

Special Issue Sponsored by



Weaponising Intimacy: A Systematic Review of Vernacular Misinformation and Digital Media Fluency in Global South Closed Networks

Author

Mercy Munene 

Kabarak University

Email: mwanjira@kabarak.ac.ke

Article History: Submitted: 8th May 2026; Accepted: 8th August 2026; Published (online): 11th August 2026

Abstract

The migration of digital misinformation from open, algorithmic platforms to encrypted messaging applications presents a critical challenge. This is particularly true in the Global South, where encrypted platforms such as WhatsApp represent the primary means of internet connectivity. However, current scholarship on media has a pronounced Western slant, with much emphasis on algorithmic content moderation and reactive fact-checking. This systematic literature review aims to evaluate the empirical studies to explore the unique architecture of closed networks that supports the spread of misinformation (defined here as false information shared without malicious intent), and assess the effectiveness of proactive digital media fluency interventions in the Global South. Data were collected from peer-reviewed articles and institutional reports on the media ecosystems in Latin America, Africa and Asia, following PRISMA 2020 guidelines. Through thematic synthesis, it is uncovered that misinformation in dark social is more likely to be driven by ‘civic duty’ and homophily than by malicious bot amplification. Moreover, automated fact-checking does not perform well in a vernacular- and location-dependent manner, creating substantial time lags for institutional verifiers. The review demonstrates that proactive digital media fluency interventions show promise relative to traditional technical verification training. The study concludes that addressing dark social misinformation requires adapting Northern frameworks and urgently integrating contextualised, proactive media fluency competencies into the education system.

Keywords: digital media, fluency, vernacular, misinformation, WhatsApp, homophily, Global South

Introduction

The rise in digital misinformation has become a core problem in today's mass communication. Academic and institutional attention has been directed first to open, algorithmic platforms, such as Twitter and Facebook, that have transparent mechanisms of virality, through which researchers and moderators can monitor and fact-check content. Recently, the battleground of online misinformation has moved to "dark social" and the encrypted, largely private nature of messaging systems such as WhatsApp and Telegram (Kanthawala & Maddox, 2022). Closed networks, unlike open networks with weak-tie algorithmic recommendations, rely on strong-tie homophily between trusted peers (family, community groups). This closed architecture essentially shields content from traditional institutional moderation, lending an unearned sense of credibility to forwarded misinformation (Garimella, 2025).

A critical examination of the state of media research and theory today shows the extreme geographic imbalance of all current studies, models, and digital literacy programmes, which are all built from a Northern perspective. For the purposes of this study, a "Northern perspective" is operationalised as pedagogical frameworks reliant on text-based verification and algorithmic moderation designed primarily for open networks. When these are translated to the Global South, they often do not correspond to distinct socio-cultural contexts. Closed messaging networks are not just an alternative to internet access; they are the main means of daily communication and news for many populations in many developing regions around the world, especially in Africa, Latin America and Asia (Wasserman & Madrid-Morales, 2019). While these strong-tie architectures can sometimes support community-driven correction (Garimella, 2025), their closed nature essentially protects content from traditional institutional moderation. Many of the rumours spread on these encrypted platforms are conveyed through vernacular audio notes, culturally specific multimedia and indigenous languages that easily evade the traditional, text-based automated fact-checking systems used by most Western institutions (Cunliffe-Jones, 2020). In this digital infodemic, therefore, it is time for a more proactive approach in fact-checking that goes beyond responding to symptoms of the infodemic to rooting out the problem at the source. This study aims to comprehensively integrate the empirical studies available on the topic of proactive digital media fluency and the mechanisms of misinformation in closed networks in the Global South by using a Systematic Literature Review (SLR) to synthesise the existing evidence in a rigorous manner. To address three major research questions, this paper systematically collects peer-reviewed literature and important institutional reports: (1) How do the unique architectural affordances of closed networks namely encryption and homophily facilitate the rapid dissemination of misinformation? (2) What specific sociocultural and vernacular factors complicate traditional, institution-led fact-checking efforts in these regions? and (3) What empirical evidence currently exists regarding the efficacy of proactive digital media fluency interventions explicitly tailored for these dark social environments?

