

## Determinants of Public Trust in Digital Media: A Systematic Literature Review

Catherine Chisala Kalumba<sup>1\*</sup>, Dr. Mwaka Siluonde<sup>1</sup>, Dr. Choolwe Mwiiya<sup>1</sup>

<sup>1</sup> Mulungushi University, Zambia

\* Corresponding Author

### Article Info

Volume 7, Issue 5

### Publication history

Accepted on 9 July 2026;

Published: 3 September 2026

### Key Words

Digital media trust, source credibility, digital literacy, misinformation, systematic review, conceptual framework, Sub-Saharan Africa

### Article Doi

10.59413/ajocs/v7.i5.10

### Abstract

**Background:** Public trust in digital media has emerged as a critical challenge in the contemporary information landscape, considering its potential implications for democratic participation, civic behavior, and information integrity. Although digital media use has increased significantly worldwide, the key determinants of audience trust in digital platforms have not been sufficiently theorized or empirically investigated, especially in African contexts.

**Methods:** This study presents a systematic literature review guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards. A structured search of peer-reviewed databases (African Journals Online, Google Scholar, Web of Science, and Scopus) was performed using predefined search strings. Studies published between 2013 and 2024 were considered, and after applying the inclusion/exclusion criteria, 52 peer-reviewed articles, book chapters, and institutional reports were selected for thematic synthesis.

**Findings:** The review identifies five principal determinants of public trust in digital media: (1) source credibility and institutional transparency, (2) digital literacy and media competence, (3) exposure to misinformation and disinformation, (4) platform design and algorithmic characteristics, and (5) socio-demographic and contextual factors. The synthesis reveals that trust formation is a dynamic, multi-layered process determined by individual competencies, platform characteristics and socio-political settings.

**Conclusions:** Building on the synthesized evidence, this study proposes an integrative Audience Trust in Digital Media (ATDM) conceptual framework that integrates the interaction of the determinants and positions trust as a socially constructed and contextually contingent outcome. The framework provides a basis for future empirical work, especially in Africa and other developing countries.

## 1. Introduction

Public trust in the media has been a longstanding concern for scholars in journalism, communication, and political science. Interest in this issue has further increased due to the digital revolution, which has brought about a major transformation in the way trust is perceived. Earlier, audiences were served by a relatively defined professional media landscape of broadcasters, newspapers and cable services. This situation has since changed, as the masses are now exposed to a more diverse and expansive digital environment saturated by citizen reporters, algorithmic curators, and anonymous content makers alongside professional journalists. This change in structure has made the assessment of information's credibility more difficult and has created what has been described as an "infodemic" (World Health Organization, 2020), a "trust deficit" (Newman et al., 2023), and a "post-truth" information environment (d'Ancona, 2017).

Understanding why audiences trust digital media is not merely a matter of satisfying academic curiosity. In reality, media trust impacts civic involvement, political activity, public health responses and social cohesion. When users are not confident in the trustworthiness of information sources, they are susceptible to misinformation and manipulation. Similarly, when citizens trust unreliable platforms or actors, the consequences for democratic discussion can be equally harmful. The COVID-19 pandemic clearly demonstrated how the collapse of trustworthy information sources provides space for the spread of dangerous misinformation (Roozenbeek et al., 2020).

In Sub-Saharan Africa, these dynamics are worsened by the unique historical, structural, and technological conditions in the region. The media landscape of many African countries is characterized by the legacy of a state-controlled media sector, political interference in the media sector, and limited press pluralism. Such factors have created highly divided and skeptical audience orientations toward media institutions (Wasserman & Madrid-Morales, 2020). The rapid expansion of mobile Internet access has democratized content production and sharing, fostering new avenues for the spread of misinformation (Banda, 2021). In particular, countries such as Zambia, Kenya, Nigeria, and South Africa occasionally witness the explosive use of social media to spread politically motivated fake news during elections (Gondwe, 2024; Manchishi et al., 2023).

Although these trends are significant, the academic literature on digital media trust remains highly fragmented. Several studies have been conducted in various fields, such as communication, political science, information science, and psychology; however, few systematic efforts have been made to integrate findings across these diverse fields. This limitation is especially prominent in Africa and developing countries. The existing literature is limited to considering misinformation or credibility perceptions in Western democracies, with less attention paid to the generalizability of existing frameworks.

This study addresses the identified gaps by conducting a systematic review of the current literature following PRISMA standards. Its main goal is to identify and synthesize the major determinants of public trust in digital media found in the existing literature and develop a conceptual framework that integrates these determinants into a coherent theoretical structure. This review contributes to the theoretical knowledge and practical design of interventions to foster trust in digital media environments, especially for researchers and policymakers in Africa and other developing countries.

