature. Third, it reviews extant knowledge on WhatsApp and reflects on the shortfalls of this literature, especially surrounding its relationship with fragmentation/polarization. The last section concludes.

TWITTERIZATION VS. PRIVATE SOCIAL MEDIA: UNDERSTANDING (THE PROBLEMS OF) EXISTING EVIDENCE

Although social media have often been accused of generating homophilic silos (Terren & Borge-Bravo, 2021), the empirical literature provides a more nuanced picture. Exposure to diversified political content on these sites tends to be greater than on classic and other digital media (Barberá et al., 2015); only a tiny portion of the population places itself within one-sided echo chambers (Fletcher et al., 2021); and algorithms recommending news in these platforms do not influence significantly the ideological equilibrium of the absorbed information (Haim et al., 2018). Moreover, social media boost exposure to the content shared by weak ties, who are more prone to post ideologically diverse news. In this light, the link between social media and polarization is ultimately ambiguous, with evidence pointing towards both escalating and decreasing trends (e.g., Tucker et al., 2018).

We detect a problem when addressing this literature though: rather than being the fruit of extensive research conducted at a cross-platform level, it relies disproportionately on Twitter and Facebook, thus virtually equating available evidence on social media with evidence on these two platforms. Indeed, despite some calls for cross-platform research (Kubin & von Sikorski, 2021; Tucker et al., 2018), so far, an overwhelming majority of studies on the impact(s) of social media mobilizes Twitter and, at the most, Facebook data exclusively. It cannot be ignored that there is an increasing interest in approaching alternative platforms, such as YouTube (Brown et al., 2022) or Reddit (Waller & Anderson, 2021). Nevertheless, as **Table 1** presents, whereas the works examining some key political issues from TikTok, Instagram, or WhatsApp use amount to 2.4%, 6.7%, and 4.5% of total research, respectively, the investigations focusing on Twitter and Facebook surpass 70%. Specifically, in terms of fragmentation (polarization) outcomes, Twitter and Facebook studies represent 60% (80%) of the total. For Cea and Palomo (2021), this imbalance is due to the fact that these platforms' application programming interface (API) made it easier for researchers to scrape (millions of) data from them.

The substantive consequence of *twitterizing* social media studies is that we may be tempted to base our understanding of social media and their polarizing potential on our understanding of Twitter and Facebook. This generalization seems problematic though. In fact, while cross-platform research is still embryonic, existing studies (Bossetta, 2018; Cinelli et al., 2021; Kligler-Vilenchik et al., 2020; Lee et al., 2021; Rossini et al., 2021b; Yarchi et al., 2021) have already alerted that social media diverge in their features and functioning, and that so they might do regarding their political effects.

A note turns critical at this stage. There is a feature that needs to be considered given its centrality to define the very possibilities of any given platform: their (non-)algorithmic rationale. Indeed, algorithmic filtering is one of the pillars of Bossetta's (2018) "digital architectures" scheme, which has been suggested to mediate political communication and behaviors on social media. Most of the latter, including Twitter or Facebook, are algorithm-led social networking sites, meaning that the algorithm operates as an active subject of the outcomes deriving from the use of those platforms. Thus, when considering whether Facebook favors fragmentation, we are ultimately considering whether this platform *creates* echo chambers; that is, whether

Table 1. Scientific works related to selected political issues in the Scopus database by social media

	Fragmentation	Polarization	Misinformation	Populism	Trust	Participation	Works per platform (all the outcomes)
Facebook	385 (19.2%)	299 (21.8%)	1,497 (23.3%)	294 (35.5%)	1,539 (40.4%)	525 (40%)	4,539 (28.8%)
Twitter	786 (39.3%)	797 (58.1%)	2,993 (46.5%)	355 (42.9%)	1185 (31.1%)	506 (38.5%)	6,622 (42%)
Instagram	98 (4.9%)	55 (4.0%)	418 (6.5%)	78 (9.4%)	348 (9.1%)	59 (4.5%)	1,056 (6.7%)
TikTok	19 (1.0%)	29 (2.1%)	213 (3.3%)	13 (1.6%)	91 (2.4%)	20 (1.5%)	385 (2.4%)
Snapchat	5 (0.2%)	5 (0.4%)	23 (0.4%)	0 (0.0%)	33 (0.9%)	4 (0.3%)	70 (0.4%)
Reddit	61 (3.0%)	64 (4.7%)	194 (3.0%)	8 (1.0%)	74 (1.9%)	11 (0.8%)	412 (2.6%)
YouTube	589 (29.4%)	80 (5.8%)	608 (9.4%)	54 (6.5%)	321 (8.4%)	111 (8.4%)	1,763 (11.2%)
WhatsApp	**32 (6) (1.5%)**	**28 (12) (2%)**	**393 (6.1%)**	**10 (1.2%)**	**184 (4.8%)**	**63 (4.8%)**	**710 (4.5%)**
Telegram	26 (1.3%)	15 (1.1%)	97 (1.5%)	16 (1.9%)	47 (1.2%)	15 (1.1%)	216 (1.4%)
Total works per outcome	2,001 (100%)	1,372 (100%)	6,436 (100%)	828 (100%)	3812 (100%)	1,314 (100%)	15,763 (100%)

Source: Own elaboration, with data from Scopus.

Please consider the following remarks to better understand **Table 1** figures:

- We look for both the relevant platform and the political issue in question in the "article title, abstract, and keywords" section.
- In our search, we identify each of the relevant political outcomes as follows:
  - Fragmentation: fragmentation OR segregation OR segmentation OR echo chamber
  - Polarization: polarisation OR polarization
  - Misinformation: misinform* OR disinform* OR fake news OR hoax*
  - Populism: populis*
  - Trust: trust OR political disaffection
  - Political participation: political participation.
- Note of caution: When following this search strategy, inevitably we come up with some works that do not necessarily and properly tackle our outcome of interest. For example, when searching "whatsapp AND populis*", there may be papers addressing some dimensions of WhatsApp unrelated to populism, but whose abstract include formulae such as "In a time dominated by populisms…" The only way out of this shortfall resides in the human curation of the Scopus results. Given that the focus of this paper is on WhatsApp and fragmentation/polarization outcomes, we have conducted this task just for the specific cells that link WhatsApp with works on fragmentation and polarization. Therefore, in the polarization column, "28" is the non-depurated number of works provided by the Scopus search, while "(12)" accounts for the number of investigations actually evaluating the relationship between WhatsApp and polarization. Despite this limitation, **Table 1** figures are still useful for our comparative purposes, as this bias should distribute itself equally across searches involving all the platforms.
- Note of caution II: As stated, our search strategy does not aim to provide a refined idea of the exact number of studies properly addressing each outcome but to give a broad overview of the cross-platform unbalance of the current state of the art. Hence, the total number of works covering these platforms and outcomes is unlikely to be 15763, for example, as there may be works repeatedly counted across several cells (if they address several platforms or outcomes). In other words, numbers should not be read in an exclusionist manner: the fact that 797 papers cover some issues related to Twitter and polarization does not prevent some of them from dealing with polarization in Facebook at the same time, for instance.

Facebook's algorithm forms echo chambers, or whether the interplay between our partisan leanings and that algorithm results in the formation of those bubbles. The same applies if we consider whether Twitter favors inflammatory content. Hence, in these platforms, individuals hold a somewhat passive position, or, more accurately, they cannot but share their agency with the algorithm. Of course, the weight of algorithmic sorting diverges across platforms: Facebook's algorithm filters content heavily, for instance, while Twitter's seems lighter in this respect (Bossetta, 2018, 2023). The point here, though, is that they all incorporate some algorithmic incidence that mediates human behavior.

In contrast, encrypted messaging apps like WhatsApp or Telegram are not led by any algorithm. In these cases, it is the individual herself that functions as the subject of the outcomes associated with social media use. **Table 2** adopts Bossetta's (2018) framework and makes this point even clearer by comparing WhatsApp with other popular platforms. In this vein, if we aim to consider whether WhatsApp favors fragmentation, for instance, we will no longer consider whether this app *creates* echo chambers, but whether individuals self-select into echo chambers within the platform. This is a little but important nuance: the capacity of a platform

Table 2. Degree of algorithmic filtering by platform (Bossetta, 2018)

	Algorithmic filtering	
	Reach	**Override**
Facebook	Heavily filtered (relevance)	Pay to promote and user-diffusion (sharing)
Twitter/X	Moderately to heavily filtered (chronology and relevance)	Pay to promote, index via hashtags, and user-diffusion (retweeting)
Instagram	Moderately to heavily filtered (chronology and relevance)	Pay to promote and index via hashtags
WhatsApp	None	No algorithm to override

Source: Own elaboration, based on Bossetta (2018).

to *produce* an outcome (sometimes in interaction with human behavior) is not the same as its suitability for a specific human behavior that produces that outcome. Therefore, the problem is not only that there exists an overrepresentation of evidence coming from Twitter and Facebook, but also that the two platforms from which that evidence is grasped belong to the specific niche of "algorithm-led" sites.

