African Journal of Social and Behavioural Sciences (AJSBS)
Volume 15, Number 9 (2025) ISSN: 2141-209X

# FAKE NEWS DISSEMINATION ON WHATSAPP: A SYSTEMATIC REVIEW OF EVIDENCE (2018–2024)

**Sikiru Olayemi Abdullahi[1]*, Temitope Seun Ibiyeye[2] & Bashir Amoda Ajijola[3]**

[1,2,3]Department of Mass Communication, Faculty of Humanities and Social Science, Al-Hikmah University, Ilorin, Kwara State. Nigeria

*olayemi595@gmail.com

**ABSTRACT:** The proliferation of fake news through encrypted messaging platforms poses growing challenges to digital communication governance. This paper systematically reviews scholarly evidence on how technological, behavioural, and policy factors influence fake news dissemination on WhatsApp between 2018 and 2024. Following PRISMA guidelines, 28 peer-reviewed studies were identified through Google Scholar, Scopus, and PubMed using search strings such as "WhatsApp misinformation", "fake news dissemination", "credibility", "speed", and "information accuracy". Inclusion criteria comprised English-language empirical studies published between 2018 and 2024. Data were synthesised thematically across four analytical dimensions: credibility, speed, accessibility, and accuracy. Results suggest that WhatsApp's end-to-end encryption, forwarding features, and group dynamics enable misinformation to circulate rapidly within trusted peer networks. Psychological factors—particularly confirmation bias, emotional triggers, and social trust—further accelerate virality. Platform interventions such as forwarding limits and message labelling mitigate spread only marginally, as users often bypass restrictions. Evidence suggests that the problem is more pronounced in regions with low media literacy, where WhatsApp serves as a primary information source. Policymakers, platform designers, and educators must collaborate to develop culturally responsive, privacy-preserving interventions. Strengthening digital literacy and integrating metadata-based detection mechanisms that can curb virality without compromising encryption. This study consolidates six years of global research, offering a comprehensive synthesis of the mechanisms, behavioural drivers, and policy gaps underpinning misinformation diffusion on WhatsApp.

**Keywords:** Fake news dissemination, WhatsApp, End-to-end encryption, Digital communication, Media literacy, Misinformation management, Peer-to-peer networks.

## INTRODUCTION

### Background to the Study

The high rate of development of digital technologies has changed the way information is created, exchanged, and consumed in the world. Although these innovations have made the world more connected and easier to communicate with, they have also promoted the fastest growth of fake news, meaning false or misleading information that is intentionally framed as factual (Wardle & Derakhshan, 2018). The spread of misinformation across the globe has become a source of concern over its implications on the stability of political systems, on the well-being of the people and on the

trust of society. To illustrate, in the case of the COVID-19 pandemic, misinformation about health spread through online channels disrupted the trust of people and caused the inability to respond medically promptly (Pennycook et al., 2020). This fact highlights the conflict between technology and human behaviour in the age when information dissemination usually exceeds its verification.

The social media and messaging tools make this problem even more difficult by exploiting algorithms that tend to push content that is more engaging, often by targeting emotional appeal, like fear or outrage. It has been discovered that emotionally charged fake news is more quickly and widely circulated than factual news since users like to share content that corresponds to their beliefs and feelings (Vosoughi et al., 2018). This has led to the corrosion of the line separating truthful information and intentional misrepresentation as a result of digital ecosystems that put virality above accuracy.

WhatsApp plays an especially powerful role among such platforms in the context of misinformation. WhatsApp serves as a mass information channel and a personal communication device (it already has more than 2 billion active users as of 2024) (Statista, 2024). Its end-to-end encryption secures the privacy of the users and at the same time covers the trail of fake information to prevent the management of the misinformation by external parties (Cole & Wagner, 2024). Since they are likely to believe what their friends, family, and community networks have forwarded to them, the app creates a closed-loop system in which unconfirmed material is propagated unhindered.