Literature Review

While digital misinformation studies is a new and evolving academic discipline, the dominant theories have been largely influenced by the architectures of open, algorithmically driven platforms that are largely used by the Global North. This Western perspective has resulted in a major theoretical deficit in how lies spread in different structural digital settings. As a result,

there are frameworks that are only applicable to alternative media ecosystems. To compensate for this imbalance, this section systematically outlines the conceptual structures of proactive digital media competency, and contrasts it with the more traditional approaches through examples of the types of learning outcomes that were identified. Moreover, it examines the structural affordances of the messenger apps used in the Global South and makes the gaps in empirical research in media studies in the Global South explicit. Finally, the context provided here reveals the need to replace broad-spectrum models of media literacy imported from other countries with interventions grounded in the local and socio-emotional contexts in which they are implemented (Wasserman & Madrid-Morales, 2019).

Mainstream scholarship has traditionally been reactive in addressing issues of digital misinformation, such as through the work of professional fact-checking and algorithmic content moderation. Yet, researchers increasingly claim that the more classic fact-checking approach is fundamentally inadequate when it comes to closed message spaces, where private content is not accessible for external monitoring or is not easily fact-checked via algorithms (Brenes Peralta et al., 2022). As a result, the literature has shifted its focus towards audiences intervening proactively, which is now fully conceptualised as "digital media fluency". Operationalised as an extension of traditional media literacy, digital media fluency incorporates nuanced technical, cognitive, and socio-emotional skills that are explicitly necessary to safely manoeuvre through peer-to-peer digital environments (Ogbodo et al., 2024). Most importantly, this cutting-edge model transcends the dated "information deficit" model, recognising that the spread of harmful misinformation is often not the result of ignorance, but may be driven by strong emotional triggers, identity signalling, or a feeling of civic duty towards communities (Garimella, 2025).

The affordances of closed messaging platforms change how misinformation is incubated and spread through networks. The adoption of end-to-end encryption generates highly opaque "dark social" environments that obscure malicious actors from external monitoring (Kanthawala & Maddox, 2022). In addition, misinformation in the Global South often does not adhere to traditional text-based communication, but rather takes the form of vernacular audio notes, highly localised memes, and culturally manipulated videos (Leaha & Canals, 2024). Platforms such as WhatsApp prioritise intimacy and effortless forwarding, making these visual and auditory falsehoods go viral at a frightening rate. Traditional, Western-centric digital literacy models cannot fully capture the unique challenges of mother-tongue broadcasting and oral traditions in digital spaces. Thus, institutional fact-checkers have serious lag problems, in constant competition with vernacular narratives that circulate in the local vernacular of different regions and user groups via regional, encrypted networks (Cunliffe-Jones, 2020).

A small number of primary studies have started to critically explore how certain WhatsApp practices and information disorders emerge in specific contexts in the Global South, but these studies have been scattered across isolated case studies and different academic fields. This fragmentation makes it extremely difficult to build scalable solutions and leaves educational institutions without a unified pedagogical plan. Without such synthesis, policymakers and educators may continually try to implement ineffective imported strategies that do not resonate with their audiences in the region (Diepeveen & Pinet, 2022). By systematically reviewing and aggregating this scattered empirical literature, this study directly addresses this academic deficiency. It aims to consolidate existing empirical evidence, evaluate the efficacy of tailored

regional interventions, and establish a robust, cohesive theoretical foundation to guide future empirical research and curriculum development within Global South digital media ecosystems.

Methodology

This study adopted a Systematic Literature Review (SLR) approach guided by the PRISMA 2020 statement. The systematic search was carried out across four main academic databases, namely Scopus, Web of Science (Core Collection), ProQuest and the Taylor & Francis Online database. To ensure contemporary relevance, the search was limited to peer-reviewed articles and key institutional reports published from 1 January 2016 (when the global digital misinformation crisis was recognised) to May 2026. The search string incorporated Boolean operators to capture the conceptual domains: ("misinformation" OR "disinformation" OR "fake news") AND ("WhatsApp" OR "Telegram" OR "closed messaging" OR "dark social") AND ("Global South" OR "Africa" OR "Latin America" OR "Asia") AND ("digital literacy" OR "media fluency" OR "fact-checking" OR "intervention"). The inclusion criteria were studies that dealt with populations or ecosystems in the Global South, that examined closed messaging platforms and that examined audience interventions (digital media fluency or verification behaviours). Exclusion criteria: studies were explicitly excluded if they focused exclusively on populations in the Global North, analysed fully open social media platforms lacking a closed-network component, or were purely technical/cryptographic analyses. Two independent researchers performed screening of the title, abstract and full text to minimise selection bias. Following the removal of 320 duplicates, 945 records were screened at the title/abstract stage. Of these, 125 reports were assessed for full-text eligibility, ultimately yielding 12 studies included in the final review. Methodological quality of the studies included in the final sample was assessed using the Mixed Methods Appraisal Tool (MMAT), so that studies with adequate empirical design, either qualitative ethnographic or quantitative experimental, were included for synthesis. Thematic analysis was used to identify patterns in the data extracted from the included studies regarding user behaviour and localised aspects of media literacy. To maintain the integrity of the distinct methodological paradigms represented in the corpus, qualitative ethnographic findings and quantitative experimental datasets were kept strictly separate during the data extraction and synthesis phases, preventing the improper pooling of disparate metrics.