From this disquisition it is apparent that different platforms might inspire different research questions and yield different results. We acknowledge that it is precisely the algorithmic nature of Facebook and Twitter that has led scholars to explore individuals' behavior on those platforms and the subsequent effects on trends like polarization, fundamentally because algorithms make these fora clear contexts to study by *uniting* experiences, behavior, and effects. However, private social media are universally widespread and inserted in our quotidianity, very often enjoying wider audiences than public social media themselves (Newman et al., 2022); they have already been proved central loci for online engagement with political content (Masip et al., 2021); as we argue in the coming section, they represent a peculiar environment with characteristics diverging from traditional public social media's, which may lead us to expect distinct outcomes too; and, despite this, we still know little about their possible interplay with fragmentation/polarizing trends (see **Table 1**). Hence, in a panorama coopted by Twitter and Facebook, our argument is that there is a strong rationale to start tackling the social media-fragmentation/polarization link by scrutinizing the singular –but academically disregarded– private social media.

The paradigmatic case when referring to these services is WhatsApp, which ranks as the third most utilized platform globally (We Are Social & Meltwater, 2024), with more than 2 billion users, and is already the most popular platform in Europe (Newman et al., 2022) and many Global South countries (We Are Social & Meltwater, 2024).[1] Indeed, this app has outranked other sites that receive greater attention in the academic and public debate. Against this backdrop, we propose that cultivating WhatsApp research could enrich extant literature. Of course, this is not to downplay other private social media like Telegram or WeChat, which are not as popular in Western democracies but also have millions of users. We equally encourage the academic community to dive into those. For the sake of practicality, we focus on WhatsApp to delimit our scope, considering that it is by far the most used platform of the kind, and that it also possesses a significantly wider, cross-country reach compared to WeChat or Telegram, whose utilization disproportionately excels in China and Russia, respectively (We Are Social & Meltwater, 2024).

AFFORDANCES, TRIBALISM, AND THEIR INTERPLAY: WHY EVIDENCE FROM TWITTER MAKES SENSE FOR TWITTER BUT NOT FOR WHATSAPP

By no means has the validity of extant findings about Twitter/Facebook been questioned. We simply contend that the overrepresentation of those platforms should not tempt us into thinking that such evidence is applicable to the wide corpus of social media, especially when different rationales operate, and different dimensions of the social media-fragmentation/polarization relationship are invoked. Indeed, let us now take

[1] For instance, in Europe, 62% and 20% of citizens use WhatsApp for general purposes and political news, respectively, which contrasts with the less than 30% and 10%, respectively, who use Twitter. Facebook usage ratios resemble those of WhatsApp; however, the gap between these platforms may expand if considering that, among 18-to-24-aged Europeans, only 50% use Facebook versus the nearly three quarters that engage with WhatsApp. Also, WhatsApp is often among the top three platforms for the consumption of political information in every country, and, with exceptions (France, Greece), it has become the most widely used platform in general, with over 70% Spaniards, Italians, Austrians, or Germans daily interacting in this site (Newman et al., 2022).

Twitter and WhatsApp as examples for comparison purposes. Provided that Twitter constitutes the starting point of the analysis, social media could be characterized as plural spaces furthering diversity thanks to, and not despite, their affordances and digital architecture. Accordingly, if WhatsApp's affordances/architecture deviate significantly from those, as we argue, the fragmentation/polarizing implications of the platform may also diverge, to the point of articulating a paradigmatic disintermediated space for the exacerbation of human drives towards political segregation and outgroup aversion. And it is this singular representation of WhatsApp that calls for addressing the platform.

Let us delve into this. Twitter is distinguished by its unrestrained publicity, visibility and searchability (Bossetta, 2018). Unless blocked, everyone, including non-users, could read –and often answer to– any tweet, with individuals being able to reach most pieces of content by utilizing the system's search function. People are enabled to follow each other without asking for authorization, enhancing the probabilities of contacting people beyond their closest ties, and giving room for asymmetric linkages (Yarchi et al., 2021, p. 7). This makes it possible for individuals to follow and interact with others without turning into members of any community, even when such others locate outside their network of friends. All of this makes contending political identities, ideologies and demands to be easily perceivable (Yarchi et al., 2021, p. 7). Though the platform also favors some content personalization and homophilic interaction, as remarked above, this does not offset its virtues for heterophilic exposure and/or conversations (Vaccari & Valeriani, 2021, pp. 24-27). In this vein, Twitter's news feed might exemplify the extent to which users are exposed to an uncontrollable mêlée of messages launched by political and non-political contacts (Bruns, 2022).

Twitter thus offers unique possibilities to interact with information/contacts from competing ideological sides, especially by exposing users to their weak ties, who have been proved to disseminate more "novel" and ideologically varied information (Granovetter, 1973). The platform's asymmetric networks, which advance diversity via those weak ties, and the (spontaneous) exposure to cross-cutting information algorithmically promoted by its news feed hinder the consolidation of totally hermetic communities of like-minded users. This does not preclude any prospect of polarization, but the encounter of perspectives from outside one's political community (Beam et al., 2018) makes those works that challenge the polarizing potential of social media reasonable. Lastly, Twitter's publicity exposes any message to public scrutiny. This is not a guarantee against misinformation posting, especially if anonymous or fake profiles are used. In fact, though evidence is again mixed, some reports (Newman et al., 2022) identify Twitter as the most concerning platform for spreading misinformation. Nonetheless, when individuals use their personal account, they tend to care for the quality of their posts –and even opt for not posting at all– to avoid the harsh scrutiny of their online communities and "losing their 'outside-Twitter' status" (Bail, 2022, pp. 70-73). Therefore, Twitter's publicity might also work as a deterrent for the sharing of dubious content, at least for some users.

A note turns critical at this stage. There exists a rich literature showing the prominent role that (political) identities play in organizing collective behavior and defining electoral decisions, media diets, and social preferences (Achen & Bartels, 2017; Campbell et al., 1960; Torcal et al., 2018). Thus, individuals have often been shown to be attached to some identity, exhibit ingroup bias and behave accordingly. This is important to understand why Twitter's architecture and affordances do not inherently lead to less fragmentative experiences or lesser polarization (Vaccari & Valeriani, 2021), but they do coherently allow for such outcomes. They do so precisely because, to some extent, they arguably *paddle* against our identity-based, tribalist drives: Twitter's logics and algorithms expose individuals to a higher degree of diversity than what they would naturally be outside the platform, and this can sometimes result in a better understanding of the other side's positions, the moderation of viewpoints, or the downsizing of aversion to political adversaries (Pettigrew & Tropp, 2006). However, what should be expected from a digital space like WhatsApp in which no algorithm exists that can display greater diversity and connections than those faced offline? It would be unsurprising if individuals reproduced their biases in such a forum, opting for self-segregating along ideological lines when it comes to discussing politics. Higher polarization could then also be the product of the self-reinforcing effect of anti-outgroup, partisan back feed in those spaces.

Let us look into the latter using WhatsApp as a paradigmatic case, as anticipated above. As a private message space lacking an algorithmic rationale, WhatsApp neither filters nor recommends content. The platform is used to chat with closer contacts (especially strong ties), easily allowing for connection between like-minded individuals through chat groups (Santini et al., 2021). Indeed, chat groups inherently revolve

around a shared concern, such that, at least regarding that concern, groups become homophilic by default. The fact that users self-select across groups –as a result of their own will to do so– encourages a higher degree of homophily where group participants are more likely to share some viewpoint that they expect to reinforce. WhatsApp is also characterized by its non-publicity. Any within-group interaction demands users to have access to the group in the first place. Inside the group, messages turn visible for all without algorithmic filters, and the only constraint to utterly free exposure is defined by the sequential structure of the conversation (Yarchi et al., 2021, p. 8). Moreover, WhatsApp ensures message encryption; content often circulates in a largely unmoderated manner, which hinders the tracking of illegitimate activities, and shared messages are not easily traceable to a source (Santini et al., 2021).

WhatsApp thus projects a playing field notably deviant from that of Twitter. The facilities to interact with strong ties as well as the structurally homophilic composition of groups, around which individuals self-select due to their interest in a theme, provide a suitable window of opportunity for the fragmentation of the public space along the lines of collectives who opine alike. On the other hand, the non-public and encrypted nature of the platform prevents the restriction of the information disseminated through it, which serves as an opportunity to propagate whichever content confirms individuals' views regardless of its authenticity.