The design of WhatsApp allows sharing information widely and quickly due to the ability to share information through group chats, broadcast lists, forwarding messages, and sharing multimedia, which cannot be edited by an editor (Jain & Meena, 2024). Despite the introduction of items like forwarding limits and message labelling in lieu of reducing virality, users are still able to bypass these measures by either copying information manually or forwarding it to smaller groups (Huang, 2024). The dynamics expose a larger dilemma of maintaining the platform's accessibility and privacy and, to the least possible extent, reducing the misuse of the platform by spreading false information.

Empirical research also demonstrates that a large proportion of users do not often check whether the content they received was actually authentic, especially in media-illiterate societies (Okocha & Akpe, 2024). Digital misinformation has real-life consequences, and in India, false information distributed through WhatsApp has led to violence (Monteiro, 2024). These examples illustrate how the convergence of technology design, user psychology, and social trust is the root cause of fake news spread.

The functionality of WhatsApp, such as encryption, the ability to interact in groups, forwarding, multimedia support, and the use of temporary updates of status, complicates communication but, at the same time, makes it harder to stop the spread of misinformation. This duality is critical in the development of effective countermeasures that can keep the privacy of users intact and guarantee the reliability of information. The review in question explores the collective role of technological, behavioural, and policy influences on the dissemination of fake news on WhatsApp between 2018 and 2024. It examines how misinformation can be spread on the platform, how the technological

design of WhatsApp, the behaviour of users, and social dynamics interact to perpetrate misinformation. The paper aims to single out the key aspects and peculiarities of fake news that is shared on WhatsApp, study the psychological and social factors that encourage users to get involved in the process of sharing and sending it, and discuss the ways in which the platform architecture fosters or restricts the speed at which misinformation is spread. Moreover, it assesses the efficiency of the current interventions that are technological, regulatory, and educational aimed at limiting the dissemination of fake news. By conducting this analysis, the research will enhance the current knowledge on the role of the WhatsApp communication ecosystem in the overall problem of misinformation in the digital era.

The paper relies on interdisciplinary insights into communication theory, technology studies and behavioural psychology to make sense of the fake news spreading through closed digital networks. In theory, it helps to continue the discussions concerning the nature of interactions between technological affordances and human cognitive biases to create digital information ecosystems. In practice, it provides information on how to construct more information-sensitive policies, media literacy programmes, and platform design solutions that would help alleviate misinformation and protect user freedoms and privacy.

Although scholarly focus on fake news on open social systems like Facebook and Twitter is expanding, it is established that there is little research examining the dynamics of fake news in encrypted and private networks such as WhatsApp. This bit of information stifles a complete grasp of the flow of misinformation that occurs outside the perceptions of people and how interventions can be effectively applied in these situations.

## LITERATURE REVIEW

The review is a synthesis of the results of twenty-eight peer-reviewed articles published in 2018-2024 that were identified with the use of a structured search process across databases, such as Scopus, Google Scholar, and ResearchGate. The keywords that the search strings included were 'WhatsApp misinformation', 'fake news dissemination', 'digital communication', and 'information verification'. The inclusion criteria were that the studies had to be a case study of WhatsApp, discuss the dynamics of misinformation, or involve behavioural and policy interventions in digital communication.

### Platform Affordances and the Fake News on WhatsApp.

The characteristics of the WhatsApp design are crucial to manipulating misinformation spread. In contrast to popular social networks like Facebook or Twitter, WhatsApp is a closed and encrypted system, and the content is validated by the users to a great extent (Pennycook et al., 2020). Researchers concur that end-to-end encryption, which was launched in 2016, favours privacy but unintentionally hampers the need to combat or monitor misinformation (Cole & Wagner, 2024). According to the research by Monteiro (2024) and Jain and Meena (2024), such an encrypted space leads to high levels of interpersonal trust that increases the credibility of forwarded messages in small, close-knit groups.

Although previous research (Vosoughi et al., 2018) focused on algorithmic virality of open platforms, recent findings (Herrero-Diz & Conde-Jimenez, 2020; Okocha & Akpe, 2024) speak of algorithmic virality of WhatsApp, the responsibility to verify is fully on the user. This informational opaqueness renders fake news hard to follow or disprove, forming what academics refer to as a dark social system ecosystem in which disinformation propagates freely.

