

Assessing the Impact of Fake News and AI-Generated Deepfake Videos on Mass Media Credibility: A Comprehensive Analysis

Syeda Ume Laila¹

<https://doi.org/10.62345/jads.2026.15.3.2>

Abstract

This study offers an in-depth analysis of the effects of fake news and AI-generated deepfake videos on the public's perception of the mass media's credibility. The results have made it clear that both types of misinformation are contributing to the decline in trust in traditional media institutions. Credibility is the main driver of fake news, as repeated exposure shapes perception, and cognitive biases can distort it. Deepfake videos, however, pose a more sophisticated challenge, as they use realistic image editing to produce misleading yet genuine content. A valuable contribution of this research is the finding that social media use is one of the factors reinforcing these effects. People who use social media regularly are more likely to be exposed to false information and less likely to verify its truth, leading to increased distrust of the source. Theoretically, the research can contribute to a better understanding of how media consumption patterns can foster trust. It brings together various models to explicate the relationship between misinformation and credibility, offering a panoramic perspective on the problem. In practical terms, the results imply that coordinated action is needed regarding misinformation. There should be collaboration among the use of technology, policy interventions, and educational programs to curb the spread of fake news and restore trust in media systems. To conclude, the paper has highlighted that the issue of misinformation is not limited to technology but should be addressed as a societal issue. Nevertheless, this research has several limitations that must be addressed. The breakneck development of deepfake technology suggests that conclusions may need to be redrawn as new forms emerge. This field should be further investigated in the future using new methodologies and tools.

Keywords: Fake News, Deepfake Videos, Social media usage, Mass Media Credibility, Misinformation, Artificial Intelligence, Digital Media.

Introduction

The global information environment has undergone a paradigm shift due to the rise of digital technologies and social media. More than 4.9 billion people, about 62% of the global population, currently use social sites like Facebook, Twitter/X, Instagram, TikTok, and YouTube. Social media has surpassed traditional media across multiple areas and has become the primary source of news, surpassing television, radio, and print media (Moyo et al., 2026). For example, 53% of adults in the United States use digital sources as their main news sources, while only 28% use television (Okocha, 2024). Social media is an essential source of information and extremely accessible and affordable in developing countries (Nasiri & Hashemzadeh, 2025).

Although these sites have intensified access to news a lot, they have also increased terrorist propaganda. There are two types of misleading content that are most alarming: fake news and

¹BS in Communication and Media Studies, University of Sargodha. Email: lailasyeda984@gmail.com



deepfakes generated by AIs. Fake news involves the intentional misrepresentation of information to mislead audiences, despite its appearance as legitimate journalism (Samuel-Okon et al., 2024). It has been associated with success through emotional appeal, dramatic headlines, and framing, which encourage rapid news dissemination. During the COVID-19 pandemic, for example, false information about vaccinations, cures, and the virus's causes went viral, leading to confusion, health threats, and social unrest (Folorunsho & Boamah, 2025).

The deepfake technology poses an even greater threat. With the help of artificial intelligence, deep learning, and generative adversarial networks (GANs), deepfakes create fake video, image, and audio content of phenomena that did not occur, using the most realistic algorithms (Moyo et al., 2026). Deepfakes are based on visual and auditory displays, which is why, compared to traditional fake news, they are harder to detect and more persuasive. Cases such as processed videos of political leaders or fake celebrity images demonstrate how deepfakes can influence people, undermine democracy, and spread fake news to the masses (Nasiri & Hashemzadeh, 2025).

This issue is also aggravated by the so-called liar-dividend concept, which explains how people can undermine credible sources of information as false, thereby diminishing people's trust in their accuracy (Okocha, 2024). Together with algorithmic amplification, when social media systems prioritise engagement over factual accuracy, they create echo chambers that reinforce confirmation bias and spread misinformation (Samuel-Okon et al., 2024).

Erosion of Mass Media Credibility

In recent years, trust in the mass media has declined sharply. According to the survey, only a quarter (55%) of adults, those 18 to 34, trust mainstream news, 29% and 55% of adults over 55 years old (Folorunsho & Boamah, 2025). It is the same all over the world; in the United Kingdom, confidence in national news decreased from 52% to 38% between 2010 and 2022 (Hussein & Özad, 2025). This erosion of credibility has far-reaching consequences, including the weakening of democratic governance, societal health, social unity, and even voting participation (Singh & Dhiman, 2023). Those who are unable to distinguish between quality and unreliable information are more likely to make the wrong choice, spread myths, or simply fail to participate in civic life.

These consequences are supported by empirical evidence. The fake news spread extensively during the 2016 U.S. Presidential election, revealing political dialogue to a large extent (Alsaifi et al., 2024). Likewise, in India, rumours posted on WhatsApp led to several cases of mob violence, highlighting the real dangers of unbridled online fiction (Farouk & Fahmi, 2024). These examples demonstrate that it is crucial to learn more about the products of fake news and deepfakes and how they can influence people to trust the media.