## 2 Methodology

This study employed a systematic literature review approach. Systematic reviews are distinguished from traditional narrative reviews by their use of explicit and reproducible methods to search for, select, evaluate, and combine evidence (Tranfield et al., 2003). The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework was used as the protocol to guide the systematic review. This choice was influenced by the fact that the framework is a widely accepted and transparent reporting guideline for systematic reviews in social science (Page et al., 2021). In addition, this approach offers methodologically sound, transparent, and replicable merits required in the development of a credible evidence base on a complex and multi-disciplinary topic.

Accordingly, a systematic literature review is a suitable approach for this paper, considering its aim of synthesizing a fragmented research area and developing theoretically grounded conclusions on the research topic related to digital media trust. Given the scope of the topic, which spans communication theory, political communication, information science, and media psychology, a structured review method is necessary to ensure a comprehensive analysis that represents the full range of available evidence.

### 2.1 Search Strategy

A structured database search was conducted using various online databases, including Scopus, Web of Science, Google Scholar, African Journals Online (AJOL), and EBSCOhost's Communication Source. These databases were chosen to ensure proper coverage of indexed journals at both the international and regional levels.

The following Boolean search strings were applied in this study:

- "digital media" AND "trust" AND "credibility"
- "online news" AND "public trust" AND "audience"
- "misinformation" AND "media trust" AND "digital literacy"
- "social media credibility" AND "determinants"
- "news credibility" AND "Africa" OR "Sub-Saharan Africa" OR "Zambia"
- "digital media" AND "trust" AND "conceptual framework"

Online searches were conducted in English. The reference lists of key retrieved articles were also hand-searched to identify additional relevant sources that were not captured through database searches using backward citation tracking. On the other hand, forward citation tracking was applied to seminal papers such as Kohring and Matthes (2007) and Wardle and Derakhshan (2017).

### 2.2 Inclusion and Exclusion Criteria

To ensure consistency and rigor in the selection process, this study applied predetermined inclusion and exclusion criteria as follows:

#### Inclusion criteria

- Peer-reviewed journal articles, book chapters, and institutional reports
- Studies published between January 2013 and December 2024
- Studies addressing public trust, credibility, or perceived reliability of digital or online media
- Studies examining at least one identifiable determinant of media trust
- Studies published in English
- Studies relating to adult audiences (18 years and above)

#### Exclusion criteria

- Studies focused exclusively on traditional (non-digital) media without reference to digital platforms
- Studies published before 2013 (to keep the study relevant to the current context of digital media landscape)
- Grey literature, conference abstracts, opinion pieces, and editorials without empirical or theoretical grounding
- Studies not directly addressing trust, credibility, or related constructs
- Duplicate publications or studies reporting the same data

### 2.3 Study Selection Process

The first database query yielded 1,247 records that were screened. After discarding duplicates ( $n = 214$ ), 1,033 records were screened for titles and abstracts.

Of these, 841 were excluded because they were either not directly related to digital media trust or credibility or fell outside the defined time period. The remaining 192 articles were retrieved in full text for eligibility assessment. Another 140 were excluded because they were not peer-reviewed or failed to meet other exclusion criteria (67), were exclusively about traditional media (48), or did not mention a determinant of trust as a primary or secondary variable (25). Therefore, 52 sources were selected for thematic synthesis. Table 1 provides a summary of the PRISMA record flow as applied in the selection process.

Table 1: PRISMA Study Selection Summary

| PRISMA Stage                                | Records (n) |
|---------------------------------------------|-------------|
| Records identified from database searches   | 1,247       |
| Duplicates removed                          | 214         |
| Records screened (title and abstract)       | 1,033       |
| Records excluded at screening stage         | 841         |
| Full-text articles assessed for eligibility | 192         |
| Full-text articles excluded                 | 140         |
| Studies included in thematic synthesis      | 52          |

## 2.4 Data Extraction and Quality Appraisal

A standard data extraction form was created to document the following data from the included studies: author(s), publication year, country or geographical region, study design, sample description, theory used, main results and trustworthiness criteria. Quality appraisal was performed based on the criteria from the Mixed Methods Appraisal Tool (MMAT) for quantitative, qualitative, and mixed-methods studies (Hong et al., 2018). Studies were evaluated in terms of methodological transparency, sampling rigor, construct validation, and their findings. All 52 included studies met the minimum quality threshold.