Results

Search Results and Study Selection

The initial systematic search yielded a total of 1,250 records. Following deduplication and title/abstract screening, 945 full-text articles were retrieved and assessed. Finally, after following the PRISMA flow (Figure 1), 12 studies were included for the final thematic synthesis, meeting all inclusion criteria.

Figure 1. PRISMA 2020 flow diagram of study selection.

Identification	Records identified from: Databases (n = 1,250) Other sources (n = 15)	Records removed before screening: Duplicate records removed (n = 320)
Screening	Records screened (n = 945)	Records excluded (n = 815)
	Reports sought for retrieval (n = 130)	Reports not retrieved (n = 5)
	Reports assessed for eligibility (n = 125)	Reports excluded (n = 113)
Included	Studies included in review (n = 12)	

Source: Adapted from Page MJ et al.,2021.

Characteristics of Included Studies

The included studies exhibited a diverse range of methodologies, reflecting the complexity of researching encrypted spaces. Qualitative approaches and mixed methods were used. Geographically, they included a significant number of studies from Sub-Saharan Africa, Latin America, and South Asia, with WhatsApp most notably represented as the platform of analysis. A summary of the key characteristics of the synthesised literature is provided in Table 1.

Table 1. Summary of included studies.

Author(s) & Year	Country / Region	Platform Analysed	Methodology	Key Findings Related to Media Fluency & Misinformation
Garimella (2025)	Unspecified (Global South context)	WhatsApp	Mixed-Methods (Quantitative data + measures)	Community-driven corrections are infrequent; users show apathy toward self-correction or feel offended by public corrections; group power dynamics and familiarity influence interventions.
Iyengar et al. (2023)	India	Digital / Mobile	Quantitative (Within-subject experimental design, n=1002)	Cognitive inoculation via gamification (e.g., the "Bad News Game") significantly improves participants' ability to identify misinformation, specifically

Author(s) & Year	Country / Region	Platform Analysed	Methodology	Key Findings Related to Media Fluency & Misinformation
				conspiracy and impersonation techniques.
Leaha & Canals (2024)	Brazil	Visual Media / Digital	Qualitative (Multimodal and hybrid ethnography)	Investigates modes of countering "alternative images" visually; highlights how public demonstrations, artistic interventions, and fact-checking agencies serve as distinct modes of visual assessment and resistance.
Cunliffe-Jones (2020)	Africa	WhatsApp & Offline networks	Case Study / Institutional Analysis	False claims spread seamlessly between offline spaces (churches, mosques) and WhatsApp; effective interventions require a holistic, four-pronged approach to foster fact-checking skills both online and offline.
Diepeveen & Pinet (2022)	Global (Multiple Geographies)	Digital Media	Qualitative (Open-ended online consultations)	Youth suggest a contextualised view of misinformation; digital literacy interventions must expand beyond verification to include cybersecurity, unequal access, and broader digital citizenship.
Riera & Zommer (2020)	Argentina	Social Media	Case Study	Outlines the impact of "Chequeado," emphasising that partnering with key actors to teach media literacy and data journalism raises the cost of spreading misinformation.
Kanthawala	India	WhatsApp	Empirical /	Focuses on individual