Ultimately, WhatsApp's architecture and affordances may pave the way for the creation of information bubbles accessible to those who aspire to reaffirm their prejudices. Closed spaces where individuals rarely engage in self-censorship behaviors, and where they are, by default, in contact only with the information they themselves provide each other. Upon these foundations, the platform can epitomize a suitable window of opportunity for the cementation of *communities of belief* –and the growing polarization of opinions– amidst today's post-truth context in which divergent versions of reality compete with each other (van Dyk, 2022). We do not argue that WhatsApp influences people to behave in a certain way; just as Twitter's, WhatsApp's (non-algorithmic) features do not automatically entail more fragmentative or polarizing outcomes, but they make way for these effects. And, in this case, they do so because the absence of an algorithm that may counterbalance selective exposure through enhanced informational diversity clears the path for identitarianism and tribal in-grouping. Also, the specific facilities for strong-tied connections and encrypted, non-edited conversations, if anything, *paddle in favor of* our tribalist drives. Individuals, who are often selective by default in choosing their media diets or political interactions to circumvent dissonance and challenges to their partisan loyalties/identities (Achen & Bartels, 2017), are yielded with a space that enable them to connect with their like-minded strong ties in real time to discuss politics in an intimate environment distant from public scrutiny and hence more prone to outrage and laughter discourse. Thus, when news are shared there, they are likely to align with group members' common political attitudes, and, if they do not, at least they are likely to be processed collectively in attitude-consistent ways (Lottridge & Bentley, 2018). Moreover, in small groups of friends, WhatsApp users might be less concerned with managing impressions or opinions, as they already have a strong familiarity with one another (Bayer et al., 2016). Consequently, individuals may be unworried about adjusting their tone or language to accommodate the preferences of diverse, (semi-)public audiences (Child & Starcher, 2016).

This reflection, by the way, gives a sense of the singularity of private social media vis-à-vis face-to-face interactions too. We observe, at least, three unique attributes underlying online personal messaging. First, its ubiquitous character, as platforms like WhatsApp afford sempiternal contact and communication at every single point in time, regardless of physical distance, which allows for exploiting the "juice" of any circumstance by talking about issues which would otherwise be disregarded. Second, communication in these platforms is *mediated* by the particular language, tone, register, and tools that characterize exchanges in digital intimate environments, which encompasses memes, stickers, and higher doses of ridicule and sensationalism evoking emotional responses and removing contextual information essential for political learning (Baulch et al., 2024b; Rowlett & Harlow, 2018). Third, though some WhatsApp conversational codes are likely to also regulate in-person exchanges, WhatsApp is comparatively a refuge of privacy within the online world itself, one of the few opportunities to escape from context collapse without simultaneously escaping from digital communications.

In this light, WhatsApp does not *create* echo chambers or polarize by itself, but sediments a superb space for individuals to unleash their partisan nature; it is the cocktail that results from the combination of the platform's unique affordances and our tribal strand that may culminate in some sort of self-segregation and

increasing levels of radicalization or outgroup aversion. Hence, empirical works categorically debunking the fragmentation and polarizing potential of social media would do well on not dismissing too early the warnings of political theorists; different platforms advance different outcomes.

WHAT DO WE (NOT) KNOW ABOUT WHATSAPP? FURTHERING A BURGEONING RESEARCH PATH

There are hence several reasons to advance a research agenda centered on studying private social media like WhatsApp:

(1) existing findings predominantly relate to two algorithm-led platforms;
(2) they generally refer to a dimension of the social media-fragmentation/polarization link dissimilar from that graspable through non-algorithm-led platforms, that is, the extent to which platforms *create* echo chambers or *make* people more angry/polarized vs the extent to which they serve as spaces for individuals to segregate and/or radicalize themselves;
(3) the concrete characteristics of WhatsApp deviate from those of traditional SNSs and bode for the configuration of hermetic, non-diverse spaces prone to further humans' tribalist drives when it comes to politics; and
(4) private social media are so widespread that they enjoy higher usage ratios than many SNSs (Newman et al., 2022), with WhatsApp being the most popular platform in Europe and most of the Global South.

The *WhatsApp literature* burgeoned very few years ago and has grown exponentially since then. However, it is still at a nascent point (see also Baulch et al., 2024a). Logistically, though there has been a modest surge in scientific studies addressing the possibilities for political talk on WhatsApp and its effects, these are still scant. An unsurprising fact if considering that there was barely a dozen works on the matter at the dawn of the new decade. As **Table 1** showed, WhatsApp research tackling fragmentation or polarization issues comprises less than 20 studies in the Scopus database. If the scope is broadened to also consider the relationship between this platform and other issues like misinformation or political participation, the number of related investigations is considerably higher but still non-comparable to the thousands of in-depth analyses approaching Twitter or Facebook (see **Table 1**). Importantly, not all those works find in WhatsApp their central focus. In fact, most of them mobilize this app as part of a cross-platform endeavor to consider broader questions associated with social media usage (Brenes Peralta et al., 2022; Goyanes et al., 2025; Klein, 2025; Pinheiro et al., 2020; Rossini et al., 2021b; Valenzuela et al., 2024). Likewise, not all those works correspond to peer-reviewed pieces published in journals; given the academic blossoming of this platform, some of them incarnate grey literature materials and congress presentations (Bhutani et al., 2019; Rosenfeld et al., 2018; Visani et al., 2022).

Substantively, WhatsApp's literature unevenly tackles only some of the conundrums linked with the use of this platform while presenting notable challenges to generalize its results beyond specific contexts (see Baulch et al., 2024a). To be more precise, we distinguish three groups of "WhatsApp studies". First, those which portray WhatsApp as a central tool for receiving, sharing and commenting on political information, broadly able to market candidates' images and shape elections in a multifaceted way (e.g., Boczkowski et al., 2018; Bozdag, 2020; Caetano et al., 2018; Cheeseman et al., 2020; Evangelista & Bruno, 2019; Garimella & Tyson, 2018; Moura & Michelson, 2017; Pont-Sorribes et al., 2020; Rossini et al., 2021a; Smith, 2025; Udupa & Wasserman, 2025; Zhu et al., 2022). Relevantly among these, Masip et al. (2021) identify WhatsApp's unique technological affordances and investigate how they shape users' engagement with news; Chadwick et al. (2023a, 2023b) conceive WhatsApp communication as "hybrid public-interpersonal," since individuals continuously switch between interpersonal and semi-public contexts, and any political discussion ends up encapsulated within quotidian, informal exchanges with close ties; similarly, Swart et al. (2018) define WhatsApp as a "dark social media," where "public connection" is possible only within a context of privacy and sender control over the audiences; Zhu et al. (2022) scrutinize how political talk can emerge from WhatsApp's everyday social interactions; and, in this vein, Kligler-Vilenchik (2021) points to the platform's close-tie sociability as a double-edged sword that simultaneously deters political talk and fosters mutual understanding when it occurs. As a matter of fact, all these investigations draw from qualitative approaches and single case studies, which may jeopardize the universalization of the findings.

Second, there are studies that concentrate on the spread of misinformation through the platform, be it by reflecting on the reception and circulation of WhatsApp fake news (Banaji & Bhat, 2020; Basavaraj, 2022; Elías & Catalan-Matamoros, 2020), by examining the structured dissemination of misinformative propaganda during political campaigns (Bursztyn & Birnbaum, 2019; Soares et al., 2021), by studying the singularities of *non-text* disinformation afforded by the platform (Cardoso et al., 2022; Garimella & Eckles, 2020; Sundar et al., 2021), by exploring users' motivations to misinform (Goyanes et al., 2025; Herrero-Diz et al., 2020) as well as their engagement in accuracy determination tactics (Brenes Peralta et al., 2022; Chadwick et al., 2023b; Kanthawala & Maddox, 2022; Kligler-Vilenchik, 2022; Malhotra, 2024), by developing machine learning, sentiment analysis or semantic search approaches for the detection of fake news (Alonso et al., 2021; Bhutani et al., 2019; Visani et al., 2022), or by evaluating the technical opportunities for countering misinformation in the platform (Kuru et al., 2022; Reis et al., 2020). There is also a non-negligible number of interuniversity research projects that analyze these dynamics from a wide variety of angles (e.g., Campbell et al., 2019; Chadwick et al., 2021; Sundar & Vasanti, 2019; Tandoc et al., 2019), though disproportionately focusing on the Global South, in general, and Brazil and India, in particular.

For the sake of accuracy, the number of studies on these matters is greater and grows exponentially almost on a weekly basis. Besides the above, for instance, recent research has emphasized the role of WhatsApp group rulemaking as a means to deal with misinformation (Chadwick et al., 2023a) and numerical misinformation more specifically (Lawson et al., 2025), and it has explored the distinctive WhatsApp uses and environments that fuel misinformation sharing and belief (Goyanes et al., 2025; Kalogeropoulos & Rossini, 2025a, 2025b; Rossini et al., 2023; Valenzuela et al., 2024). To sum up, this panorama suffices to show that, in the field of political communication, WhatsApp's literature is unbalanced in favor of the misinformation phenomenon and at the expense of equally relevant outcomes, such as fragmentation or polarization, our focus in this essay. This is detrimental because, as we suggested, there are good reasons to speculate that WhatsApp's unique features could lead to higher polarization.