Recurring content categories are identified in empirical research, and they include political propaganda, misinformation about health, and sensational rumours. The use of politically motivated fake news during the 2018 elections in Brazil had an impact on the people due to the dissemination of such information through WhatsApp groups (Masip et al., 2021). Otherwise, false health information about cures and vaccines spread throughout the COVID-19 pandemic (Pennycook et al., 2020; Apuke & Omar, 2021). These instances demonstrate that the WhatsApp affordances, such as privacy, multimedia integration, and quick forwarding, are made in such a way to build a viral, not verbal, ecosystem.

**User Behaviour and Cognitive Biases.**

The second theme of discussion is the role of the human mind and interpersonal interaction in supporting the spread of fake news. According to scholars, confirmation bias has always been the primary force, as people accept and forward the information that agrees with the already formed picture (Mahamed et al., 2023). This bias is further fuelled by the trust that peer networks invoke, where one will usually tend to believe messages by relatives or close friends irrespective of their origin (Asghar et al., 2021).

Various research studies (Banaji et al., 2019; Munusamy et al., 2024) tie the illusion-of-truth effect and fear of missing out (FOMO) to impulsive distribution of false information. In the case of crisis, messages that are framed as urgent or time-sensitive are perceived as having a higher level of emotion, and this will lead to faster spreading of the message without validation. This is especially acute in the fake news about health and security when the emotional appeal takes over the critical thinking.

WhatsApp group dynamics are also important. Such individuals as administrators and socially influential people can be called super spreaders of misinformation (Kalogeropoulos & Rossini, 2023). Echo chambers are homogeneous groups of people who have similar ideological or cultural beliefs and help to strengthen misinformation (Santini et al., 2021). Conflicting results appear in terms of demographics, as older users are more vulnerable to it because of low levels of digital literacy (Dauda & Abubakar, 2020; Morosoli et al., 2022), whereas younger users are affected by emotional and identity-based information (Shafi & Ravikumar, 2018).

It is also behaviour tendencies that are influenced by the cultural settings. Collectivists such as India and Nigeria find it harder to question the information they receive in messages because they want to protect social harmony (Herrero-Diz & Conde-Jimenez, 2020; Shahzad et al., 2023). Conversely, members of individualist cultures have a greater level of doubt and are less hesitant to question questionable information (Munusamy et al., 2024). In such a way, the behaviour aspect of misinformation spreading is not only cognitive but highly social and cultural as well.

**Interventions of Policies and Regulation.**

Some of the countermeasures taken by WhatsApp and regulating organisations have been in response to the global problem of fake news. In 2018, WhatsApp implemented a feature of message forwarding restrictions and marking of frequently forwarded items (Paris & Pasquetto, 2024). Research findings indicate a moderate effect, whereby fewer cases of mass forwarding occur when there is a significant misinformation spurt like the COVID-19 pandemic (Masip et al., 2021). Nevertheless, the critics believe that these actions are a shallow attempt because users may still copy and paste the content to evade restrictions (Pereira, 2022).

Authorities around the globe have advocated increased regulation, but they usually have been in conflict with WhatsApp and its user privacy policy and encryption (Omar, 2023). As an example, the proposal of India to trace forwarded messages in 2021 caused controversy due to the possible violation of freedom of expression (Syamsi, 2023). Researchers such as Santini et al. (2021) and Helm and Nasu (2021) suggest using a multi-stakeholder framework to balance privacy rights and accountability by cooperating with civil society, policymakers, and technology companies.

There is unanimity in the literature that top-down regulatory measures do not suffice. Rather, more sustainable solutions can be found through improving digital literacy and spreading the idea of fact-checking efforts organised by communities. Nevertheless, implementation is incomplete, particularly in the areas that have little access to valid information sources (Okocha & Akpe, 2024).

**Technological Countermeasures and Future Directions.**

The last theme is centred on technological changes meant to curb fake news on encrypted systems. Other scholars like Jain and Meena (2024) suggest metadata-related systems of detection that are capable of detecting suspicious patterns of sharing without breaking encryption. Likewise, AI-assisted methods examining the frequency of dissemination and the similarity of content have demonstrated the potential to encounter viral fake news in the initial phase (Harris et al., 2024).