Author(s) & Year	Country / Region	Platform Analysed	Methodology	Key Findings Related to Media Fluency & Misinformation
& Maddox (2022)			Behavioural	tactics of accuracy determination and how users "hide in the echo chamber," leading to fact-checking failures on encrypted apps.
Matanji et al. (2026)	Kenya & Senegal	Digital / General	Empirical / Comparative	Analyses the interplay between proactive media literacy and reactive fact-checking responses across two distinct African information ecosystems.
Adjei-Mensah et al. (2024)	Ghana	Information Networks	Empirical	Demonstrates how misinformation explicitly shapes critical decision-making among cocoa farmers, highlighting vulnerabilities in rural, specialised communities.
Ogbodo et al. (2024)	Nigeria	Digital / General	Empirical / Perspectives	Highlights the necessity of digital literacy as the primary tool for fighting fake news proliferation, specifically from the perspective of Nigerian netizens.
Wasserman & Madrid-Morales (2019)	Kenya, Nigeria, South Africa	Digital / General	Quantitative / Exploratory	A foundational study establishing baseline levels of media trust and fake news awareness across three major Sub-Saharan media markets.
Brenes Peralta et al. (2022)	Latin America (Implied)	WhatsApp	Empirical	Compares individual evaluation against formal fact-checking in determining users'

Author(s) & Year	Country / Region	Platform Analysed	Methodology	Key Findings Related to Media Fluency & Misinformation
				recognition and willingness to share COVID-19 fake news.

Thematic Findings

The synthesised literature revealed three distinct themes regarding how misinformation operates within closed networks.

The Weaponisation of Homophily and "Civic Duty"

A dominant theme is that misinformation in closed networks is rarely driven by malicious intent at the user level; rather, observational evidence suggests a strong association with homophily. Researchers (Garimella, 2025; Kanthawala & Maddox, 2022) have identified a behavioural pattern in which users forward unconfirmed information out of a sense of responsibility to alert their circle. Furthermore, the synthesised data show demographic detail of this behaviour. Often, youth who are more digitally fluent see misinformation but deliberately do not correct it in their family's WhatsApp group because they know they would be seen as disrespecting the elders. Motivation to self-correct, or offence at being corrected in public, can be poor among users (Garimella, 2025).

Structural Evasion via Vernacular and Audio-Visual Media

Closed-network misinformation structurally evades Western-centric fact-checking. The literature explicitly notes a shift toward vernacular audio notes and culturally specific imagery (Leaha & Canals, 2024). Platforms such as WhatsApp use encryption, so institutional fact-checkers are unable to reach in and scan such files. Furthermore, the complexities and ethics of real-time translation complicate these efforts; automated systems often fail to grasp the cultural nuances, humor, or specific idioms embedded in vernacular reporting, rendering direct textual translation ineffective. Misinformation is typically localised and targeted at specific communities, such as the cocoa farmers in Ghana, where farmers' choices regarding their crops are influenced by misinformation (Adjei-Mensah et al., 2024). False claims propagate easily between offline community spaces and WhatsApp, suggesting that social trust is bridging the digital divide (Cunliffe-Jones, 2020). The increasing tendency towards decentralising moderation away from centralised institutions and toward localised, community-driven options is a reaction against this

latency. Moreover, emerging research indicates that future fact-checking should incorporate AI systems dedicated to the processing of indigenous languages and oral inflections to identify vernacular audio notes before they become fully saturated, which would help flag these notes more effectively.

Efficacy of Proactive Digital Media Fluency Interventions

Traditional media literacy is often insufficient for encrypted environments. Interventions focused on behavioural inoculation have shown promising results, such as the adaptation of gamified tools in India (Iyengar et al., 2023). Moreover, as noted by youth voices, interventions need to be contextual, addressing broader issues such as cybersecurity and unequal access (Diepeveen & Pinet, 2022). In the Sub-Saharan context, combating the infodemic demands a strategic and coordinated combination of proactive media literacy and reactive fact-checking (Matanji et al., 2026; Riera & Zommer, 2020).

Discussion

This review identifies a major mismatch between the nature of the dark social threat and current pedagogical responses. The results validate that vectors for digital misinformation are fundamentally different in the Global South from those in the Global North.

Networked Intimacy and the Failure of the Deficit Model

The synthesis fundamentally challenges the prevailing "information deficit" model, which assumes that users share fake news simply because they lack correct facts. Instead, the data are congruent with theories of "networked intimacy" and Zizi Papacharissi's framework of affective publics. If a message is sent by a trusted church elder or family member, the emotional load and the trust between people become more important than the cognitive verification process, which is not always available on closed platforms. Addressing closed-network misinformation as a technological issue misses these important socio-cultural dynamics.