## 2.5 Analytical Approach

Thematic synthesis was employed as the analytical method described by Thomas and Harden (2008). This method was performed in three iterative steps: coding the findings from the included studies line-by-line, developing descriptive themes, and generating analytical themes as overarching interpretations of the evidence. This approach enabled the inductive identification of patterns across diverse studies while maintaining fidelity to the original data. The emerging themes were the five principal determinants that shaped public trust in digital media, as detailed in the Findings section.

# 3 Findings

The thematic synthesis identified five principal determinants of public trust in the digital media. Each of the identified determinants is presented below as an analytical theme, supported by evidence from the literature.

## 3.1 Source Credibility and Institutional Transparency

The reviewed literature identified “source credibility” as the most frequently mentioned factor of public trust in digital media. The literature draws from the work of Hovland, Janis, and Kelley (1953) and extends it for the digital world by Kohring and Matthes (2007) to conceptualize source credibility in two dimensions of perceived expertise and perceived trustworthiness. Perceived expertise indicates how knowledgeable and competent a source is, whereas perceived trustworthiness refers to how honest, unbiased, and well-intentioned a source is.

In digital media environments, assessing source credibility is complicated by the collapse of traditional gatekeeping systems. Tandoc Jr. (2019) observes that with the rising number of online content created by both professionals and non-professionals, there is a significant credibility vacuum in the media environment. This means that audiences are expected to evaluate the credibility of content using heuristic cues, including website design, domain names, social media profile indicators, and follower counts, instead of relying on institutional cues. Studies have consistently established that professionally produced content from trusted news outlets tends to attract more trust than user-generated content, even when considering content quality (Humphrecht et al., 2022; Lindgren et al., 2024).

The current literature further shows that the transparency of institutions is closely linked to the credibility of sources, and in the post-digital age, it is gaining increasing attention. When audiences can easily identify or are aware of an editor's policy, ownership, correction process, and funding sources, they tend to have greater trust in such sources (Meier, 2023). However, where a lack of transparency is especially prevalent, such as in partisan digital news outlets and social media, the outcome is skepticism and mistrust (Molina et al., 2021). In essence, the literature reviewed showed that journalistic professionalism acts as a cue of credibility, with trust judgments being positively influenced by journalistic adherence to ethical codes, the use of named sources, and the provision of verifiable evidence.

In the African context, institutional trust is further mediated by historical experiences of media capture and political interference. Several reviewed studies note that in countries with a legacy of state-controlled media, such as Zambia, Zimbabwe, and Uganda, audience skepticism toward legacy media institutions persists even after liberalization, as citizens continue to question the editorial independence of established outlets (Banda, 2021; Mabweazara, 2021). This skepticism may paradoxically lead audiences to extend greater trust to decentralized digital sources, including social media, despite their low accountability standards.

### 3.2 Digital Literacy and Media Competence

Digital literacy, broadly defined as the ability to access, evaluate, produce, and share information through digital technologies, was the second key factor in strengthening public trust in digital media, as it was directly associated with trust judgments and as a moderating factor influencing all other determinants.

The reviewed literature makes a clear distinction between functional and critical digital literacies. While functional literacy is the ability to access digital media, critical literacy most effectively predicts trust judgments. Pennycook and Rand (2019) show the resistance of analytically oriented people, those who reason rather than intuit, to misinformation and their ability to “calibrate” their level of trust appropriately. Likewise, Guess et al. (2021) revealed that people who undergo media literacy training can differentiate between credible and non-credible content to a greater extent.

At the population level, Mabweazara (2021) established that digital literacy is a socio-stratified variable that describes the digital skills of younger, more educated, and urban groups. Stratification also has implications for the formation of trust: audiences with less digital literacy might trust peripheral signals (such as emotional resonance, peer sharing, or aesthetics) more than content-based assessments, making them more likely to accept content of low and high quality without critically evaluating it. In a study of Zambian audiences, Gondwe (2024) revealed that knowledge of digital literacy has a significant impact on how audiences trust news and how they can distinguish between it and misinformation, with less digitally literate audiences exhibiting less discriminating trust patterns.

Another notable factor in the current literature is that trust and digital literacy are not directly correlated. Some studies caution against expecting a direct link between increased digital literacy and more appropriate trust. Highly literate users can also be selective about trusting sources, as they might selectively engage in critical scrutiny of sources that challenge their beliefs or values while being more likely to trust sources that reinforce their beliefs (Trilling et al., 2017). This dynamic related to confirmation bias suggests that digital literacy interventions should consider cognitive and affective aspects, alongside technical skills, when it comes to understanding media evaluation.