Indeed, the state of scientific evidence is much more fragile when it comes to the third group of WhatsApp studies, namely, those which examine the effects of WhatsApp (political) usage on certain outcomes, including polarization. Consuming news and discussing politics through private messaging apps like WhatsApp have been suggested to undercut individuals' political knowledge (Gil de Zúñiga & Goyanes, 2021; Yamamoto et al., 2018) and media literacy (Wang et al., 2022). Moreover, those behaviors seem to strengthen users' activism (Gil de Zúñiga et al., 2021), their extra-institutional political participation in offline and online arenas (Ojo et al., 2025; Scherman et al., 2022; Vaccari & Valeriani, 2018; Yamamoto et al., 2018), and their proclivity to protest illegally (Gil de Zúñiga & Goyanes, 2021; Klein, 2025). Meng & Wang (2024), on their part, find no influence of those WhatsApp usages on generalized trust, while Valenzuela et al. (2021) and Vermeer et al. (2021) do capture a significant positive effect of them on learning about politics.

More relevantly, Valeriani and Vaccari (2018) find a slightly positive link between discussing politics on WhatsApp and ideological extremism, while, building upon the Chilean context, Scherman et al. (2022) show that WhatsApp usage increases perceived political polarization, but not affective polarization. Drawing from (panel) survey methods, the latter conclusion is also reached by Lee et al. (2021) for Japan and the US, and by Valenzuela et al. (2021) for Chile. Similarly, in their India-based survey experiment, Neyazi et al. (2025) observe no discernible effects of WhatsApp uncivil messages on affective polarization.[2] Some research (Combs et al., 2023; Kibet & Ward, 2018; Kligler-Vilenchik et al., 2020; Yarchi et al., 2021; Zumárraga-Espinosa, 2024) even suggests a depolarizing impact of private messaging on the political content produced by users. However, this set of studies often draws from artificially created WhatsApp groups, which involve heterogeneous networks of ideologically diverse and highly educated citizens who engage in civil conversations. These groups hence bring together conducive conditions for that depolarization to occur, but this does not seem reflective of the ordinary conditions under which individuals express politically in the platform on a daily basis. Lastly, building on a Hong Kong-based survey, Davidson and Kobayashi (2023) do show that WhatsApp political rumoring increases outgroup aversion and thereby affective polarization, and, in a relatively similar spirit, Kothur and Pandey (2025) note that WhatsApp indeed affects polarization by strengthening the related effect of overall

[2] According to Rossini et al. (2025), what this type of WhatsApp messages do predict is avoidance behavior, with users being prone to disengage, and uncivil discussions proving themselves "underminers" of close social bonds.

social media news consumption. And this is all: concerning the relationship between WhatsApp and polarization, we just count with this dozen of investigations with mixed results and stemming from not always generalizable contexts.[3] On the one hand, they conform an extraordinarily narrow corpus of evidence (as compared to Twitter/Facebook), which prevents reaching stark conclusions on the matter; on the other, even the few existing works face serious weaknesses.

Thus, there exists a thriving research agenda associated with the study of WhatsApp as a political communication tool; however, this boom has not been paralleled with extensive analyses on the actual fragmentation of WhatsApp political talk or the polarizing implications of using this platform. These research paths are still comparatively minor and present key limitations. We thus suggest that there is a need to fill these voids as a way to strengthen this emerging literature. On the one hand, to assess political talk fragmentation, a first endeavor revolves around understanding how exactly political talk takes place on WhatsApp. The works by Chadwick et al. (2023a, 2023b), Kligler-Vilenchik (2021), Masip et al. (2021), Swart et al. (2018), and Zhu et al. (2022) embody meaningful pioneer approaches to this. Nevertheless, they account for several shortfalls. First, they all excel in conceptualizing WhatsApp's communicative nature and in defining the singular overall framework within which the platform affords (political) discussions, which comprises a combination of specific technological affordances, individual human factors, and group social dynamics in the context of *quotidianity* and *interpersonality*. However, with the partial exceptions of Kligler-Vilenchik (2021) and Masip et al. (2021), they fail to examine how such a unique framework conditions certain dimensions of political talk compared to other online and offline spaces (including opinion expression, modulation, or formation), or whether it tends to encourage a prototypical type of political conversation at the cost of others. The latter, secondly, would require a fine-grained reflection on the various types of conversations that the platform potentially affords, as it is unlikely that public affairs are always addressed in the same way and with the same interlocutors (see Vaccari & Valeriani, 2021), especially on a non-curated platform technically unable to homogenize experiences. Some works (Caetano et al., 2018; Garimella & Tyson, 2018) have sought to characterize WhatsApp's political discussion considering the influence of the platform's singularity. Nevertheless, given the encrypted nature of WhatsApp, these approaches rely on public groups, so that future efforts should revolve around the small private groups in which most people participate (Rossini et al., 2021a).

A third point of interest hitherto overlooked refers to the extent to which the various ways of discussing public affairs on WhatsApp constitute relevant phenomena at all, as well as whether (and how) different population segments harness differently WhatsApp's uniqueness (i.e., who holds which political conversation on the platform). Studies such as Masip et al.'s (2021), Swart et al.'s (2018) or Chadwick et al.'s (2023a, 2023b) reveal that some people do comment on public issues on WhatsApp, but, beyond stereotypical portrayals, we know neither the actual share of the population who engage in (each of) these conversations, nor the specific subjects more prone to do so. Likewise, we ignore whether the latter diverge from the "traditional suspects" involved in other social media or offline fora, which turns critical to design tailored policy remedies to tackle potential democratic threats fueled by WhatsApp. Kalogeropoulos's (2021) research is a partial attempt to fill this void; it seeks to delineate who spreads political information on the platform, but its focus is just on unidirectional news sharing. When it comes specifically to the discussion of public affairs, only the investigations by Rossini et al. (2021a) –for Brazil– and Valeriani and Vaccari (2018) –for Germany, Italy, and the UK– incarnate an exception to the general lack of literature on these aspects, as they do give insights into platform conversational use across the population. Even in this case, however, the investigations neither profile WhatsApp "political users" nor discern between different types of political discussions. Overall, one of the reasons underlying this shortcoming relates to the utter predominance of qualitative approaches, which complicates the precise measurement of the actual prevalence of conversations and prototypical users. Further research should therefore stress the importance of quantitative strategies in this regard.

These proposals for a better understanding of WhatsApp political talk turn critical for the ulterior assessment of the degree of fragmentation of this space. So far, we lack evidence documenting the variety of spheres within the platform where different conversations take place, as well as their level of personal or

[3] This shortage of research focusing on the relationship of WhatsApp with polarization is particularly striking in light of the in-parallel rocketing number of studies that present this platform as a repository of ingredients which indeed drive polarization, such as hate speech (Udupa & Wasserman, 2025) and its normalization via humor (Bayarri Toscano & Fernández-Villanueva, 2025) or memes (Chagas, 2024).

ideological homogeneity. In other public fora, like Twitter, curation algorithms expose different users to different content, but individuals *play* all in the same field, with posts being all visible in the platform feed. This does not occur in WhatsApp, where playing fields are potentially infinite, and individuals could make strategic decisions when discussing politics depending on the types of WhatsApp groups they participate in and the people they engage with, as Kligler-Vilenchik's (2021) observational study suggests. Thus, to evaluate whether political talk is *echo-chamberized* on WhatsApp, it is necessary to pay attention to the idiosyncrasy of the groups where that talk takes place. Accordingly, investigations should go for the systematized characterization of those playing fields –contacts and groups– that accommodate specific types of conversations: is ideological homophily a precondition to engage with politics on WhatsApp groups, or can they be politically diverse, and when? (When) do WhatsApp groups require some "personal homophily" based on close ties to discuss politics?

Qualitative research (Chadwick et al., 2023a, 2023b; Masip et al., 2021; Udupa & Wasserman, 2025) has suggested that political discussions are segmented on WhatsApp, making the echo chamber metaphor applicable to this context. However, no quantitative endeavor has been conducted to this date that supports and amplifies the generalizability of such claims.[4] When it comes to political conversations, a look into WhatsApp groups' physiognomy that considers their size, ideological affinity and tie strength would document more rigorously fragmentation fears. To that end, a transition seems convenient towards more mixed or quantitative approaches that enable us to corroborate the hypotheses of qualitative exploratory works and reach more robust conclusions. With the same objective, those would ideally draw from cross-cultural data, as another key limitation of existing qualitative studies resides in their sole reliance on participants and experiences from one country. Hence, in light of the difficulties for penetrating this platform, we suggest advancing the path initiated by Kalogeropoulos and Rossini (2025a, 2025b), who employ survey-based techniques to grasp the diverging compositional features of WhatsApp groups, though they neither establish a relevant typology of groups, nor consider whether different group types are conducive for different types of political conversations.

Ultimately, we perceive that overall considerations of fragmentation in WhatsApp need to consider a variety of elements: (1) the diversity of political conversations that could potentially arise in the platform; (2) the nature of such conversations (do their dynamics resemble or deviate from echo chambers/constructive exchanges/toxic relief?); (3) the specific groups in which people self-select for each conversation (not all conversations may occur in equally segmented environments), and the characteristics of their interlocutors (close ties? Ideologically like-minded?); and (4) the actual prevalence of fragmented experiences amidst the broader conversational context of WhatsApp.