However, the literature is still unsure about the ethical and practical possibility of such tools. Their advantage is that they guarantee better monitoring, but their drawback is that they will compromise the privacy of users, which is one of the principles on which WhatsApp was designed. To address this dilemma, recent papers (Herrero-Diz & Conde-Jimenez, 2020; Medeiros & Singh, 2020) recommend combining privacy-law-abiding AI and participatory verification frameworks, in which users voluntarily review questionable content.

The literature on technological countermeasures highlights the shift to proactive measures rather than reactive ones, where human and automated countermeasures should be used together. The merger of behavioural understanding, technological development, and ethical policy systems is becoming more widely regarded as a crucial tool towards the struggle against fake news within closed digital ecosystems.

### Synthesis and Gaps in the Literature

Through the studies reviewed, a unifying theme is evident: human behaviour, technological design and policy frameworks all provide the basic components of misinformation spreading on WhatsApp. Nevertheless, there are still great lapses. Not many studies use the longitudinal or cross-regional approach, and most of them pay much attention to the specific context of one country, including India or Brazil. Further, the intervention strategies are not empirically tested on encrypted environments.

This review will help seal these gaps by combining technological and behavioural as well as policy perspectives to offer a holistic view of misinformation dynamics on WhatsApp. It is based on previous descriptive research that provides an analytical summary of the intersection of privacy, trust, and virality to continue the circulation of fake news between 2018 and 2024.

## METHODOLOGY

The paper under consideration demonstrates a systematic review design that is facilitated by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 framework. The intention was to generalise previous empirical studies on how fake news has been spread through WhatsApp between 2018 and 2024 and make clear, reproducible findings.

The search was organised in Google Scholar, Scopus, and PubMed. Search strings were a combination of keywords with Boolean operators of the following form:

("WhatsApp misinformation" OR "fake news distribution" OR "WhatsApp rumours") AND (credibility) OR (speed) OR (accuracy) OR (accessibility).

The search was done in publications between January 2018 and December 2024. Other manual screenings of reference lists of key papers were used to supplement results.

### Inclusion Criteria

- Articles, conference papers, and reports concerning WhatsApp misinformation peer-reviewed.
- Articles in the English language published within 2018-2024.
- Study that is methodologically well described (qualitative, quantitative, or mixed).

### Exclusion Criteria

- Opinion pieces, editorials or an unverified grey literature.
- Surveys that examine platforms other than WhatsApp.
- Servers with duplicates or unavailable full texts.

**Screening and Selection**

A total of 124 publications were found in the search, and 48 publications were screened by keywords (abstract and title). Following the use of inclusion criteria, 28 articles were to be analysed in their entirety.

**A PRISMA flow**

- Records identified: 124
- After screening: 48
- Omitted due to irrelevance or redundancy: 20
- Final studies included: 28

**Quality Appraisal**

The quality of methods was measured by the Mixed Methods Appraisal Tool (MMAT). All the studies were rated based on the clarity of the objectives, suitability of the design, sufficiency of sampling, and clarity of analysis. Only those studies that scored 60 per cent or above were put in the synthesis.

**Data Synthesis and Extraction.**

The thematic synthesis approach was used to systematise findings based on four dimensions of analysis that had been identified as relevant in previous studies:

- **Credibility:** relevance of sources, certifiability of content;
- **Speed:** rate of spread of messages;
- **Accessibility:** ease in getting messages over networks;
- **Accuracy:** efficiency of action and judgement of the user.

The analyzed researches were summarized as a results matrix (Table 1) and interpreted using comparative analysis, which shows both convergence, divergence, and gaps in the research.

**RESULTS**

A total of 28 studies were included in the study and were examined under varying geographical, methodological, and thematic settings. The majority of publications were based in Asia (primarily India and Malaysia), Latin America (Brazil), and Africa (Nigeria), where the WhatsApp usage and rates of misinformation are the most widespread.