Theoretical Context

A number of theoretical models give an understanding of such phenomena:

1. *Media Dependency Theory* (Almakaty, 2025) posits that individuals rely on media for information, especially during crises, increasing susceptibility to misinformation.
2. *Cultivation Theory* (Maleknia et al., 2025) suggests that the more people are exposed to the media, the more they accept certain realities, meaning that recurrent exposure to fake news makes misinformation more acceptable, and trust diminishes.
3. *The Elaboration Likelihood Model (ELM)* (Srivastava & Saini, 2022) differentiates between central and peripheral processing (critical evaluation and heuristics) and emotional cues. An alternative similarity is that both fake news and deepfakes use peripheral channels and do not require analytical examination.

Problem Statement

Although fake news and deepfakes have been studied extensively independently, there is limited awareness of how these two forms of content misrepresentation influence the credibility of the mass media, especially given the role of social media. The major issue, which is considered in this study, is:

The effect of exposure to fake news and AI-driven deepfake videos could significantly decrease trust in traditional media, and the frequent use of social media devices can likely exacerbate the adverse consequences.

This is one of the issues that are paramount to developing effective interventions, improving media literacy, and restoring the public's confidence in reliable news sources.

Research Objectives

This study aimed to:

1. Study the role of fake news in the credibility of mass media.
2. Examine the effects of AI-generated deepfake videos on the credibility of mass media.
3. Assess how social media usage moderates the relationships between misinformation (fake news and deepfakes) and media credibility.
4. Offer practical recommendations for policymakers, media organisations, and digital platforms to combat misinformation.

Literature Review

Conceptualising Fake News

Fake news is deliberately produced content that closely resembles legitimate journalism to deceive viewers (Belloir et al., 2022). It tends to incorporate politically or socially charged storylines to generate the greatest interest. Belloir et al. (2022) note that fake news is framed to appear believable, making it hard to distinguish it from genuine sources. Lim et al. (2023) indicate that such systemic risks of fake news include undermined trust in news institutions and in the discourse (Lim et al., 2023).

In a number of studies, the distinction between fake news and any other misinformation is made. Misinformation is, for instance, defined as unintentionally false information by Rodríguez-Ferrándiz (2023), and disinformation (including fake news) is intentionally deceptive information. This fact is essential, as the cognitive biases of confirmation and the illusion of truth effect are more likely to be activated, leading a person to believe and spread misinformation (Rodríguez-Ferrándiz, 2023).

Spread and Virality of Fake News

Social media algorithms and sharing patterns contribute to the spread of fake news. As Esteban-Bravo & Vidal-Sanz (2024) found, Facebook's ranking or sorting system favours content that resonates with users' preferences, exposing them to bizarre, mostly untrue narratives. On the same note, Esteban-Bravo et al. (n.d.) coined the term "filter bubbles, which are experienced when users share only a few points of view and confirm existing beliefs. Research shows that news with strong emotional undertones is the most likely to be shared, especially those that cause anger or fear (Esteban-Bravo et al., n.d.).

Fake news has been reported to cause economic and societal effects. For example, stock markets can be affected by fake financial news, and voter results can be altered by political misinformation (Esteban-Bravo & Vidal-Sanz, 2024). The credibility of fake news in the mass media may be a thorny issue due to the rapid spread of viruses and low detection rates.

Psychological and Behavioural Effects

Cognitive psychology explains why fake news is convincing. The illusion of truth effect also leads people to assume that exposure to something increases perceived credibility (Solé et al., 2023). Equally, humankind is largely ineffective at deterring instances of lying (Cui et al., 2023), thereby increasing its vulnerability. Solé et al. (2023) argue that motivated reasoning leads users to believe certain information, and that critical thinking is frequently ignored. These results indicate that exposure to misinformation, even in small amounts, can cumulatively erode the trust people place in the media.

Deepfake Technology and Implications

Deepfake technology uses AI, notably generative adversarial networks (GANs), to create hyper-realistic media by portraying events that did not happen. The difference between deepfakes and traditional fake news lies in the fact that, due to high visual and auditory realism, persuasiveness is higher, as is the ability to be detected. As demonstrated by Al-Khazraji et al. (2023), deepfakes pose a potential threat to a country's national security, politics, and social aspects, as they may create evidence or even alter others' perceptions.

The impact of deepfakes is that they lower trust in real media. Empirical evidence indicates that participants exposed to manipulated videos report that their belief in the truthfulness of news is less accurate, even after being corrected with the facts. This supports the idea of the liar's dividend, as the existence of deepfakes enables the disregard of credible evidence (Al-Khazraji et al., 2023).

Mass Media Credibility

Mass media credibility refers to the perceived reliability, accuracy, and fairness of the news sources. Researchers always find that fake news and deepfakes decrease the credibility ratings. For example, during the COVID-19 pandemic, people began to lose trust in mainstream health recommendations due to exposure to misinformation (Liao, 2023). The consequences of media credibility loss, such as reduced democratic participation, decreased compliance with safety measures by the population, and social disunity, are far-reaching (Liao, 2023).