On the other hand, we encourage scholars to further explore the link between WhatsApp usage and the various dimensions of polarization.[5] There exist strong theoretical reasons to expect increasing polarization from certain WhatsApp uses; however, only very few works have addressed this matter, and they have done so problematically, which makes extant evidence, besides mixed, objectionable. For example, as anticipated, Kligler-Vilenchik et al. (2020) and Yarchi et al. (2021) obtain their results from artificial and ideologically diverse groups created *ad hoc* for their investigations and which do not resemble the quotidian dynamics of the platform. Other survey analyses, such as Lee et al.'s (2021) or Scherman et al.'s (2022), do not face this problem but disregard the complexity of everyday WhatsApp use. By measuring how polarization is affected by the "frequency of using WhatsApp (to discuss politics)," these studies assume that WhatsApp's technical possibility to choose with whom to connect automatically leads to opinion-reinforcing and polarizing interactions. However, there is not a sole way of using WhatsApp for politics, and it seems naïve to believe that there is a

[4] Indeed, in their systematic review of echo chamber research within social media, Hartmann et al. (2025) already point to the lack of studies that document empirically the primacy of echo-chamber spaces in private messaging apps like WhatsApp.

[5] Ideological and affective dimensions of polarization might result from different processes and have different implications (Reiljan, 2020). However, this distinction is sometimes overlooked, and no study addresses them together. We barely count with works concentrating on the ideological dimension –Valeriani and Vaccari (2018) could be an exception–, while those dwelling on the affective side (Lee et al., 2021; Neyazi et al., 2025) do not abound either. Comprehensive approaches that conjointly assess and connect the distinct link between WhatsApp and each of those dimensions should be especially welcomed by the academic community.

univocal effect associated with all the possible ways of doing so. In attempting to provide a more realistic picture, we propose to define the concrete political experiences afforded by the platform. This includes both passive –receiving political content– and active experiences –sharing political content or maintaining different types of political conversations (e.g., echo chamber discussions to ridicule/mess with the political adversary, "Habermasian talks" to better understand other viewpoints)–. Rather than assuming that polarization stems directly from WhatsApp's conducive affordances, we suggest estimating the relationship between distinct conversations and experiences –which may serve as useful independent variables– and polarization.

In this vein, it is also naïve to believe that even specific WhatsApp usages possess uniform effects across different subgroups within the population. This precision is critical to grasp whether some WhatsApp usages can widen or reduce extant inequalities in polarization or outgroup aversion levels. Nonetheless, with the exception of Davidson and Kobayashi (2023), the literature has usually overlooked this detail. In this line, coming research should not only determine which population segments are prone to radicalize or develop hateful outgroup feelings because of their WhatsApp experiences, but also size the moderation effect of the platform environments where each experience occurs. In other words: (how) do the compositional features of WhatsApp groups influence the link between diverse WhatsApp experiences and polarization? The convenience to advance in this latter direction was also stressed by Kalogeropoulos and Rossini (2025a, 2025b). They already created some indicators to measure groups' composition, although their goal consisted of grasping the moderator role of group physiognomies in misinformation sharing on WhatsApp.

Lastly, it seems equally naïve to think that WhatsApp experiences affect polarization uniformly across contexts and nations. Research has hitherto focused on the Global South via single-country case studies (e.g., Brazil, India, Southeast-Asian countries). However, our knowledge of these dynamics for Western and European democracies is paradoxically almost non-existent. This raises questions about the generalizability of existing results. We contend that future studies should make a cross-country effort and pay greater attention to the role that systemic features play in shaping the relationship between WhatsApp experiences and polarization. Given the impossibility to scrape primary data across private groups, all the above endeavors are likely to demand survey-based strategies. Yet, we argue that WhatsApp provides a unique opportunity to embrace mixed methods approaches, combining survey-rooted quantitative data with interview- or focus-group-centered qualitative inputs that allow for a better understanding of the observed effects and their explanatory mechanisms. Interestingly, despite a few quantitative and qualitative works already exist, no study has made both approaches dialogue.

CONCLUDING REMARKS

This paper has argued that the current focus on public social media and their algorithm-driven content curation falls short of capturing the full spectrum of influences that digital platforms exert on our public sphere. Private social media like WhatsApp are also important spaces for engaging with politics, deviate from traditional SNSs and arguably set the conditions for more insular and echo chamber-enhancing interactions. However, they are significantly under-researched, limiting our understanding of the role of various forms of digital communication in shaping political fragmentation and polarization.

Expanding this research agenda is, on the one hand, a useful academic exercise to build a clearer picture of our interaction with private social media like WhatsApp. For example, one of the few existing pieces of evidence related to WhatsApp usage is that it boosts political participation (Valenzuela et al., 2021; Yamamoto et al., 2018). If using this platform were also to augment polarization, this could explain why these apps promote political action, i.e., by leveraging partisan commitments instead of other normatively desirable mechanisms. On the other hand, deepening this research agenda is a convenient endeavor to design effective strategies against the challenges posed by the *platformization* of the public sphere. If private social media were indeed to enhance tendencies towards tribalism and segregation, policy interventions aimed at countering this problem should not be solely based on dominant Twitter-centric regulatory approaches. By embracing a more precise research agenda, we can untangle the complex web of digital influences on our public sphere.

Author contributions: JMM: conceptualization, supervision, project administration, formal analysis, writing – original draft, writing – review & editing; **CRM**: supervision, writing – review & editing. Both authors approved the final version of the article.

Funding: This study was supported by the EPUDEM project "The erosion of the democratic public sphere: Causes, consequences, and reform proposals" (project-ID: PP2022_5) funded by Universidad Pontificia Comillas.

Acknowledgments: The authors would like to thank participants in the "Politics and Emotions" panel of the 17th Congress of Spain's Political Science Association for their valuable feedback on earlier versions of this paper. The authors would also like to thank the anonymous reviewers for their careful reading of our manuscript and their insightful comments and suggestions.

Ethics declaration: This study is a literature review of previously published research. As it does not involve human participants, animals, sensitive personal data, or any experimental interventions, formal approval from an ethics committee was not required. All sources used in this review were publicly available and properly cited.

Declaration of interest: The authors declared no competing interest.

Data availability: Data generated or analyzed during this study are available from the authors on request.

REFERENCES

Achen, C. H., & Bartels, L. M. (2017). *Democracy for realists*. Princeton University Press. https://doi.org/10.2307/j.ctvc7770q

Alonso, M. A., Vilares, D., Gómez-Rodríguez, C., & Vilares, J. (2021). Sentiment analysis for fake news detection. *Electronics, 10*(11), 1348-1368. https://doi.org/10.3390/electronics10111348

Bail, C. (2022). *Breaking the social media prism: How to make our platforms less polarising*. Princeton University Press. https://doi.org/10.2307/j.ctv2mgg2nm

Banaji, S., & Bhat, R. (2020). *WhatsApp vigilantes: An exploration of citizen reception and circulation of WhatsApp misinformation linked to mob violence in India*. London School of Economics and Political Science.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science, 26*(10), 1531-1542. https://doi.org/10.1177/0956797615594620

Basavaraj, K. (2022). Misinformation in India's 2019 national election. *Journal of Quantitative Description Digital Media, 2*, 1-43. https://doi.org/10.51685/jqd.2022.021

Baulch, E., Johns, A., & Matamoros-Fernández, A. (2024a). A critical review of media and communications scholarship on messaging apps. In M. M. Skoric, & N. Pang (Eds.), *Research handbook on social media and society* (pp. 270-286). Edward Elgar Publishing. https://doi.org/10.4337/9781800377059.00031

Baulch, E., Matamoros-Fernández, A., & Suwana, F. (2024b). Memetic persuasion and WhatsAppification in Indonesia's 2019 presidential election. *New Media & Society, 26*(5), 2473-2491. https://doi.org/10.1177/14614448221088274

Bayarri Toscano, G., & Fernández-Villanueva, C. (2025). "Funny weapons": The norms of humour in the construction of far-right political polarisation. *Social Inclusion, 13*, 1-19. https://doi.org/10.17645/si.10211

Bayer, J. B., Ellison, N. B., Schoenebeck, S. Y., & Falk, E. B. (2016). Sharing the small moments: Ephemeral social interaction on Snapchat. *Information, Communication & Society, 19*(7), 956-977. https://doi.org/10.1080/1369118X.2015.1084349

Beam, M. A., Hutchens, M. J., & Hmielowski, J. D. (2018). Facebook news and (de)polarization: Reinforcing spirals in the 2016 US election. *Information, Communication & Society, 21*(7), 940-958. https://doi.org/10.1080/1369118X.2018.1444783

Bhutani, B., Rastogi, N., Sehgal, P., & Purwar, A. (2019). Fake news detection using sentiment analysis. In *Proceedings of the 2019 12th International Conference on Contemporary Computing* (pp. 1-5). IEEE. https://doi.org/10.1109/IC3.2019.8844880

Boczkowski, P. J., Matassi, M., & Mitchelstein, E. (2018). How young users deal with multiple platforms: The role of meaning-making in social media repertoires. *Journal of Computer-Mediated Communication, 23*(5), 245-259. https://doi.org/10.1093/jcmc/zmy012