### 3.3 Exposure to Misinformation and Disinformation

The third principal determinant is the impact of exposure to misinformation and disinformation on public trust in the media. Wardle and Derakhshan's (2017) taxonomy defines misinformation as false information shared without deceptive intent and disinformation as deliberately crafted information that is meant to mislead. The literature reviewed discusses both forms and their impact on trust as similar in the sense that they both lead to a loss of trust in digital media.

Repeated exposure to false, misleading, or manipulative information has been shown to erode trust in digital media as a whole, not just in specific news organizations or platforms where misinformation is found. Several studies on the reviewed corpus report such a 'spillover' effect. Roozenbeek et al. (2020) show that for those having access to COVID-19 misinformation in both social media and traditional media, exposure to misinformation led to a decrease in trust in both types of media. Molina et al. (2021) also propose that the more the public is exposed to stories that correct previous inaccuracies, the more they may become cynical of the media's credibility if they think corrections are politically driven.

The "fire hose" effect refers to the overwhelming amount of incorrect information presented to users on social media platforms, with algorithmic recommendation systems favoring engagement over accurate information (Tandoc Jr., 2019; Guess et al., 2021), which can make it difficult for users to verify the accuracy of the information they are consuming. This is especially true in politically charged situations, such as elections or public health crises.

In the African context, Manchishi et al. (2023) report that WhatsApp and Facebook are the most important platforms that spread politically motivated misinformation in Zambia, where politically motivated misinformation spreads rapidly during elections. This is extended by Masavah and Turpin (2024), who define weaker regulations, fewer fact-checking facilities, and weaker digital literacy as structural amplifiers of the effects of misinformation on the erosion of trust.

### 3.4 Platform Design and Algorithmic Characteristics

A fourth and increasingly prominent factor that affects digital media trust is the design of digital platforms, such as their visual design, user experiences, curation systems, and logic of algorithms. This theme differs from the previous ones, which are mainly about the audience, in that its focus is on the structural affordances of digital environments as trust-shaping forces.

Several dimensions of platform design and trust have been studied. Sundar (2008), the author of a heuristic-systematic model of credibility judgment that is still basic, lists several platform heuristics used to determine trust, such as platform professionalism, clarity of navigation, and the existence of external links or citations. These results have been confirmed in various other social media settings, indicating that professional profile pictures, verified badges, and a large number of followers are used to signal the credibility of accounts (Humprecht et al., 2022).

The impact of algorithmic content curation on shaping trust is a growing concern, as outlined in the literature. Algorithms found on platforms such as Facebook, YouTube, and X (formerly Twitter) are not neutral; they engage in ranking, filtering, and recommending content according to some sort of engagement signal that may have little or nothing to do with accuracy or credibility. As several authors have noted, this is a form of structural amplification of sensational, emotionally arousing, and frequently misleading reporting over measured and accurate reporting (Tandoc Jr., 2019; Trilling et al., 2017).

Other literature reviewed also mentions transparency regarding algorithmic and editorial processes as a trust-relevant variable at the platform level. The general public who knows that content is being algorithmically curated and knows the commercial reasoning behind that curation typically exhibits less trust in algorithmically served content than in editorially curated content (Meier, 2023). This discovery has implications for communicating with audiences about platform practices of content selection processes and how audiences learn critical audience practices for platform and content literacy.

### 3.5 Socio-Demographic and Contextual Factors

The fifth determinant emerging from the thematic synthesis relates to the diversity of socio-demographic and contextual factors that influence the development of trust judgments towards digital media. Unlike the preceding themes, this theme acknowledges that the influence of all other determinants is transacted through the contextual nature of each individual audience member and the macro-level contexts within which each member of the audience is situated.

Age is among the most widely studied sociodemographic factors that predict digital media trust. Younger generations are more likely to be digitally literate and more comfortable with digital platforms, but not necessarily more tied to or well-versed in digital trust. Some research even suggests that younger users might be more skeptical of digital news while simultaneously consuming it (Newman et al., 2023). This habit is primarily related to their increased exposure to misinformation environments. In contrast, older audiences might have more general trust in trusted media outlets and might be more susceptible to misinformation online because of their limited knowledge of digital media (Guess et al., 2021).

Trust formation is significantly influenced by the level of education, with digital literacy and critical thinking being the primary pathways. The audience's educational level is associated with higher source verification behaviors and more sophisticated trust calibration by media type (Gondwe, 2024; Humprecht et al., 2022). The literature reviewed has given less systematic focus to gender differences in digital media trust, but some researchers assert that women tend to be more cautious about sharing unverified information, which may represent different social norms on information sharing.