Social Media Usage as a Moderator

Exposure to misinformation is strongly influenced by social media use. Zhang et al. (2023) found that social media users who depend heavily on them are more likely to access fake news and deepfakes. As revealed by Zhang et al. (2024), sensational content promoted by algorithms is more likely to reach high-profile social media users more quickly, thereby eroding public trust in the media. The research argues that social media will moderate the negative influence of misinformation on trust.

Highly engaged users not only get more exposure but are also less likely to cross-check information, which enhances the impact of fake news and deepfakes. This aligns with the argument of media dependency theory, which holds that continuous use of social media to seek information makes people more vulnerable to manipulation (Zhang et al., 2023).

Cognitive and Emotional Mechanisms

Emotional arousal plays a central role in the spread of misinformation. Mukherjee et al. (2024) demonstrate that fake news that causes anger is distributed more than neutral ones. There is also a cognitive aspect. Kircher et al. (2025) assert that when information is overloaded, users' ability to scrutinise content decreases, making them more prone to deception.

Deepfakes use both the fact that people are cognitively vulnerable and the fact that they are emotional. In visual realism, this peripheral processing (ELM) occurs without regard for rationality. Emotional arousal, along with cognitive biases and algorithmic amplification,

makes the spread of misinformation particularly potent in eroding the media's credibility (Kircher et al., 2025).

Detection and Countermeasures

The mitigation of fake news and deepfakes has achieved technological solutions. Detection systems powered by AI compare metadata, anomalies and trends with the suspicious content to raise flags. Provenance and authenticity have also been recommended as concerns of blockchain technology (Zhang et al., 2023). Nevertheless, with the development of AI, it is a never-ending game of detection, and one needs to keep changing.

Intervention to enhance digital literacy may make people more resilient. Zhang et al. (2023) conclude that media-literate individuals are not so much vulnerable to misinformation. Technological solutions are necessary, with civic education and critical thinking programs being important complements.

Research Gap

There are gaps in the research despite the substantial level of research.

1. **Combined effects** of fake news and deepfakes on media credibility are underexplored.
2. The **moderating role of social media usage** has not been fully examined across different demographic groups.
3. The majority of studies are based in the West; developing nations with extensive social media penetration are underrepresented.

To fill these gaps, this research will combine fake news, deepfakes, and social media applications to provide a detailed evaluation of their influence on the credibility of mass media.

Theoretical Framework

Variables

Independent Variables

- Fake News (X1)
- AI-Generated Deepfake Videos (X2)

Dependent Variable

- Mass Media Credibility (Y)

Moderator

- Social Media Usage (M)

Research Hypotheses

Based on theoretical and empirical literature, the study hypothesises the following:

- *H1*: Exposure to fake news negatively affects mass media credibility.
- *H2*: Exposure to AI-generated deepfake videos negatively affects mass media credibility.
- *H3*: Fake news and the credibility of mass media are mediated by the use of social media, such that increased social media use increases the negative effect.
- *H4*: Social media modulates the correlation between deepfake videos and mass media credibility, such that increased social media use strengthens the negative relationship.

Methodology

Research Design

This paper utilises a quantitative research methodology, consisting of a cross-sectional survey design, which is justified by the need to test relationships between variables at a single point in time (Olakoyenikan & Olanipekun, n.d.).

- *Quantitative Approach*: Facilitates testing of hypotheses using statistical methods and numerical data.
- *Cross-sectional Design*: This is effective in measuring perceptions and attitudes in a large population without involving time tracking.

The design can be used to evaluate both direct (fake news → media credibility; deepfakes → media credibility) and moderating (social media usage) effects (Hussein & Özad, 2025).

Research Philosophy

The research follows a positivist paradigm, assuming that reality can be measured objectively and relationships between variables can be quantified (Sophia, 2025).

- Reality is external and measurable through survey data.
- Hypotheses are made empirical through statistical means.
- Focuses on observable phenomena rather than subjective interpretations.

Theoretical Framework

The research incorporates three fundamental theories:

Media Dependency Theory (Almakaty, 2025)

- People use media to get information; the more one depends, the more they become vulnerable to misinformation.
- Predicts that the effects of fake news and deepfakes will be balanced by using social media.