**Table 1: Summary of Reviewed Studies (2018–2024)**

| Author(s) | Year | Country/Region | Methodology | Key Findings |
|---|---|---|---|---|
| Vosoughi, Roy & Aral | 2018 | Global | Quantitative (Network Analysis) | False news spreads six times faster than factual news, primarily due to emotional novelty. |
| Farooq | 2018 | India | Content Analysis | WhatsApp serves as a political propaganda tool influencing public perception. |
| Banaji et al. | 2019 | India | Qualitative (Case Study) | Group-based trust encourages unverified message sharing during communal crises. |
| Caetano et al. | 2019 | Brazil | Computational Study | High forwarding chains show network clustering typical of closed systems. |
| Medeiros & Singh | 2020 | India | Policy Review | Forwarding limits reduce mass dissemination but do not stop message replication. |
| Pennycook & Rand | 2020 | Global | Experimental | "Accuracy prompts" improve discernment but have temporary behavioural effects. |
| Herrero-Diz & Conde-Jiménez | 2020 | Spain | Survey | Young users share misinformation for social validation more than belief. |
| Dauda & Abubakar | 2020 | Nigeria | Survey | Lower digital literacy correlates with higher fake news forwarding. |
| Lee & Lee | 2021 | Korea | Experimental | Repetition increases perceived truthfulness of misinformation (illusion of truth). |
| Helm & Nasu | 2021 | Global | Legal Review | Privacy–security paradox complicates misinformation regulation. |
| Masip et al. | 2021 | Europe | Digital Ethnography | Closed WhatsApp environments hinder fact-checking interventions. |
| Kalogeropoulos & Rossini | 2023 | Europe | Content & Network Analysis | Echo chambers dominate group conversations; |

| | | | | |
|---|---|---|---|---|
| | | | | authority of administrators amplifies falsehoods. |
| Shahzad, Iqbal & Khan | 2023 | Pakistan | Systematic Review | Cultural identity and language strongly shape fake news diffusion. |
| Syamsi | 2023 | Indonesia | Policy Analysis | Traceability laws threaten encryption and user trust. |
| Paris & Pasquetto | 2024 | Global | Qualitative | Message labelling improves awareness but not verification. |
| Monteiro | 2024 | Brazil | Case Study | WhatsApp misinformation shaped political discourse in 2018 and 2022 elections. |
| Munusamy et al. | 2024 | Malaysia | Systematic Review | Anxiety and emotional arousal drive misinformation during crises. |
| Harris et al. | 2024 | Global | Technical Review | Metadata analysis can detect patterns without compromising privacy. |
| Morosoli et al. | 2022 | Europe | Survey | Older adults are more vulnerable due to high trust and low digital skills. |
| Vese | 2022 | Global | Conceptual Review | Encryption ensures privacy but obstructs governance and moderation. |
| Okocha & Akpe | 2024 | Nigeria | Qualitative | WhatsApp misinformation reduces civic trust and public participation. |
| Omar | 2023 | Malaysia | Policy Study | Regulatory efforts depend on national digital literacy and infrastructure. |
| Pennycook et al. | 2020 | Global | Experimental | Behavioural "accuracy nudges" reduce misinformation sharing in controlled settings. |
| Kalogeropoulos & Rossini | 2023 | Global | Mixed Methods | Group administrators and influencers shape group message acceptance. |
| Santini, Tucci & Salles | 2021 | Brazil | Case Study | Fake news campaigns during elections polarize citizens. |
| Banaji et al. | 2019 | India | Qualitative | Group echo chambers reinforce confirmation bias. |

| | | | | |
|---|---|---|---|---|
| Monteiro | 2024 | Brazil | Case Study | Misinformation networks evolve faster than countermeasures. |
| Masip et al. | 2021 | Spain | Digital Journalism | Users rely more on peer validation than institutional verification. |

**Thematic Synthesis by Analytical Dimension**

**A. Credibility**

In a majority of research, perceived credibility is socially constructed as opposed to being evidence-based. Listening to messages shared by influential people or authority figures (including group administrators or community leaders) is forced upon the unconscious (Kalogeropoulos & Rossini, 2023). Credibility perception is also improved by emotional and ideological alignment (Apuke & Omar, 2021). It is a social trust mechanism that does not allow external fact-checking to be as effective as users, in the first place, are interpersonally validated rather than institutionally.