Bossetta, M. (2018). The digital architectures of social media: Comparing political campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 US election. *Journalism & Mass Communication Quarterly, 95*(2), 471-496. https://doi.org/10.1177/1077699018763307

Bossetta, M. (2023). Social media digital architectures: A platform-first approach to political communication and participation. In S. Coleman, & L. Sorensen (Eds.), *Handbook of digital politics* (pp. 226-270). Edward Elgar Publishing. https://doi.org/10.4337/9781800377585.00025

Boulianne, S., Hoffmann, C. P., & Bossetta, M. (2024). Social media platforms for politics: A comparison of Facebook, Instagram, Twitter, YouTube, Reddit, Snapchat, and WhatsApp. *New Media & Society, 27*(11), 6006-6037. https://doi.org/10.1177/14614448241262415

Bozdag, C. (2020). Managing diverse online networks in the context of polarization: Understanding how we grow apart on and through social media. *Social Media + Society, 6*(4). https://doi.org/10.1177/2056305120975713

Brenes Peralta, C. M., Sánchez, R. P., & González, I. S. (2022). Individual evaluation vs fact-checking in the recognition and willingness to share fake news about COVID-19 via WhatsApp. *Journalism Studies, 23*(1), 1-24. https://doi.org/10.1080/1461670X.2021.1994446

Brown, M. A., Bisbee, J., Lai, A., Bonneau, R., Nagler, J., & Tucker, J. A. (2022). *Echo chambers, rabbit holes, and algorithmic bias: How YouTube recommends content to real users*. SSRN. https://doi.org/10.2139/ssrn.4114905

Bruns, A. (2022). Echo chambers? Filter bubbles? The misleading metaphors that obscure the real problem. In M. Pérez-Escolar, & J. M. Noguera-Vivo (Eds.), *Hate speech and polarization in participatory society* (pp. 33-48). Routledge. https://doi.org/10.4324/9781003109891-4

Bursztyn, V. S., & Birnbaum, L. (2019). Thousands of small, constant rallies: A large-scale analysis of partisan WhatsApp groups. In *Proceedings of the 2019 IEEE/ ACM International Conference on Advances in Social Networks Analysis and Mining* (pp. 484-488). IEEE. https://doi.org/10.1145/3341161.3342905

Caetano, J. A., de Oliveira, J. F., Lima, H. S., Marques-Neto, H. T., Magno, G., Meira Jr, W., & Almeida, V. A. (2018). *Analysing and characterizing political discussions in WhatsApp public groups*. arXiv. https://doi.org/10.48550/arXiv.1804.00397

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. Willey.

Campbell, S., Kuru, O., Bayer, J., Baruh, L., & Ling, R. (2019). *Beyond the forward: The social shaping of (mis)information through WhatsApp* [Research project, University of Michigan].

Cardoso, G., Sepúlveda, R., & Narciso, I. (2022). WhatsApp and audio misinformation during the COVID-19 pandemic. *Profesional de la Información, 31*(3), 1-14. https://doi.org/10.3145/epi.2022.may.21

Cea, N., & Palomo, B. (2021). Disinformation matters: Analyzing the academic production. In G. López-García, D. Palau-Sampio, B. Palomo, E. Campos-Domínguez, & P. Masip (Eds.), *Politics of disinformation: The influence of fake news on the public sphere* (p. 5-22). John Wiley & Sons. https://doi.org/10.1002/9781119743347.ch1

Chadwick, A., Hall, N. A., & Vaccari, C. (2023a). Misinformation rules!? Could "group rules" reduce misinformation in online personal messaging? *New Media & Society, 27*(1), 106-126. https://doi.org/10.1177/14614448231172964

Chadwick, A., Vaccari, C., & Hall, N. A. (2023b). What explains the spread of misinformation in online personal messaging networks? Exploring the role of conflict avoidance. *Digital Journalism, 12*(5), 574-593. https://doi.org/10.1080/21670811.2023.2206038

Chadwick, A., Vaccari, C., Hall, N. A., Akolgo, P., & Lawson, B. (2021). *The everyday misinformation project* [Research project, Loughborough University].

Chagas, V. (2024). Far-right memespheres and platform affordances: The effects of environmental opacity on the spread of extremist memes on Twitter and WhatsApp. *Journal of Applied Communication Research, 52*(4), 517-533. https://doi.org/10.1080/00909882.2024.2355114

Cheeseman, N., Fisher, J., Hassan, I., & Hitchen, J. (2020). Nigeria's WhatsApp politics. *Journal of Democracy, 31*, Article 145. https://doi.org/10.1353/jod.2020.0037

Child, J. T., & Starcher, S. C. (2016). Fuzzy Facebook privacy boundaries: Exploring mediated lurking, vague-booking, and Facebook privacy management. *Computers in Human Behavior, 54*, 483-490. https://doi.org/10.1016/j.chb.2015.08.035

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 118*(9), 1-8. https://doi.org/10.1073/pnas.2023301118

Combs, A., Tierney, G., Guay, B., Merhout, F., Bail, C. A., Hillygus, D. S., & Volfovsky, A. (2023). Reducing political polarization in the United States with a mobile chat platform. *Nature Human Behaviour, 7*(9), 1454-1461. https://doi.org/10.1038/s41562-023-01655-0

Dahlgreen, P. (2005). The Internet, public spheres, and political communication: Dispersion and deliberation. *Political Communication, 22*(2), 147-162. https://doi.org/10.1080/10584600590933160

Davidson, B., & Kobayashi, T. (2023). Did you hear? Rumour communication via instant messaging apps and its impact on affective polarisation. In C. Soon (Ed.), *Mobile communication and online falsehoods in Asia: Trends, impact and practice* (pp. 173-196). Springer. https://doi.org/10.1007/978-94-024-2225-2_10

De Blasio, E., Marianne, K., Wolf, S., & Sorice, M. (2020). The ongoing transformation of the digital public sphere: Basic considerations on a moving target. *Media and Communication, 8*(4), 1-5. https://doi.org/10.17645/mac.v8i4.3639

Elías, C., & Catalan-Matamoros, D. (2020). Coronavirus in Spain: Fear of 'official' fake news boosts WhatsApp and alternative sources. *Media and Communication, 8*(2), 462-466. https://doi.org/10.17645/mac.v8i2.3217

Evangelista, R., & Bruno, F. (2019). WhatsApp and political instability in Brazil: targeted messages and political radicalization. *Internet Policy Review, 8*(4), 1-23. https://doi.org/10.14763/2019.4.1434

Fletcher, R., Robertson, C. T., & Nielsen, R. K. (2021). How many people live in politically partisan online news echo chambers in different countries? *Journal of Quantitative Description: Digital Media, 1*. https://doi.org/10.51685/jqd.2021.020

Garimella, K., & Eckles, D. (2020). Images and misinformation in political groups: Evidence from WhatsApp in India. *The Harvard Kennedy School Misinformation Review, 1*(5). https://doi.org/10.37016/mr-2020-030

Garimella, K., & Tyson, G. (2018). WhatsApp doc? A first look at WhatsApp public group data. In *Proceedings of the International AAAI Conference on Web and Social Media*. https://doi.org/10.1609/icwsm.v12i1.14989

Gil de Zúñiga, H., & Goyanes, M. (2021). Fuelling civil disobedience in democracy: WhatsApp news use, political knowledge, and illegal political protest. *New Media & Society, 5*(12), 3500-3520. https://doi.org/10.1177/14614448211047850

Gil de Zúñiga, H., Ardèvol-Abreu, A., & Casero-Ripollés, A. (2021). WhatsApp political discussion, conventional participation and activism: Exploring direct, indirect and generational effects. *Information, Communication & Society, 24*(2), 201-218. https://doi.org/10.1080/1369118X.2019.1642933

Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press. https://doi.org/10.12987/9780300235029

Goyanes, M., Jordá, B., & de Fuentes, J. M. (2025). From lack of concerns to complicity: How misinformation concerns influences fake news sharing on WhatsApp via the partisan legitimation of fake news. *Social Science Quarterly, 106*(4), Article e70039. https://doi.org/10.1111/ssqu.70039

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology, 78*(6), 1360-1380. https://doi.org/10.1086/225469

Habermas, J. (1962). *The structural transformation of the public sphere*. Polity.