The political and media systems in which audiences are situated have strong conditioning influences on the formation of trust at the contextual level. Baseline trust in the media is greater and more institutionally embedded in countries with strong democratic institutions, regulations, and media traditions (Nordic countries: Newman et al., 2023). In comparison, people living in contexts with high levels of political polarization, media capture, and low press freedom typically have more fragmented and contingent orientations of trust, using personal networks and the recommendations of others as sources of trust when they do not have trusted institutional signals (Madrid-Morales, 2020; Banda, 2021).

Furthermore, literature evidence shows that geographic location within countries also matters significantly when it comes to trust in digital media. Rural audiences tend to have more uniform and abstract trust orientations than their urban counterparts, whose higher broadband Internet penetration, access to a variety of media, and formal educational institutions make them more likely to have more differentiated and context-sensitive trust orientations. This urban-rural digital divide in trust formation is especially noticeable in African research (Mabweazara, 2021; Gondwe, 2024), and is salient regarding the need for media literacy interventions and trust-building strategies to be contextually tailored.

## 4 Proposed Conceptual Framework

### 4.1 Rationale for framework development

The five determinants identified in the literature review are not independent of each other. Instead, they act dynamically, some as antecedents of others and some as direct and indirect influences on the development of trust. While many theories account for some aspects of audience credibility assessment, such as Source Credibility Theory (Hovland et al., 1953; Kohring & Matthes, 2007) and the Elaboration Likelihood Model (Petty & Cacioppo, 1986), they were developed in pre-digital or early digital contexts and fail to fully explain the structural and technological complexities of today's digital media environments.

The literature reviewed indicates the following needs for an integrative framework: (1) trust as an outcome variable; (2) both individual- and structural-level determinants; (3) digital literacy and contextual factors as mediation and moderation variables; and (4) the framework should be generalizable across various national, cultural, and technological contexts. This need is addressed in the proposed Audience Trust in Digital Media (ATDM) framework.

### 4.2 Framework Architecture

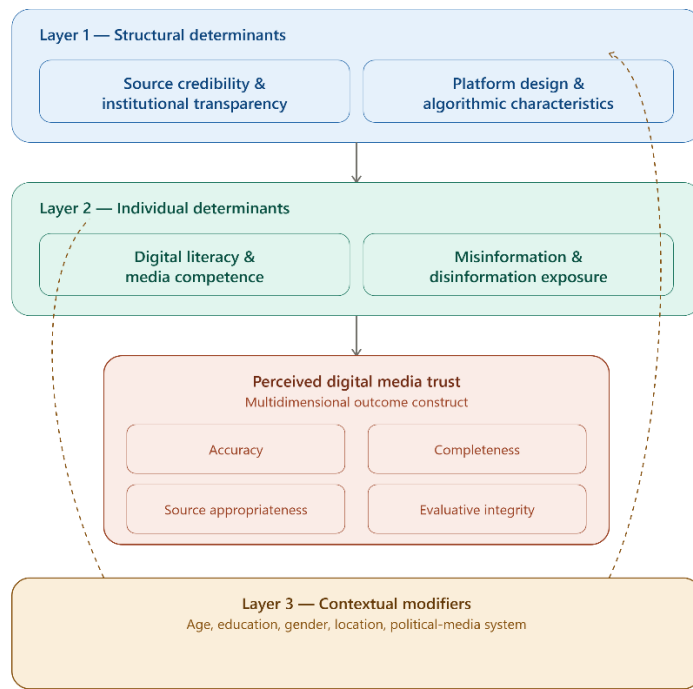
The ATDM framework organizes the five determinants of trust in digital media into three analytical layers: Structural Determinants, Individual Determinants, and Contextual Modifiers. These layers together account for the formation of Perceived Digital Media Trust as the central outcome construct.

**Layer 1** – Structural Determinants reflect structural conditions at the institutional and platform levels: source credibility and institutional transparency (Determinant 1) and platform design and algorithmic characteristics (Determinant 4). These factors shape the information context in which audiences live and set expectations for the quality of their trust judgment. Regulatory policy, journalistic norms, and platform governance can shape their practices, but to a large extent, they are not directly under the control of individual audience members.

**Layer 2** – Individual Determinants includes aspects of digital literacy and media competence (Determinant 2) and exposure to misinformation and disinformation (Determinant 3) that apply at the level of audience competencies/exposure. These factors intervene between structural conditions and trust outcomes. Digital literacy is a protective resource that allows audiences to critically assess structural conditions and defy the erosion of trust associated with misinformation. On the other hand, misinformation exposure is a stressor that drains the resources of trust, especially in environments with low digital literacy levels.