**B. Speed**

Fake news is greatly accelerated by the forwarding feature and multimedia content (videos, pictures, voice notes). When users are in a highly anxiety-inducing environment, like during a health crisis or election, they tend to circulate content without any verification (Munusamy et al., 2024). The mass sharing was reduced by up to 70 percent in certain areas, but new redistribution types in manual form appeared (Medeiros & Singh, 2020).

**C. Accessibility**

The simple nature of WhatsApp and its efficiency in data consumption have rendered it a hegemony of communication, especially in developing nations where people have limited access to broadband (Dauda & Abubakar, 2020). Multilingual flexibility enables it to customise messages depending on local dialects and socio-cultural frameworks and increase its relevance and shareability (Shahzad et al., 2023). Subsequently, WhatsApp can usually serve as both a source of information and a source of misinformation in low-literacy conditions.

**D. Accuracy**

The addition of label features for frequently forwarded messages, as well as message limits, made users more aware but did not impact their sharing behaviors (Paris & Pasquetto, 2024). Simple accuracy nudges, which require users to rate the truthfulness of a message, as has been experimentally demonstrated, provide short-term benefits in discernment (Pennycook & Rand, 2020). However, this result fails to persist with broader literacy support.

**Discussion**

The results of the review support the fact that the spread of fake news on WhatsApp is facilitated by the convergence of technological possibilities, behavioural patterns and regulatory limitations. These three lenses are used to understand the results in this section.

The end-to-end encryption of WhatsApp is one of the clear conflicts between the privacy enhancement of the main users and the regulation of the information. The closed-network architecture that guarantees confidentiality, on the other hand, impedes moderation and traceability (Vese, 2022). The data confirms the argument that privacy technologies are unwillingly helping to promote misinformation societies, which is reminiscent of the discussion of the digital rights dilemma found by Helm and Nasu (2021).

The group and broadcast form includes the characteristics of public social media but without visibility or transparency of algorithms. This invisibility of virality enables misinformation to continue operating below the regulatory radar and requires creativity in approaches to restricting personal privacy (Harris et al., 2024).

The results of behaviour are very much consistent with psychological models of cognitive bias and social contagion. Confirmation bias is the reason why people are more inclined to share information that fits their worldview, and emotional contagion spreads fear-inducing information faster (Apuke & Omar, 2021; Munusamy et al., 2024). Acceptance of misinformation is supported by the illusion-of-truth effect (Lee & Lee, 2021) when repeated in various groups.

These biases are also enhanced by echo chambers, which make group administrators and the more prolific posters significant gatekeepers. As demonstrated by Kalogeropoulos and Rossini (2023), misinformation usually gets incorporated into social identity, so when one attempts to provide a corrective message, this may be seen as an assault instead of a correction.

Platform-level interventions that relay a restriction, message labelling, and awareness information have been moderately successful. Nevertheless, the misinformation is still persistent, implying that technical controls cannot be enough. The users react to the policy changes depending on the cultural, linguistic, and emotional context (Shahzad et al., 2023).

Regulations imposed by the government requiring traceability pose a threat of losing user trust and violating civil liberties (Syamsi, 2023). Therefore, co-regulation, i.e., collaboration mechanisms between governments, technology companies, and civil society organisations, is the best option (Helm & Nasu, 2021).

The problem of freedom of expression and the necessity to suppress misinformation that is harmful still remains a matter of ethical issues. Excessive surveillance may cause censorship, and the lack of control allows the harm to continue. The findings demonstrate that an alternative governance framework with privacy preservation is necessary that incorporates community-induced reporting, algorithmic identification of metadata trends, and localised digital literacy programmes.

## Conclusion

In this systematic review, 28 peer-reviewed articles that were published between 2018 and 24 were analysed to learn how technological, behavioural, and policy factors influence the dissemination of fake news through WhatsApp. The results indicate that the design, especially end-to-end encryption, group chats, and forwarding, of the platform serves to create a personalised but extremely viral system of communication where misinformation thrives on trust, emotion, and repetition. The review states three interdependent forces that keep this issue alive, including technological affordances which facilitate the replication of messages quickly, behavioural bias which promotes the sharing of messages without thinking, and policy constraints that do not combat adaptive user behaviour. Even though other measures like message labelling and forwarding restrictions have kept the mass forwarding under control, circumvention and unequal digital literacy still undermine their effectiveness, a fact that proves the claim made that misinformation on WhatsApp is a multi-layered socio-technical problem that requires multi-layered solutions.