Habermas, J. (2006). Political communication in media society: Does democracy still enjoy an epistemic dimension? The impact of normative theory on empirical research. *Communication Theory, 16*(4), 411-426. https://doi.org/10.1111/j.1468-2885.2006.00280.x

Habermas, J. (2022). Reflections and hypotheses on a further structural transformation of the political public sphere. *Theory, Culture & Society, 39*(4), 145-171. https://doi.org/10.1177/02632764221112341

Haim, M., Graefe, A., & Brosius, H. B. (2018). Burst of the filter bubble? Effects of personalization on the diversity of Google News. *Digital Journalism, 6*(3), 330-343. https://doi.org/10.1080/21670811.2017.1338145

Hartmann, D., Wang, S. M., Pohlmann, L., & Berendt, B. (2025). A systematic review of echo chamber research: comparative analysis of conceptualizations, operationalizations, and varying outcomes. *Journal of Computational Social Science, 8*(2), Article 52. https://doi.org/10.1007/s42001-025-00381-z

Herrero-Diz, P., Conde-Jiménez, J., & Reyes de Cózar, S. (2020). Teens' motivations to spread fake news on WhatsApp. *Social Media + Society*, *6*(3). https://doi.org/10.1177/2056305120942879

Kalogeropoulos, A. (2021). Who shares news on mobile messaging applications, why and in what ways? A cross-national analysis. *Mobile Media & Communication, 9*(2), 336-352. https://doi.org/10.1177/2050157920958442

Kalogeropoulos, A., & Rossini, P. (2025a). Don't talk to strangers? The role of network composition, WhatsApp groups, and partisanship in explaining beliefs in misinformation about COVID-19 in Brazil. *Journal of Information Technology & Politics, 22*(1), 113-130. https://doi.org/10.1080/19331681.2023.2234902

Kalogeropoulos, A., & Rossini, P. (2025b). Unraveling WhatsApp group dynamics to understand the threat of misinformation in messaging apps. *New Media & Society, 27*(3), 1625-1650. https://doi.org/10.1177/14614448231199247

Kanthawala, S., & Maddox, J. (2022). Hiding in the echo chamber: Fact-Checking failures and individual tactics of accuracy determination on WhatsApp in India. *Asian Journal of Communication, 32*(2), 174-191. https://doi.org/10.1080/01292986.2021.2023594

Kibet, A., & Ward, S. (2018). Socially networked heterogeneity: The influence of WhatsApp as a social networking site on polarisation in Kenya. *African Journalism Studies, 39*(4), 42-66. https://doi.org/10.1080/23743670.2018.1537979

Klein, O. (2025). Mobilising the mob: The multifaceted role of social media in the January 6th US capitol attack. *Javnost-The Public, 32*(1), 33-50. https://doi.org/10.1080/13183222.2025.2469028

Kligler-Vilenchik, N. (2021). Friendship and politics don't mix? The role of sociability for online political talk. *Information, Communication & Society, 24*(1), 118-133. https://doi.org/10.1080/1369118X.2019.1635185

Kligler-Vilenchik, N. (2022). Collective social correction: Addressing misinformation through group practices of information verification on WhatsApp. *Digital Journalism, 10*(2), 300-318. https://doi.org/10.1080/21670811.2021.1972020

Kligler-Vilenchik, N., Baden, C., & Yarchi, M. (2020). Interpretative polarization across platforms: How political disagreement develops over time on Facebook, Twitter, and WhatsApp. *Social Media + Society, 6*(3). https://doi.org/10.1177/2056305120944393

Kothur, L., & Pandey, V. (2025). Role of social media news consumption in cultivating opinion polarization. *Information Technology & People, 38*(2), 757-786. https://doi.org/10.1108/ITP-12-2022-0942

Kubin, E., & von Sikorski, C. (2021). The role of (social) media in political polarization: A systematic review. *Annals of the International Communication Association, 45*(3), 188-206. https://doi.org/10.1080/23808985.2021.1976070

Kuru, O., Campbell, S., Bayer, J., Baruh, L., & Ling, R. (2022). Encountering and correcting misinformation on WhatsApp: The roles of user motivations and trust in messaging group members. *Disinformation in the Global South, 6*, 88-107. https://doi.org/10.1002/9781119714491.ch7

Lawson, B. T., Chadwick, A., Hall, N. A., & Vaccari, C. (2025). The trustworthiness of peers and public discourse: Exploring how people navigate numerical dis/misinformation on personal messaging platforms. *Information, Communication & Society, 28*(4), 633-650. https://doi.org/10.1080/1369118X.2024.2400141

Lee, S., Rojas, H., & Yamamoto, M. (2021). Social media, messaging apps, and affective polarization in the United States and Japan. *Mass Communication and Society, 25*(5), 673-697. https://doi.org/10.1080/15205436.2021.1953534

Lorenz-Spreen, P., Oswald, L., Lewandowsky, S., & Hertwig, R. (2023). A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature Human Behaviour, 7*(1), 74-101. https://doi.org/10.1038/s41562-022-01460-1

Lottridge, D., & Bentley, F. R. (2018). Let's hate together: How people share news in messaging, social, and public networks. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-13). https://doi.org/10.1145/3173574.3173634

Malhotra, P. (2024). Misinformation in WhatsApp family groups: Generational perceptions and correction considerations in a meso-news space. *Digital Journalism, 12*(5), 594-612. https://doi.org/10.1080/21670811.2023.2213731

Masip, P., Suau, J., Ruiz-Caballero, C., Capilla, P., & Zilles, K. (2021). News engagement on closed platforms. Human factors and technological affordances influencing exposure to news on WhatsApp. *Digital Journalism, 9*(8), 1062-1084. https://doi.org/10.1080/21670811.2021.1927778

McCay-Peet, L., & Quan-Haase, A. (2017). What is social media and what questions can social media research help us answer. In L. Sloan, & A. Quan-Haase (Eds.), *The SAGE handbook of social media research methods* (pp. 35-51). SAGE. https://doi.org/10.4135/9781473983847.n2

McKay, S., & Tenove, C. (2021). Disinformation as a threat to deliberative democracy. *Political Research Quarterly, 74*(3), 703-717. https://doi.org/10.1177/1065912920938143

Meng, X., & Wang, Y. (2024). To trust or not to trust? Exploring the roles of Facebook and WhatsApp use and network diversity. *Mass Communication and Society, 27*(6), 1366-1388. https://doi.org/10.1080/15205436.2023.2248956

Moura, M., & Michelson, M. R. (2017). WhatsApp in Brazil: Mobilising voters through door-to-door and personal messages. *Internet Policy Review, 6*(4), 1-18. https://doi.org/10.14763/2017.4.775

Newman, N., Fletcher, R., Robertson, C., Eddy, K., & Nielsen, R. K. (2022). *Reuters Institute digital news report 2022*. University of Oxford.

Neyazi, T. A., Kuru, O., & Mukerjee, S. (2025). The Influence of sexism and incivility in WhatsApp political discussions on affective polarization: Evidence from a 2022 multi-party election in India. *The International Journal of Press/Politics, 30*(3), 797-818. https://doi.org/10.1177/19401612241242431

Ojo, O. A., Edwards, B. I., & Oyewole, J. A. (2025). Educational technology and digital democracy: A case study on social media's influence on undergraduate political engagement in Nigeria. *Environment and Social Psychology, 10*(7), Article 3612. https://doi.org/10.59429/esp.v10i7.3612

Persily, N., & Tucker, J. A. (Eds.). (2020). *Social media and democracy: The state of the field, prospects for reform*. Cambridge University Press. https://doi.org/10.1017/9781108890960

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology, 90*(5), Article 751. https://doi.org/10.1037/0022-3514.90.5.751

Pinheiro, D. A., De Araújo, M. G. L., De Souza, K. B., Campos, B. D. S., De Oliveira, E. M., Lima, R. S. M., & De Souza, G. B. (2020). Sharing fake news about health in the cross-platform messaging app WhatsApp during the COVID-19 pandemic: A pilot study. *International Journal of Scientific Research and Management, 8*, 403-410. https://doi.org/10.18535/ijsrm/v8i10.mp01

Pont-Sorribes, C., Besalú, R., & Codina, L. (2020). WhatsApp como canal de información política en España: Credibilidad, perfil de usuarios y compartición de contenidos [WhatsApp as a channel for political information in Spain: Credibility, user profile and content sharing]. *Profesional de la Información, 29*(6). https://doi.org/10.3145/epi.2020.nov.19

Reiljan, A. (2020). 'Fear and loathing across party lines' (also) in Europe: Affective polarisation in European party systems. *European Journal of Political Research, 59*(2), 376-396. https://doi.org/10.1111/1475-6765.12351

Reis, J., Melo, P. D. F., Garimella, K., & Benevenuto, F. (2020). Can WhatsApp benefit from debunked fact-checked stories to reduce misinformation? *The Harvard Kennedy School Misinformation Review, 1*(5). https://doi.org/10.37016/mr-2020-035

Rosa, H. (2022). Social media filters and resonances: Democracy and the contemporary public sphere. *Theory, Culture & Society, 39*(4), 17-35. https://doi.org/10.1177/02632764221103520

Rosenfeld, A., Sina, S., Sarne, D., Avidov, O., & Kraus, S. (2018). WhatsApp usage patterns and prediction of demographic characteristics without access to message content. *Demographic Research, 39*, 647-670. https://doi.org/10.4054/DemRes.2018.39.22

Rossini, P., Baptista, É. A., de Oliveira, V. V., & Stromer-Galley, J. (2021a). Digital media landscape in Brazil: Political (mis)information and participation on Facebook and WhatsApp. *Journal of Quantitative Description: Digital Media, 1*. https://doi.org/10.51685/jqd.2021.015