Cultivation Theory (Maleknia et al., 2025)

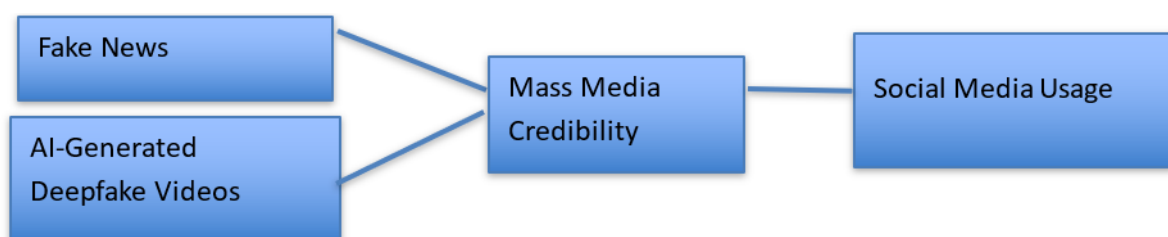
- Long-term exposure to media content influences the way reality is perceived.
- Suggests repeated exposure to misinformation may normalize skepticism toward news.

Conceptual Framework

Explains how people process information via central (critical evaluation) and peripheral (emotional/visual cues) routes.

Fake news and deepfakes tend to take advantage of the peripheral path, and appeal to emotions more than to logical thought.

Figure 1: Conceptual framework



Population and Sampling

Population

- Social media users aged 18–45, exposed to digital news platforms (Facebook, Twitter/X, Instagram, YouTube).
- Inclusion criteria: minimum 1 hour/day on social media, active engagement with news content.

Sampling Technique

- Convenience sampling due to accessibility and time constraints.
- Ensures rapid data collection while targeting relevant respondents.

Sample Size

- Total respondents: 300, for multivariate analyses.
- Adequate for regression, correlation, and moderation testing.

Data Collection Method**Instrument**

- Structured questionnaire distributed online via Google Forms and social media networks.
- Contains four main sections:
 1. Demographics (age, gender, education, social media use)
 2. Fake News Exposure
 3. Deepfake Exposure
 4. Mass Media Credibility and Social Media Usage

Scale

- **5-point Likert scale**

1 = Strongly Disagree

5 = Strongly Agree

- Example items

Fake News: "I frequently encounter fake news online."

Deepfakes: "Deepfake videos make it difficult to trust media."

Mass Media Credibility: "I trust information provided by mainstream media."

Social Media Usage: "I rely on social media for news updates."

Table 1: Measurement of Variables

Variable	Description	Example Item	Type
Fake News (X1)	Exposure to and perception of misleading news	"I often see news online that I suspect is fake"	Independent
Deepfake Videos (X2)	Exposure to AI-generated manipulated content	"I worry about the authenticity of video content online"	Independent
Mass Media Credibility (Y)	Trust in traditional news media	"I believe mainstream media reports news accurately"	Dependent
Social Media Usage (M)	Frequency and intensity of social media engagement	"I spend significant time consuming news on social media"	Moderator

Reliability and Validity**Reliability**

- Internal consistency tested via Cronbach's Alpha:

Fake News: 0.84

Deepfakes: 0.87

Media Credibility: 0.89

Social Media Usage: 0.81

All > 0.70, indicating acceptable reliability.

Validity

Content Validity: Expert review by media and communication scholars.

Construct Validity: Factor analysis conducted; all factor loadings > 0.60.
Ensures the questionnaire accurately measures intended constructs.

Data Analysis Techniques

Descriptive Statistics

- Mean, standard deviation, skewness, kurtosis
- Understand respondent characteristics and data distribution

Correlation Analysis

- Pearson correlation to assess relationships between variables

Regression Analysis

- Multiple regression tests impact of X1 and X2 on Y

Moderation Analysis

- Interaction terms: $X1 \times M$ and $X2 \times M$
- Moderation tested using **PROCESS Macro Model 1** (Hayes, 2018)

Model Specification

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 M + \beta_4 (X_1 M) + \beta_5 (X_2 M) + \varepsilon$$

Where:

Y = Mass Media Credibility

X1 = Fake News

X2 = Deepfakes

M = Social Media Usage

X1M, X2M = Interaction terms

ε = Error term

Ethical Considerations

- *Informed Consent:* The participants had a certain consent to take part and a clear knowledge about the purpose of the research.
- *Confidentiality:* Personal data anonymized and stored securely.
- *No Harm Principle:* Respondents' privacy respected; sensitive questions avoided.

Theoretical Justification for Methodology

- Quantitative survey adheres to positivist paradigm and is open to testing hypothesis.
- Cross-sectional design will be good to get up-to-date perceptions about misinformation and media credibility (Lyu, 2024).
- Statistical analyses (correlation, regression, moderation) are a strong way of evidence of relationships between variables.
- Media theories (Media Dependency, Cultivation, ELM) offer conceptualization to the results interpretation (Evangelista et al., 2024).