## Recommendations

Based on the evidence reviewed, it is possible to suggest the following recommendations:

1. WhatsApp ought to consider partnering with independent fact-checking agencies to create in-app verification, which will notify users when the shared content is similar to the known misinformation trends.
2. In line with Dauda and Abubakar (2020) and Banaji et al. (2019), the focus of the interventions should be on low-digital literacy populations, in particular, older users and rural communities, where the culturally-specific education programs teaching skills of verification and critical evaluation should be provided.
3. Metadatabased AI systems would prevent breaches of encryption to reveal abnormal dissemination chains, and in this regard, there would be a balance between privacy and accountability, which is in line with Harris et al. (2024).
4. Governments, civil society, and platform operators must collaborate with one another through frameworks that provide unilateral regulation in the form of censorship and encourage transparency (Helm & Nasu, 2021).
5. Evidence-based interventions by policy-makers to examine the actual impacts of WhatsApp anti-misinformation protections should be funded by longitudinal and experimental research, making the process proactive instead of reactive.

## Implications to Practice and Research

1. Communication services that make use of encryption should establish responsible innovation streams that incorporate ethics and governance in its early design. The review singles out the necessity of adaptive design thinking that provides a balance between usability, privacy and verifiability.
2. Digital inclusion and media literacy should be the major priority of regulatory bodies as a long-term solution rather than a punitive measure. The focus is to be on empowering the user to make informed choices rather than focusing on the surveillance of the message.

3. Universities, schools, and non-profits can also be instrumental by incorporating critical media literacy into curricula and mass education. Emotional intelligence and critical thinking can be built to counter the vulnerability to emotionally charged misinformation.
4. Future studies need to undertake cross-country comparative study designs to investigate cross-cultural differences in misinformation dynamics. Furthermore, experimental research, which tests behavioral interventions such as gamified accuracy prompts or group-administered accountability systems, may expand the knowledge of the interaction between design and psychology in closed networks.

**REFERENCES**

Apuke, O. D., & Omar, B. (2021). User motivation in fake news sharing during the COVID-19 pandemic: An application of the uses and gratification theory. *Online Information Review, 45*(6), 1201–1216. https://doi.org/10.1108/OIR-03-2020-0116

Banaji, S., Bhat, R., Agarwal, A., Passanha, N., & Pravin, M. S. (2019). *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.

Caetano, J. A., Magno, G., & Gonçalves, M. (2019). Characterizing attention cascades in WhatsApp groups. *Proceedings of the 10th ACM Conference on Web Science*, 45–54. https://doi.org/10.1145/3292522.3326035

Cole, K. K., & Wagner, J. M. (2024). Crisis pregnancy centers: Digital rhetoric, misinformation, and trust. *Journal of Digital Communication Research, 6*(2), 88–102.

Dauda, S., & Abubakar, A. (2020). Students' behavioural intention towards adopting WhatsApp as a source of news. *Arts and Social Science Research, 8*(2), 59–72.

Farooq, G. (2018). Politics of fake news: How WhatsApp became a potent propaganda tool in India. *Media Watch, 9*(1), 48–64. https://doi.org/10.15655/mw/2018/v9i1/49279

Harris, S., Hadi, H. J., Ahmad, N., & Alshara, M. A. (2024). Fake news detection revisited: An extensive review of theoretical frameworks, dataset assessments, and forward-looking research. *Technologies, 12*(11), 222. https://doi.org/10.3390/technologies12110222

Helm, R. K., & Nasu, H. (2021). Regulatory responses to "fake news" and freedom of expression: Normative and empirical evaluation. *Human Rights Law Review, 21*(2), 302–329. https://doi.org/10.1093/hrlr/ngab010

Herrero-Diz, P., & Conde-Jiménez, J. (2020). Teens' motivations to spread fake news on WhatsApp. *Social Media + Society, 6*(4), 1–15. https://doi.org/10.1177/2056305120978471

Huang, Y. (2024). Fake news reaching young people on social networks: Distrust challenging media. *Journal of Media Studies, 31*(3), 115–133.