Rossini, P., Mont'Alverne, C., & Kalogeropoulos, A. (2023). Explaining beliefs in electoral misinformation in the 2022 Brazilian election: The role of ideology, political trust, social media, and messaging apps. *The Harvard Kennedy School Misinformation Review, 4*(3). https://doi.org/10.37016/mr-2020-15

Rossini, P., Stromer-Galley, J., Baptista, E. A., & Veiga de Oliveira, V. (2021b). Dysfunctional information sharing on WhatsApp and Facebook: The role of political talk, cross-cutting exposure and social corrections. *New Media & Society, 23*(8), 2430-2451. https://doi.org/10.1177/1461444820928059

Rossini, P., Stromer-Galley, J., Baptista, E. A., & Veiga de Oliveira, V. (2025). Discussions that divide: The interplay between experiencing uncivil conversations and politically motivated avoidance on WhatsApp and Facebook. *Journal of Broadcasting & Electronic Media, 69*(4), 437-453. https://doi.org/10.1080/08838151.2025.2519733

Rowlett, J., & Harlow, S. (2018). Selfies and sensationalism on the campaign trail: A visual analysis of snapchat's political coverage. *Visual Communication Quarterly, 25*(2), 82-92. https://doi.org/10.1080/15551393.2017.1388724

Ruess, C., Hoffmann, C. P., & Boulianne, S. (2023) Online political participation: The evolution of a concept. *Information, Communication & Society, 26*(8), 1495-1512. https://doi.org/10.1080/1369118X.2021.2013919

Santini, R. M., Tucci, G., Salles, D., & de Almeida, A. R. D. (2021). Do you believe in fake after all? WhatsApp disinformation campaign during the Brazilian 2018 presidential election. In G. López-García, D. Palau-Sampio, B. Palomo, E. Campos-Domínguez, & P. Masip (Eds.), *Politics of disinformation: The influence of fake news on the public sphere* (pp. 49-66). John Wiley & Sons. https://doi.org/10.1002/9781119743347.ch4

Scherman, A., Etchegaray, N., Browne, M., Mazorra, D., & Rojas, H. (2022). WhatsApp, polarization, and non-conventional political participation: Chile and Colombia before the social outbursts of 2019. *Media and Communication, 10*(4), 77-93. https://doi.org/10.17645/mac.v10i4.5817

Seeliger, M., & Sevignani, S. (2022). A new structural transformation of the public sphere? An introduction. *Theory, Culture & Society, 39*(4), 3-16. https://doi.org/10.1177/02632764221109439

Smith, M. M. (Ed.). (2025). *Political marketing on social and mobile media: Remaking the paradigm?* Taylor & Francis. https://doi.org/10.4324/9781003533078

Smyrnaios, N. (2015). Google and the algorithmic infomediation of news. *Media Fields Journal, 10*(10).

Smyrnaios, N., & Baisnée, O. (2023). Critically understanding the platformization of the public sphere. *European Journal of Communication, 38*(5), 435-445. https://doi.org/10.1177/02673231231189046

Soares, F. B., Recuero, R., Volcan, T., Fagundes, G., & Sodré, G. (2021). Bolsonaro's firehose: How COVID-19 disinformation on WhatsApp was used to fight a government political crisis in Brazil. *The Harvard Kennedy School Misinformation Review,* 2(1).

Stewart, E., & Hartmann, D. (2020). The new structural transformation of the public sphere. *Sociological Theory, 38*(2), 170-191. https://doi.org/10.1177/0735275120926205

Sundar, S. S., & Vasanti, P. N. (2019). *Seeing is believing: Is video modality more powerful in spreading fake news?* [Research project, The Pennsylvania State University].

Sundar, S. S., Molina, M. D., & Cho, E. (2021). Seeing is believing: Is video modality more powerful in spreading fake news via online messaging apps? *Journal of Computer-Mediated Communication, 26*(6), 301-319. https://doi.org/10.1093/jcmc/zmab010

Sunstein, C. R. (2018). *#Republic: Divided democracy in the age of social media*. Princeton University Press. https://doi.org/10.1515/9781400890521

Swart, J., Peters, C., & Broersma, M. (2018). Shedding light on the dark social: The connective role of news and journalism in social media communities. *New Media & Society, 20*(11), 4329-4345. https://doi.org/10.1177/1461444818772063

Tandoc, E., Sian, L. C., & Joanna, S. S. C. (2019). *Is correction fluid? How to make fact-checks on WhatsApp more effective* [Research project, Nanyang Technological University].

Terren, L., & Borge-Bravo, R. (2021). Echo chambers on social media: A systematic review of the literature. *Review of Communication Research, 9*, 99-118. https://doi.org/10.12840/ISSN.2255-4165.028

Torcal, M., Martini, S., & Orriols, L. (2018). Deciding about the unknown: The effect of party and ideological cues on forming opinions about the European Union. *European Union Politics, 19*(3), 502-523. https://doi.org/10.1177/1465116518769754

Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., & Nyhan, B. (2018). *Social media, political polarization, and political disinformation: A review of the scientific literature*. Hewlett Foundation. https://doi.org/10.2139/ssrn.3144139

Udupa, S., & Wasserman, H. (Eds.) (2025). *WhatsApp in the world: Disinformation, encryption, and extreme speech*. NYU Press. https://doi.org/10.18574/nyu/9781479833306.001.0001

Vaccari, C., & Valeriani, A. (2018). Digital political talk and political participation: Comparing established and third wave democracies. *SAGE Open, 8*(2). https://doi.org/10.1177/2158244018784986

Vaccari, C., & Valeriani, A. (2021). *Outside the bubble: Social media and political participation in Western democracies*. Oxford University Press. https://doi.org/10.1093/oso/9780190858476.001.0001

Valenzuela, S., Bachmann, I., & Bargsted, M. (2021). The personal is the political? What do WhatsApp users share and how it matters for news knowledge, polarization and participation in Chile. *Digital Journalism, 9*(2), 155-175. https://doi.org/10.1080/21670811.2019.1693904

Valenzuela, S., Muñiz, C., & Santos, M. (2024). Social media and belief in misinformation in Mexico: A case of maximal panic, minimal effects? *The International Journal of Press/Politics, 29*(3), 667-688. https://doi.org/10.1177/19401612221088988

Valeriani, A., & Vaccari, C. (2018). Political talk on mobile instant messaging services: A comparative analysis of Germany, Italy, and the UK. *Information, Communication & Society, 21*(11), 1715-1731. https://doi.org/10.1080/1369118X.2017.1350730

van Dyk, S. (2022). Post-truth, the future of democracy and the public sphere. *Theory, Culture & Society, 39*(4), 37-50. https://doi.org/10.1177/02632764221103514

Vermeer, S. A., Kruikemeier, S., Trilling, D., & de Vreese, C. H. (2021). WhatsApp with politics?! Examining the effects of interpersonal political discussion in instant messaging apps. *The International Journal of Press/Politics, 26*(2), 410-437. https://doi.org/10.1177/1940161220925020

Visani, C., Sorathiya, V., & Lavadiya, S. (2022). Fake news polarization for sentiment analysis. In S. Malik, R. Bansal, & A. Tyagi (Eds.), *Impact and role of digital technologies in adolescent lives* (pp. 255-263). IGI Global. https://doi.org/10.4018/978-1-7998-8318-0.ch019

Waller, I., & Anderson, A. (2021). Quantifying social organization and political polarization in online platforms. *Nature, 600*(7888), 264-268. https://doi.org/10.1038/s41586-021-04167-x

Wang, A. H. E., Yeh, Y. Y., Wu, C. K., & Chen, F. Y. (2022). Media literacy and partisan convergence across social network sites. *The Social Science Journal*. https://doi.org/10.1080/03623319.2022.2127174

We Are Social & Meltwater. (2024). Digital 2024 April global statshot report. *Meltwater-We Are Social*. https://datareportal.com/reports/digital-2024-april-global-statshot

Yamamoto, M., Kushin, M. J., & Dalisay, F. (2018). How informed are messaging app users about politics? A linkage of messaging app use and political knowledge and participation. *Telematics and Informatics, 35*(8), 2376-2386. https://doi.org/10.1016/j.tele.2018.10.008

Yarchi, M., Baden, C., & Kligler-Vilenchik, N. (2021). Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media. *Political Communication, 38*(2), 98-139. https://doi.org/10.1080/10584609.2020.1785067

Zhu, Q., Esteve-Del-Valle, M., & Meyer, J. K. (2022). Safe spaces? Grounding political talk in WhatsApp groups. *New Media & Society, 26*(9), 5423-5444. https://doi.org/10.1177/14614448221136080

Zumárraga-Espinosa, M. (2024). El rol despolarizador de la heterogeneidad de red y la discusión política diversa en medios sociales: Un análisis de las plataformas Facebook, Twitter y WhatsApp en las elecciones de Ecuador 2021 [The depolarizing role of network heterogeneity and diverse political discussion in social media: An analysis of the Facebook, Twitter, and WhatsApp platforms in the 2021 Ecuador elections]. *Redes: Revista Hispana para el Análisis de Redes Sociales, 35*(2), 165-183. https://doi.org/10.5565/rev/redes.1024