**Layer 3** – Contextual Modifiers includes socio-demographic and contextual factors (Determinant 5) that are moderators in both the structure and individual layers. The relationships between the other four determinants of trust and trust outcomes are influenced by the intensity and direction of these relationships with age, education, gender, geographic location, and political-media system characteristics.

The central outcome construct, identified as Perceived Digital Media Trust, is conceptualized as a multidimensional construct comprising four sub-dimensions, as established by Kohring and Matthes (2007). These subdimensions include trust in the factuality of reporting (accuracy), trust in the selectivity of topics (completeness), trust in the selectivity of sources (source appropriateness), and trust in the journalistic assessment of facts (evaluative integrity). Multidimensional conceptualization is desired over single-item and unidimensional measures because of the nuances of how audiences judge the reliability of digital media content.



Dashed lines: Layer 3 moderates the relationships in Layers 1 and 2  
 Solid lines: direct influence flowing toward perceived trust

Figure 1: ATDM Framework – Layers, Determinants, and Intervention Entry Points

### 4.3 Key Propositions of the Framework

The ATDM framework generates several theoretically-informed propositions that could serve as a basis for further empirical research, including the following:

- Proposition 1: Source credibility and institutional transparency exert a direct and positive effect on perceived digital media trust, such that audiences who perceive digital media sources as expert, transparent, and accountable demonstrate higher levels of trust across all sub-dimensions.
- Proposition 2: Digital literacy has a direct positive impact on the perceived level of trust in digital media, and a positive moderating influence on the impact of source credibility on the perceived level of trust, with digitally literate audiences more likely to see and reward credible sources of digital information.
- Proposition 3: Exposure to misinformation exerts a direct and negative effect on perceived digital media trust, and this negative effect is amplified in contexts of low digital literacy.
- Proposition 4: The mechanism and nature of a platform and algorithmic features moderate the relationship between source credibility and audience trust perceptions, which means that algorithmic environment distortions weaken the impact of journalistic professionalism on audience trust perceptions.
- Proposition 5: Socio-demographic and contextual factors moderate all of the above relationships, with education, digital literacy, and media system quality serving as the most significant moderating variables in African and developing country contexts.

### 4.4 Contribution and Applicability of the Framework

- The proposed ATDM framework offers diverse contributions to digital media trust theory. First, it offers an integrative account that combines all three perspectives (structural, individual, and contextual) in a single analytical matrix, overcoming the fragmentation in the current literature. Second, the framework is sensitive to context; it acknowledges that the processes of trust formation are not the same in different national contexts and that they are influenced by the historical, political, and technological contexts of the countries involved in the study. Third, it identifies which determinants can be addressed directly at the audience level (such as digital literacy and critical media evaluation) and which at the institutional/policy level (such as source transparency, platform governance, and regulatory frameworks). This distinction is useful when designing strategies for building trust that are appropriately targeted.
- This framework is particularly relevant in the contexts of African and developing countries. This is because the dynamic between low institutional trust, high exposure to misinformation, and the volatility of digital literacy and platform environments in these regions creates unique experiences of trust that are not adequately represented in frameworks developed in the context of Western democratic societies. The ATDM framework can be applied as a basis for operationalizing the variables, testing them, and developing predictive models in future empirical research in Zambia and similar settings.

### 4.5 Discussion

#### Principal Insights from the Review

This systematic review and the resulting conceptual framework provide valuable insights for both academic and practical applications. The first and most



foundational revelation is that trust in digital media is not a unitary or static phenomenon but is multidimensional, dynamic, and deeply contextual in nature. Thus, audiences are not fixed in their trust or distrust of digital media. Instead, they adjust their levels of trust depending on the platform, type of information, and social context. Additionally, audiences change their trust calibrations over time due to their experiences with information quality and institutional behavior.

The second insight concerns the centrality of digital literacy as a protective resource in contemporary digital information environments. The reviewed literature consistently identifies digital literacy as a direct predictor of appropriate trust calibration and a buffer against the trust-eroding effects of misinformation exposure. This finding has immediate practical implications: investments in digital literacy education are among the most cost-effective interventions available to governments, civil society organizations, and media institutions seeking to rebuild audience trust. However, the current literature emphasizes that digital literacy programs must move beyond functional skills toward critical competencies, including the ability to recognize algorithmic curation, understand ownership structures, and apply source verification practices.