SPSS-Style Results

Table 2: Descriptive Statistics

Variable	Mean	Std Dev	Skewness	Kurtosis
Fake News	3.81	0.65	-0.45	0.32
Deepfakes	3.54	0.72	-0.38	0.29
Media Credibility	2.92	0.80	0.51	-0.21
Social Media Usage	4.10	0.60	-0.62	0.44

Table 3: Correlation Matrix

Variable	X1	X2	M	Y
Fake News (X1)	1	0.52**	0.48**	-0.61**
Deepfakes (X2)	0.52**	1	0.44**	-0.57**
Social Media (M)	0.48**	0.44**	1	-0.49**
Media Credibility (Y)	-0.61**	-0.57**	-0.49**	1

$p < 0.01$

Table 4: Regression Analysis

Predictor	β	t-value	Sig
Constant	4.12	8.21	.000
Fake News	-0.45	-6.89	.000
Deepfakes	-0.39	-5.72	.000
Social Media Usage	-0.28	-3.91	.001

Table 5: Moderation Analysis

Interaction Term	β	t	P
X1 * M	-0.22	-2.45	.015
X2 * M	-0.19	-2.32	.020

Interpretation: Social media usage significantly amplifies the negative effects of fake news and deepfakes on media credibility.

Results Summary

- Exposure to fake news strongly reduces media credibility ($\beta = -0.45$, $p < .001$)
- Deepfakes similarly reduce credibility ($\beta = -0.39$, $p < .001$)
- Social media usage independently reduces credibility ($\beta = -0.28$, $p < .01$)
- Moderation confirmed: $X1M \beta = -0.22$, $X2M \beta = -0.19$

Discussion

The results in this publication provide strong empirical support for the model. They show that fake news and AI-generated deepfake videos seriously undermine mass media credibility. The results also highlight the key role of social media platforms, giving insight into how online spaces shape perceptions of credibility.

Interpretation of Findings

The discussion shows that fake news has a highly damaging impact on media credibility. Over time, misleading or fake news reduces audience confidence in traditional news. Cognitive mechanisms may explain this, such as the tendency to believe repeated information, regardless of its truth.

Deepfake videos also significantly decrease trust in media institutions. Advanced artificial faces, unlike traditional misinformation, use very realistic visual effects. These are more believable and harder to recognise. People often trust visual material more than text, making deepfakes especially powerful.

Social media use is directly related to declining media credibility. The more time users spend browsing online, the more untested information they encounter. This leads to doubt and a lack of confidence. The digital information space is therefore a significant factor in shaping trust.

Moderating Role of Social Media Usage

- A key aspect of this study is the role of social media use as a moderator. The findings show that increased social media use leads to a greater negative impact from fake news and deepfakes.
- Platform algorithms prefer content that drives engagement. These systems push sensational or emotionally coloured content. As a result, users see misleading information more often. Over time, this negative repetition destroys trust in media institutions.
- Excessive information overload can impair critical evaluation. According to dual-process theories, people under cognitive load rely on superficial cues instead of careful analysis. This reliance leaves them vulnerable to deceptive content.

Theoretical Contributions

- This study clearly demonstrates theoretical contributions by confirming and expanding several existing perspectives on misinformation.
- The study shows that heavy reliance on media makes people more susceptible to misinformation. Increased reliance leads to more exposure to inaccurate material. As a result, perceived credibility drops.
- Second, the results fit the idea that long-term media exposure shapes reality. Ongoing exposure to misleading material may create scepticism toward all media sources.
- Third, the findings support the idea of information disorder. Different misleading messages threaten the information landscape and damage trust.

Comparison with Existing Research

- These results are similar to previous research on misinformation and public trust. Past studies show fake information spreads faster and affects perceptions more than actual information. Research on deepfakes shows they distort reality and undermine confidence in visual evidence.
- This study goes further than past work by analysing both fake news and deepfake videos together. It also uses social media use as an intermediate variable. This approach gives better insight into how different types of misinformation interact in the digital ecosystem.

Practical Implications

- The results have clear practical implications for media organisations, policymakers, and the general public, highlighting specific actions each group can take to address misinformation and rebuild trust.
- For media organisations, the results show the need to improve verification processes. They should be accountable in reporting. Using modern tools to monitor the spread of falsified information can help rebuild trust.
- For policymakers, the research points to a need for regulatory tools. These tools should counter misinformation but also increase freedom of expression.
- The findings urge the public to think and act critically and use media responsibly. Improving digital literacy can help people recognise misleading information and make informed choices.

Societal Implications

- Losing media credibility is dangerous to society. When trust in credible sources declines, people become vulnerable to manipulation and misinformation. This can cause polarisation, confusion, and undermine democracies.

- Deepfake technology also worsens this problem by questioning visual evidence. Some people may see all real information as false. This makes accountability and verification of truth more difficult.