Shifting to Socio-Emotional Fluency

Adapting Northern tools for socio-emotional fluency in the Global South context must involve not only technical skills but also socio-emotional regulation: pausing, recognising emotional manipulation in a vernacular context, and coping with the complex social dynamics of correcting peers in strong-tie spaces such as navigating the ingrained respect for elders within family groups without causing social friction. Rather than entirely discarding Northern frameworks, existing

tools for prebunking and inoculation require rigorous localised adaptation. Specific learning outcomes related to vernacular media verification should be systematically integrated into education, replacing imported basic-education models with appropriate tertiary frameworks. Furthermore, emerging research indicates that future fact-checking should incorporate AI systems dedicated to processing indigenous languages and oral inflections.

Limitation of the Study

While this systematic review provides a robust synthesis, several limitations must be acknowledged. First, the Boolean search string was based largely on English-language terminology, which may have inadvertently missed highly relevant research published in native languages in regions such as Francophone Africa or Latin America. Secondly, the 12 studies included here provide rich qualitative and in-depth understandings, but given the small number of empirical, peer-reviewed interventions in this particular niche, the field of Global South "dark social" scholarship is still in its infancy.

Conclusion

The transition of misinformation to encrypted networks represents a profound shift in global communication. This review reaffirms the role of architecture in shaping sharing behaviour, heavily influenced by homophily and "civic duty". Reliance on vernacular multimedia means that conventional fact-checking is often unsuitable, resulting in serious latency problems. Proactive digital media fluency with a localised focus on regulation at the socio-emotional level therefore shows vital promise as a necessary countermeasure. Greater emphasis must be placed on making modules of closed-network dynamics a formalised part of the curriculum at higher education institutions, alongside the development of AI fact-checking tools specific to indigenous languages.

Funding Statement

This research received no external funding.

Conflict of Interest Statement

The authors declare that there is no conflict of interest.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Licensing Statement

© 2026 The Author(s). This article is published by *Kabarak Journal of Research & Innovation*

(KJRI) under the **Creative Commons Attribution 4.0 International (CC BY 4.0) License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Adjei-Mensah, J., Tham-Agyekum, E. K., Ndosi, J., Ankuyi, F., Nimoh, F., & Ankrah, D. A. (2024). Fake news on the farm: How misinformation shapes cocoa farmers' decisions in Jukwa, Ghana. *Journal of Agricultural & Food Information*, 1–22. <https://doi.org/10.1080/10496505.2026.2633141>
- 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>
- Cunliffe-Jones, P. (2020). From church and mosque to WhatsApp — Africa Check's holistic approach to countering 'fake news'. *The Political Quarterly*, 91(3), 596–599. <https://doi.org/10.1111/1467-923X.12899>
- Diepeveen, S., & Pinet, M. (2022). User perspectives on digital literacy as a response to misinformation. *Development Policy Review*, 40(S2), Article e12671. <https://doi.org/10.1111/dpr.12671>
- Garimella, K. (2025). Community-driven fact-checking on WhatsApp: Who fact-checks whom, why, and with what effect? *Proceedings of the Association for Information Science and Technology*, 62(1), 218–228. <https://doi.org/10.1002/pra2.1250>
- Iyengar, A., Gupta, P., & Priya, N. (2023). Inoculation against conspiracy theories: A consumer-side approach to India's fake news problem. *Applied Cognitive Psychology*, 37(2), 290–303. <https://doi.org/10.1002/acp.3995>
- 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>
- Leaha, M. A., & Canals, R. (2024). Resisting alternative images: An ethnography of visual disinformation in Brazil. *Cultural Anthropology*, 39(4), 533–563. <https://doi.org/10.14506/ca39.4.03>
- Matanji, F., Tully, M., Mudavadi, K. C., Diop, L., & Madrid-Morales, D. (2026). Media literacy and fact-checking as proactive and reactive responses to misinformation in Kenya and Senegal. *African Journalism Studies*, 47(1), 1–18.

- Ogbodo, J. N., Nwaoboli, E. P., Anorue, L. I., & Onyebuchi, C. A. (2024). Combating fake news and misinformation through digital media literacy: Perspectives from Nigerian netizens. *Journal of Communication Inquiry*.
- Riera, A., & Zommer, L. (2020). Using fact checking to improve information systems in Argentina. *The Political Quarterly*, 91(3), 600–604. <https://doi.org/10.1111/1467-923X.12895>
- Wasserman, H., & Madrid-Morales, D. (2019). An exploratory study of "fake news" and media trust in Kenya, Nigeria and South Africa. *African Journalism Studies*, 40(1), 107–123. <https://doi.org/10.1080/23743670.2019.1627230>