The third finding concerns the platforms' structural role. The reviewed literature provides valuable evidence that platform design and algorithmic characteristics are not merely neutral conduits for content but are active determinants of the information environments in which trust judgements are made. This implies that trust is not solely a property of content or audiences but is also a function of the technological architectures that mediate between them. This has important implications for media regulation and platform accountability, suggesting that transparency requirements, algorithmic auditing, and content labelling standards can meaningfully support public trust.

### Implications for Research in African Contexts

The ATDM framework and the evidence synthesis on which it is based have specific implications for research in African and Sub-Saharan contexts, including Zambia, as follows: The reviewed literature highlights several features of the African digital media environment that require particular theoretical attention: the persistence of historical distrust toward legacy institutions; the dominance of mobile-first, closed-platform information environments (particularly WhatsApp); the structural undersupply of professional, independent journalism in many markets; and the acute digital literacy divide that separates urban from rural audiences and educated from less-educated citizens.

Future studies in the African context must, therefore, be sensitive to the moderating effect of the political-media system, the importance of platform-specific dimensions of trust rather than abstract "digital media" trust frameworks, and the differing trust impacts of information spread through peer networks in contexts of high use of WhatsApp. The ATDM framework offers a theoretical framework for this research program that can be adapted and modified to fit the context of this study.

### Limitations of the Review

This review has several limitations. First, the limitation of publications written in English only may have resulted in the omission of relevant scholarship in French, Portuguese, and African languages, potentially underrepresenting scholarship from Francophone Africa. Second, the use of published peer-reviewed sources, which is necessary for quality assurance, may also create publication bias if studies reporting positive results are more likely to be published than those reporting null results.

Moreover, while the search period (2013-2024) is necessary to reflect the current digital media environment, it may result in the exclusion of earlier systematic searches that might have provided important foundational insights. Furthermore, the proposed ATDM framework is conceptual rather than empirical and is based on the evidence reviewed. Thus, its propositions need to be tested with empirical data, and its applicability to a variety of national contexts must be established in a context-sensitive way.

---

## 5 Conclusion and Recommendations

This systematic literature review synthesizes evidence from 52 scholarly sources to identify five main factors that influence public confidence in digital media: source credibility and transparency of the platforms, digital literacy and media competence, exposure to misinformation and disinformation, platform design and algorithmic characteristics, and sociodemographic and contextual factors. The review shows that trust formation in digital media contexts is a complex, multi-layered process involving interactions between individual competencies, structural conditions, and contextual factors.

In response to these findings, this study proposes the Audience Trust in Digital Media (ATDM) conceptual framework, which organizes the five determinants into three analytical layers (structural, individual, and contextual) and positions perceived digital media trust as a multidimensional outcome construct. The framework offers a theoretically grounded and contextually adaptable foundation for future empirical inquiry, including research on digital media credibility and audience behavior in Zambia.

Overall, this review has three principal implications. First, a multi-pronged approach is needed at the institutional level (increased transparency, accountability, and professional standards), platform level (algorithmically governed and designed platforms), and audience level (digital literacy education and critical media engagement). Second, heavy contextualization is needed: the context matters significantly; context-specific adaptation is needed for trust frameworks and interventions to be applicable in African media environments. Third, the ATDM framework and this review can serve as a starting point for empirical research, including mixed-methods studies that can test and refine the propositions conveyed in the framework.

As digital media continue to expand in reach and influence throughout Sub-Saharan Africa and around the world, understanding the conditions under which audiences extend or withhold trust will become an increasing subject of academic and policy interest. This review represents a key milestone toward meeting that imperative while also providing the conceptual foundation for a program of audience-centered and contextually grounded research on the factors shaping digital media trust.

---

## References

- Banda, F. (2021). Digital journalism and media ethics in Africa: Navigating new frontiers of practice and accountability. *African Journalism Studies*, 42(3), 1–18.
- d'Ancona, M. (2017). *Post-Truth: The New War on Truth and How to Fight Back*. London: Ebury Press.