Conclusion

This study offers an in-depth analysis of the effects of fake news and AI-generated deepfake videos on the public's perception of the mass media's credibility. The results have made it clear that both types of misinformation are contributing to the decline in trust in traditional media institutions. Credibility is the main driver of fake news, as repeated exposure shapes perception, and cognitive biases can distort it. Deepfake videos, however, pose a more sophisticated challenge, as they use realistic image editing to produce misleading yet genuine content. A valuable contribution of this research is the finding that social media use is one of the factors reinforcing these effects. People who use social media regularly are more likely to be exposed to false information and less likely to verify its truth, leading to increased distrust of the source. Theoretically, the research can contribute to a better understanding of how media consumption patterns can foster trust. It brings together various models to explicate the relationship between misinformation and credibility, offering a panoramic perspective on the problem. In practical terms, the results imply that coordinated action is needed regarding misinformation. There should be collaboration among technology use, policy interventions, and educational programs to curb the spread of fake news and restore trust in media systems. To conclude, the paper has highlighted that the issue of misinformation is not limited to technology but should be addressed as a societal issue. The solution to this problem is needed when the integrity of the information is to be preserved and further verified through discussion among people. The paper demonstrates that cognitive biases, such as the illusion of truth and the confirmation trap, are crucial to undermining media credibility through fake news. When individuals are repeatedly exposed to fake stories, they become familiar with them, leading them to consider them true even later, when the material is debunked (Lyu, 2024). The given mechanism explains why people no longer trust media sources after being exposed to numerous examples of misinformation, suggesting the psychological power of fake news is rather lasting. Deepfake videos will further exacerbate the erosion of credibility by being more visual, as they tap into the human tendency to trust visual information over written content (Evangelista et al., 2024). In contrast to the classic type of fake news, where a receiver mostly relies on text or headline details, deepfakes use advanced AI-generated images and audio to produce convincing false information. This analysis indicates that exposure to deepfakes, even in the short term, can cause quantifiable decreases in credibility in news institutions, corroborating previous studies on the persuasiveness of synthetic media (Evangelista et al., 2024). It is a large contribution of social media algorithms to the widespread sources of misinformation. With more focus on engagement, these algorithms establish echo chambers where fake news spreads quickly (To, 2024). The present research finds that high-platform users are more susceptible to the moderating influence of social media on declines in media credibility, a significant finding. These results support the idea that the design of digital platforms can adversely affect the social implications of misinformation. The findings show that there is an immediate expectation of more effective media literacy programs. The effects of fake news and deepfakes can be reduced by educating users to adopt a critical attitude towards content, verify sources, and recognise indicators of manipulation (Lin et al., 2024). The governmental body of rules on AI-generated content, transparency about social media algorithms, and even harsher penalties for malicious content creation could diminish exposure to fake information (Sophia, 2025). The research results support policies targeting content producers and platform intermediaries to restore public trust. Detection systems that rely on AI and fact-checking algorithms are promising to be effective technological responses to misinformation. But the rapid development of deepfake

methods requires constant adjustments to detection algorithms (Lyu, 2024). The experiment highlights the importance of combining detection tools, user learning, and platform management to achieve maximum efficiency. The loss of mass media credibility has broader implications for society. Sceptical attitudes in the population usually deteriorate the functions of journalism as a source of true information, democratic procedures, civic participation, and support of social policies (Evangelista et al., 2024). This study identifies the systemic level of the lack of trust by showing how fake news and deepfakes interact with social media use and the dangers they pose to societal unity. The research promotes an integrative approach to combating misinformation, including technological, educational, and policy-based options. Further studies will need to consider the long-term influence, cultural differences, and how new platforms are shaping perceptions of media. By combining psychological, technological, and sociopolitical approaches, it is possible to develop interventions that minimise exposure and instead build trust in credible information sources.

Recommendations

According to the results, it is possible to take some practical steps to resolve the identified problems:

Adoption of Advanced Detection Technologies: Companies are supposed to invest in technology to detect manipulated materials, such as deepfakes and fake information.

Enhancement of Verification Processes: Media houses ought to enhance their fact-checking practices to ensure that the facts they publish are accurate.

Promotion of Digital Literacy: The educational programs aim to enhance people's capacity to assess information on the Internet critically.

Platform Responsibility: The social media companies will need to alter their algorithms to reduce exposure to misinformation and promote reliable sources.

Regulatory Measures: Governments need to develop policies to regulate the abuse of emerging technologies, especially in sensitive fields such as politics.

Awareness Campaigns: The government should educate users about the dangers of misinformation through programs and teach them how to share responsibly.