Jain, M. A., & Meena, A. K. (2024). Political communication in the age of misinformation: Addressing challenges in the digital era. *Asian Journal of Media and Communication, 14*(2), 67–81.

Kalogeropoulos, A., & Rossini, P. (2023). Unravelling WhatsApp group dynamics to understand the threat of misinformation in messaging apps. *New Media & Society, 25*(6), 1789–1807. https://doi.org/10.1177/14614448221091112

Lee, E. H., & Lee, T. D. (2021). A socio-behavioural approach to understanding the spread of disinformation. *Asian Communication Research, 18*(3), 201–218. https://doi.org/10.1080/19767286.2021.1935824

Masip, P., Suau, J., Ruiz-Caballero, C., & Capilla, P. (2021). News engagement on closed platforms: Human factors and technological affordances influencing exposure to news on WhatsApp. *Digital Journalism, 9*(7), 1023–1040. https://doi.org/10.1080/21670811.2020.1846579

Medeiros, B., & Singh, P. (2020). Addressing misinformation on WhatsApp in India through intermediary liability policy, platform design modification, and media literacy. *Journal of Information Policy, 10*(1), 276–297. https://doi.org/10.5325/jinfopoli.10.2020.0276

Monteiro, J. M. (2024). Democracy under attack: Social media and misinformation in Brazil. *Global Media Journal, 27*(2), 49–67.

Morosoli, S., Van Aelst, P., & Humprecht, E. (2022). Identifying the drivers behind the dissemination of online misinformation: A study on political attitudes and individual characteristics. *SAGE Open, 12*(1), 1–14. https://doi.org/10.1177/21582440221090212

Munusamy, S., Syasyila, K., Shaari, A. A. H., & Pitchan, M. A. (2024). Psychological factors contributing to the creation and dissemination of fake news among social media users: A systematic review. *BMC Psychology, 12*(4), 215–233. https://doi.org/10.1186/s40359-024-01212-x

Okocha, D. O., & Akpe, S. M. (2024). Civic engagement, public participation, and trust in digital space. *Journal of Media and Society, 19*(4), 213–230. https://doi.org/10.1007/s43636-024-00267-4

Omar, B. (2023). Countering fake news on WhatsApp in Malaysia: Current practices, future initiatives, and challenges ahead. *Mobile Communication and Online Falsehoods in Asia, 5*(2), 67–83.

Paris, B., & Pasquetto, I. (2024). Hidden virality and the everyday burden of correcting WhatsApp mis- and disinformation. *Cambridge Studies on Governing Social Media, 8*(1), 45–60.

Pennycook, G., & Rand, D. G. (2020). Fighting misinformation on WhatsApp: Evidence from behavioural interventions. *Psychological Science, 31*(8), 770–780. https://doi.org/10.1177/0956797620939054

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science, 31*(7), 770–780. https://doi.org/10.1177/0956797620939054

Santini, R. M., Tucci, G., & Salles, D. (2021). Do you believe in fake after all? WhatsApp disinformation campaign during the Brazilian 2018 presidential election. *Journal of Communication Inquiry, 45*(3), 245–261. https://doi.org/10.1177/0196859920961042

Shahzad, K., Iqbal, A., & Khan, S. A. (2023). Determinants of fake news diffusion on social media: A systematic literature review. *Global Knowledge Management Journal, 11*(2), 89–105.

Statista. (2024). Number of WhatsApp users worldwide. *Statista Reports.* https://www.statista.com/statistics/260819/number-of-monthly-active-whatsapp-users

Syamsi, V. V. (2023). Network society, fake news, and a hazard in the post-truth era: The case of WhatsApp and Facebook restriction by the Indonesian government in May 2019. *Asia-Pacific Research in Social Sciences and Humanities, 4*(1), 77–90.

Vese, D. (2022). End-to-end encryption and the limits of digital governance. *Journal of Information Ethics, 31*(1), 14–27. https://doi.org/10.3172/JIE.31.1.14

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. (2018). *Information disorder: Toward an interdisciplinary framework for research and policy making*. Council of Europe Report DGI(2017)09.