- Gondwe, G. (2024). Audience perception of fake news in Zambia: Examining the relationship between media literacy and news believability. *International Communication Research*, 47. Available at: [https://scholar.harvard.edu/sites/scholar.harvard.edu/files/ggondwezunda/files/audience\\_perception\\_of\\_fake\\_news2.pdf](https://scholar.harvard.edu/sites/scholar.harvard.edu/files/ggondwezunda/files/audience_perception_of_fake_news2.pdf) [Accessed 28 October 2025].
- Guess, A. M., Nagler, J., and Tucker, J. A. (2021). The consequences of online misinformation exposure: Evidence from Facebook and Twitter. *Political Communication*, 38(3), 1–23.
- Hong, Q. N., Fábregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M. P., Griffiths, F., Nicolau, B., O' Cathain, A., Rousseau, M. C., Vedel, I., and Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information*, 34(4), 285–291. <https://doi.org/10.3233/EFI-180221>
- Hovland, C. I., Janis, I. L., and Kelley, H. H. (1953). *Communication and Persuasion: Psychological Studies of Opinion Change*. New Haven: Yale University Press.
- Humprecht, E., Castro Herrero, L., Blassnig, S., Brüggemann, M., and Engesser, S. (2022). Media systems in the digital age: An empirical comparison of 30 countries. *Journal of Communication*, 72(2), 145–164.
- Kohring, M., & Matthes, J. (2007). Trust in news media: Development and validation of a multidimensional scale. *Communication Research*, 34(2), 231–252. <https://doi.org/10.1177/0093650206298071>
- Lindgren, E., Lindholm, T., Vliegthart, R., Boomgaarden, H. G., Damstra, A., Strömbäck, J., and Tsifti, Y. (2024). Trusting the facts: The role of framing, news media as a (trusted) source, and opinion resonance in perceived truth in statistical statements. *Journalism and Mass Communication Quarterly*, 101(4), 981–1004. <https://doi.org/10.1177/1077699022117117>
- Mabwezara, H. M. (2021). Towards reimagining the 'digital divide': Impediments and circumnavigation practices in the appropriation of the mobile phone by African journalists. *Information, Communication and Society*, 24(3), 344–364. <https://doi.org/10.1080/1369118X.2020.1834602>
- Manchishi, K., Pindayi, B., and Mambwe, E. (2023). Financial sustainability of social media-driven publications in Zambia. In *New Journalism Ecologies in East and Southern Africa* (pp. 207–223). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-031-23625-9\\_12](https://doi.org/10.1007/978-3-031-23625-9_12)
- Masavah, V. M., and Turpin, M. (2024). Fake news in developing countries: Drivers, mechanisms, and consequences. In *the International Conference on Implications of Information and Digital Technologies for Development* (pp. 127–142). Cham: Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-66986-6\\_10](https://doi.org/10.1007/978-3-031-66986-6_10)
- Meier K. (2023). Transparency in journalism: A systematic review of the literature. *Journalism Practice*, 17(5), 993–1011. <https://doi.org/10.1080/17512786.2021.1984956>
- Molina, M. D., Sundar, S. S., Le, T., and Lee, D. (2021). "Fake news" is not simply false information: A concept explication and taxonomy of online content. *American Behavioral Scientist*, 65(2), 180–212. <https://doi.org/10.1177/0002764219878224>
- Newman N., Fletcher R., Robertson C. T., Eddy, K., and Nielsen, R. K. (2023). Reuters Institute Digital News Report 2023 (<https://www.digitalnewsreport.org/>). Oxford: Reuters Institute for the Study of Journalism.
- Page, M. J., & McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Pennycook, G., and Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences*, 116(7), 2521–2526. <https://doi.org/10.1073/pnas.1806781116>
- Petty, R. E., and Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19, 123–205.
- Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L. J., Recchia, G., van der Bles, A. M., and van der Linden, S. (2020). Susceptibility to misinformation about COVID-19 across 26 countries. *Royal Society Open Science*, 7(10), 201199. <https://doi.org/10.1098/rsos.201199>
- Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger and A. J. Flanagin (Eds.) *Digital Media, Youth, and Credibility* (pp. 73–100). Cambridge, MA: MIT Press.
- Tandoc Jr., E. C. (2019). The facts of fake news: A research review. *Sociology Compass*, 13(9), e12724. <https://doi.org/10.1111/soc4.12724>
- Thomas, J., and Harden, A. (2008). Methods for thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8(1), 1–10. <https://doi.org/10.1186/1471-2288-8-45>
- Tranfield, D., Denyer, D., Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, 14(3), 207–222. <https://doi.org/10.1111/1467-8551.00375>
- Trilling, D., Van Klingeren, M., and Tsifti, Y. (2017). Selective exposure, political polarization, and possible mediators: Evidence from the Netherlands. *International Journal of Public Opinion Research*, 29(2), 189–213. <https://doi.org/10.1093/ijpor/edw003>
- Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policymaking. Strasbourg: Council of Europe. <https://rm.coe.int/information-disorder-report/1680764666>
- Wasserman, H., and Madrid-Morales, D. (2020). Credibility in the African context. *Journalism Practice*, 14(10): 1265–1279.
- World Health Organization (WHO). (2020). Infodemic management: Infodemics. Geneva: WHO. <https://www.who.int/health-topics/infodemic.int/health-topics/infodemic>