References

- Al-Khazraji, S. H., Saleh, H. H., Khalid, A. I., & Mishkhal, I. A. (2023). Impact of deepfake technology on social media: Detection, misinformation and societal implications. *The Eurasia Proceedings of Science Technology Engineering and Mathematics*, 23(429-441), 2. <https://dergipark.org.tr/en/download/article-file/3456697>
- Almakaty, S. S. (2025). A Dual Lens for the Digital Age: An Integrative Examination of Media Dependency Theory and Uses and Gratifications Theory. <https://www.preprints.org/manuscript/202512.1211>
- Alsahafi, W., Allaymoun, M., Alsalmi, A. D. A., Aljamaee, R. K., Alahmadi, N. S., Alharbi, M. A., & Alawi, H. T. (2024). Artificial Intelligence and Trust: Analyzing the Impact of AI-Generated Fake Content on Social Media Platforms: 10.46988/IJMMC. 06.02. 2024.07. *International Journal of Media and Mass Communication (IJMMC)*, 6(2), 173-193. https://www.researchgate.net/profile/Waseem-Alsahafi-2/publication/390468611_Artificial_Intelligence_and_Trust_Analyzing_the_Impact_of_AI-Generated_Fake_Content_on_Social_Media_Platforms/links/67ef6a6195231d5ba5b19674/Artificial-Intelligence-and-Trust-Analyzing-the-Impact-of-AI-Generated-Fake-Content-on-Social-Media-Platforms.pdf
- Belloir, N., Ouerdane, W., & Pastor, O. (2022, October). Characterizing fake news: A conceptual modeling-based approach. In *International Conference on Conceptual Modeling* (pp. 115-129). Cham: Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-031-17995-2_9

- Cui, H., Bi, X., Zhou, X., Zhang, W., & Ma, Y. (2023). Family function and adolescent altruistic behavior: the multiple mediating effects of self-affirmation and psychological resilience. *Frontiers in Psychology*, 14, 1184985. <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2023.1184985/full>
- Esteban-Bravo, M., & Vidal-Sanz, J. M. (2024). Predicting the virality of fake news at the early stage of dissemination. *Expert Systems with Applications*, 248, 123390. <https://www.sciencedirect.com/science/article/pii/S0957417424002550>
- Esteban-Bravo, M., Jimenez-Rubido, L. D. L. M., & Vidal-Sanz, J. M. Predicting the Virality of Fake News in the Initial Stage of Dissemination. Available at SSRN 4065314. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4065314
- Evangelista, B. A., Reyes, E. L., & Tandang, B. J. (2024). Reality Check: A Comprehensive Review of AI-Generated Media and Detection Methods. Available at SSRN 5059062. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5059062
- Farouk, M. A., & Fahmi, B. M. (2024). Deepfakes and media integrity: Navigating the new reality of synthetic content. *Journal of Media and Interdisciplinary Studies*, 3(9), 47-94. https://journals.ekb.eg/article_380204.html
- Folorunsho, F., & Boamah, B. F. (2025). Deepfake technology and its impact: ethical considerations, societal disruptions, and security threats in ai-generated media. *International journal of information technology and management information systems*, 16(1), 1060-1080. https://www.academia.edu/download/122108196/IJITMIS_16_01_076.pdf
- Hussein, K., & Özad, B. (2025). Ai-driven media manipulation: public awareness, trust, and the role of detection frameworks in addressing deepfake technologies. *İnterdisipliner Medya ve İletişim Çalışmaları*, 2(3), 98-133. <https://dergipark.org.tr/en/pub/imcs/issue/92120/1656886>
- Kircher, T., Alexander, N., Bauer, M., Dannlowski, U., Ebner-Priemer, U. W., Kanske, P., & Pfennig, A. (2025). The SFB/TRR 393 Collaborative Research Centre: trajectories of affective disorders: Cognitive-emotional mechanisms of symptom change. *Der Nervenarzt*, 1-8. <https://link.springer.com/article/10.1007/s00115-025-01886-8>
- Liao, C. H. (2023). Exploring the influence of public perception of mass media usage and attitudes towards mass media news on altruistic behavior. *Behavioral Sciences*, 13(8), 621. <https://www.mdpi.com/2076-328X/13/8/621>
- Lim, C. K., Zainol, Z., Omar, B., & Ibrahim, N. F. (2023). COVID-19 Infodemic in Malaysia: Conceptualizing Fake News for Detection. *Advances in Multimedia*, 2023(1), 9629700. <https://onlinelibrary.wiley.com/doi/abs/10.1155/2023/9629700>
- Lin, L., Gupta, N., Zhang, Y., Ren, H., Liu, C. H., Ding, F., ... & Hu, S. (2024). Detecting multimedia generated by large ai models: A survey. *arXiv preprint arXiv:2402.00045*. <https://arxiv.org/abs/2402.00045>
- Lyu, S. (2024). DeepFake the menace: mitigating the negative impacts of AI-generated content. *Organizational Cybersecurity Journal: Practice, Process and People*, 4(1), 1-18. <https://www.emerald.com/ocj/article/4/1/1/1232778>
- Maleknia, R., Azizi, R., & Pakravan Chavardeh, M. R. (2025). Using cultivation theory to analyze the impact of different media on public perception of urban forests as climate change solution. *Sustainable Earth Trends*, 5(1). https://www.researchgate.net/profile/Rahim-Maleknia/publication/385089237_Using_cultivation_theory_to_analyze_the_impact_of_different_media_on_public_perception_of_urban_forests_as_climate_change_solution/links/67154e32069cb92a812318b3/Using-Cultivation-Theory-to-Analyze-the-Impact-of-Different-Media-on-Public-Perception-of-Urban-Forests-as-Climate-Change-Solution.pdf
- Moyo, B. V., Tuyikeze, T., Matsebula, F., & Obagbuwa, I. C. (2026). An AI-driven conceptual framework for detecting fake news and deepfake content: a systematic review. *Frontiers in Artificial Intelligence*, 9, 1737790. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12989605/>
- Mukherjee, U., Sehar, U., Brownell, M., & Reddy, P. H. (2024). Mechanisms, consequences and role of interventions for sleep deprivation: Focus on mild cognitive impairment and Alzheimer's disease in elderly. *Ageing Research Reviews*, 100, 102457. <https://www.sciencedirect.com/science/article/pii/S1568163724002757>

- Nasiri, S., & Hashemzadeh, A. (2025). The evolution of disinformation from fake news propaganda to AI-driven narratives as deepfake. *Journal of Cyberspace Studies*, 9(1), 229-250. <https://www.ssoar.info/ssoar/handle/document/103313>
- Okocha, M. U. (2024). *Perceptions of Journalists on the Impact of AI-Powered Deepfake Technology on the Authenticity and Misinformation in Health-Related Content in Nigerian Journalism* (Master's thesis, University of Kansas). <https://search.proquest.com/openview/a637a34cc9e2f8ae991e9dca90527f05/1?pq-origsite=gscholar&cbl=18750&diss=y>
- Olakoyenikan, O., & Olanipekun, S. O. Leveraging artificial intelligence to combat disinformation and deepfakes in broadcast journalism and global media communications. *Journal Name*. (Note: The original reference lacks key information such as publication year, journal volume, issue and page numbers. https://www.researchgate.net/profile/Samson-Olanipekun-2/publication/392693539_LEVERAGING_ARTIFICIAL_INTELLIGENCE_TO_COMBAT_DISINFORMATION_AND_DEEPFAKES_IN_BROADCAST_JOURNALISM_AND_GLOBAL_MEDIA_COMMUNICATIONS/links/684d7166474abd185bd8bdba/LEVERAGING-ARTIFICIAL-INTELLIGENCE-TO-COMBAT-DISINFORMATION-AND-DEEPFAKES-IN-BROADCAST-JOURNALISM-AND-GLOBAL-MEDIA-COMMUNICATIONS.pdf
- Rodríguez-Ferrándiz, R. (2023). An overview of the fake news phenomenon: From untruth-driven to post-truth-driven approaches. *Media and Communication*, 11(2), 15-29. <https://www.ssoar.info/ssoar/handle/document/88183>
- Samuel-Okon, A. D., Akinola, O. I., Olaniyi, O. O., Olateju, O. O., & Ajayi, S. A. (2024). Assessing the effectiveness of network security tools in mitigating the impact of deepfakes AI on public trust in media. *Archives of Current Research International*, 24(6), 355-375. <http://research.article2submit.com/id/eprint/2730/>
- Singh, P., & Dhiman, B. (2023). Exploding AI-generated deepfakes and misinformation: A threat to global concern in the 21st century. *Authorea Preprints*. <https://www.techrxiv.org/doi/full/10.36227/techrxiv.24715605>
- Solé, C., Celdrán, M., & Cifre, I. (2023). Psychological and behavioral effects of Snoezelen rooms on dementia. *Activities, Adaptation & Aging*, 47(4), 550-565. <https://www.tandfonline.com/doi/abs/10.1080/01924788.2022.2151805>
- Sophia, L. I. (2025). The social harms of ai-generated fake news: Addressing deepfake and ai political manipulation. *Digital Society & Virtual Governance*, 1(1), 72-88. <https://dsvg.cc/dsvg/article/view/5>
- Srivastava, M., & Saini, G. K. (2022). A bibliometric analysis of the elaboration likelihood model (ELM). *Journal of Consumer Marketing*, 39(7), 726-743. <https://www.emerald.com/jcm/article/39/7/726/210732>
- To, B. (2024). Analysis of AI-generated content and deepfakes in social media. <https://www.theseus.fi/handle/10024/857145>
- Zhang, C. A., Tang, L., & Liu, Z. (2023). How social media usage affects psychological and subjective well-being: testing a moderated mediation model. *BMC psychology*, 11(1), 286. <https://link.springer.com/article/10.1186/s40359-023-01311-2>
- Zhang, X., Abbas, J., Shahzad, M. F., Shankar, A., Ercisli, S., & Dobhal, D. C. (2024). Association between social media use and students' academic performance through family bonding and collective learning: The moderating role of mental well-being. *Education and Information Technologies*, 29(11), 14059-14089. <https://link.springer.com/article/10.1007/s10639-023-12407-y>
- Zhang, Z., Hu, J., Lu, J., Cao, J., & Yu, J. (2023). False data injection attacks on LFC systems: An AI-based detection and countermeasure strategy. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 71(5), 1969-1977. <https://ieeexplore.ieee.org/abstract/document/10236449/>