S3H Working Paper Series

Number 4: 2025

Impact of media literacy on deepfake awareness and perception: A Study using ELM

Maria Aurangzeb

Dr. Sundus Mustaqeem

October 2025

School of Social Sciences and Humanities (S3H)
National University of Sciences and Technology (NUST)
Sector H-12, Islamabad, Pakistan

S3H Working Paper Series

Number 4: 2025

Impact of media literacy on deepfake awareness and perception: A Study using ELM

Maria Aurangzeb
Student, School of Social Sciences and Humanities, NUST
E-mail: maria.mmc23s3h@student.nust.edu.pk
, abbasimariaa92@gmail.com

Dr. Sundus Mustaqeem
Assistant Professor, School of Social Sciences and Humanities, NUST
E-mail: sundus@s3h.nust.edu.pk

October 2025

School of Social Sciences and Humanities (S3H)
National University of Sciences and Technology (NUST)
Sector H-12, Islamabad, Pakistan

S3H Working Paper Series

Contents

List of Tables

Abstract

Deepfake technology has emerged as a significant threat in today's digital age, posing a risk to trust in media, privacy, and societal norms. This study analyses the influence of media literacy on the awareness and understanding of deepfake technology among young people in Pakistan, specifically those aged 18–35. Using the Elaboration Likelihood Model (ELM) as a guide, the researcher collected data from 301 participants through an online survey, employing the snowball sampling method. The analysis employed several techniques to assess the reliability of the survey scales and descriptive statistics to gain demographic insights. Findings showed that while participants had a moderate grasp of media literacy and deepfake awareness, media literacy did correlate with a better understanding and awareness of the risks associated with deepfakes. Additionally, media literacy was significantly associated with a greater reliance on central route processing, where participants critically analyzed deepfake content, indicating more thoughtful engagement with media. Furthermore, Future research should explore qualitative methods to gain a deeper understanding of users' perception and behavior about deepfake media.

Keywords: Deepfake, Media Literacy, Elaboration Likelihood Model, Digital Users, Youth, Pakistan, Misinformation

1. Introduction

We exist in what many scholars describe as a *post-truth society*, where misinformation often overshadows objective facts and distorts public understanding across journalism, politics, and social discourse (Harari, 2018; McIntyre, 2018). A prominent outcome of this environment is the rise of *deepfakes,* AI-generated images, audio, or videos that closely mimic authentic content (Vaccari & Chadwick, 2020). Their realism undermines audiences' ability to differentiate truth from fabrication, eroding trust in digital communication. Deepfakes present multifaceted risks, including personal harm such as reputational damage and psychological distress, as well as broader societal threats like misinformation campaigns, political manipulation, and targeted harassment (Chesney & Citron, 2019; Cruz, 2024).

In Pakistan, these concerns are particularly significant. With over 66 million active social media users, the digital landscape is highly vulnerable to synthetic media (DataReportal, 2025). Deepfakes have influenced political narratives, such as AI-generated appearances of former Prime Minister Imran Khan, and have been weaponized to defame female leaders, amplifying gendered harassment in conservative contexts (Dawn, 2024). Yet many young users lack the media literacy required to verify content, making them susceptible to misinformation and manipulation.

Research indicates that media literacy enhances individuals' capacity to evaluate digital information critically (Jones-Jang et al., 2021; Hameleers, 2023). However, limited studies explore how this applies within Pakistan or how literacy interacts with cognitive processing pathways outlined in the *Elaboration Likelihood Model* (ELM), which distinguishes between central, analytical evaluation and peripheral, cue-based judgment (Petty & Cacioppo, 1986).

This study aims to examine the influence of media literacy on deepfake awareness and evaluation among Pakistani youth aged 18–35. Specifically, it investigates whether greater literacy fosters critical engagement through central processing and reduces reliance on superficial cues, thereby strengthening users' ability to discern and respond to deceptive media content.

1.1 Research Questions:

i. What is the level of awareness of deepfake technology among digital users aged 18 to 35 in Pakistan?

ii. What are the dominant perceptions of Pakistanis regarding the societal impact of deepfake technology?

iii. To what extent does media literacy influence individuals' ability to critically evaluate deepfake content through a central processing route?

1.2. Research Objectives

i. To assess the level of awareness of deepfake technology among digital users aged 18-35 in Pakistan.

ii. To identify the dominant societal perceptions of deepfake technology in Pakistan.

iii. To explore whether media literacy affects individuals' ability to critically evaluate and recognize deepfake content using the central processing route or the Peripheral route of the Elaboration Likelihood Model.

2. Literature Review

Literature on misinformation and digital deception has expanded considerably in recent years as artificial intelligence transforms the creation and circulation of media content. Scholars have explored how algorithmic technologies and social media ecosystems facilitate the spread of false or manipulated information, undermining public trust and media credibility (Maras & Alexandrou, 2018; Mirsky & Lee, 2021). Within this domain, *deepfakes,* synthetic videos or images generated through machine-learning models such as generative adversarial networks (GANs), represent a particularly disruptive innovation. They challenge traditional notions of authenticity, enabling the fabrication of persuasive yet entirely false narratives. Early studies focused on the technical processes of deepfake generation, while more recent research examines their psychological, social, and political implications, including their influence on perception, credibility, and information processing (Chesney & Citron, 2019; Vaccari & Chadwick, 2020). In Pakistan, the rapid growth of digital media and limited public awareness amplify these challenges, yet empirical investigation remains scarce. Accordingly, this review synthesizes theoretical and empirical contributions related to deepfake creation, dissemination, and impact, highlighting how media literacy and cognitive processing models can inform strategies to counter the spread of synthetic misinformation.

2.1. Understanding Deepfakes

The contemporary information environment is increasingly characterized by misinformation, selective truth-telling, and manipulation, leading scholars to describe this era as a *post-truth society* (Harari, 2018; McIntyre, 2018). Within this context, *deepfakes,* synthetic media generated through artificial intelligence (AI), represent one of the most disruptive and concerning developments in digital communication. Deepfakes are hyper-realistic, yet fabricated images, videos, or audio recordings created using advanced generative models such as Generative Adversarial Networks (GANs). These models operate by training neural networks to imitate authentic patterns of human expression, movement, and speech, thereby blurring the boundary between reality and fabrication (Vaccari & Chadwick, 2020; Mirsky & Lee, 2021).

Deepfake creation has evolved rapidly, with open-source software and machine learning tools reducing both the technical skill and cost required for production. As a result, the capacity to generate highly convincing digital forgeries has become accessible to almost anyone with basic computing knowledge. These advancements have accelerated the proliferation of synthetic content across digital platforms, fundamentally challenging the authenticity and reliability of online information (Diakopoulos & Johnson, 2021).

The social and ethical implications of deepfakes are multifaceted. At the individual level, manipulated videos can lead to serious emotional and reputational harm, including harassment, blackmail, or defamation (Chesney & Citron, 2019). At the societal level, deepfakes have been exploited in political propaganda, financial scams, and misinformation campaigns, threatening public trust and democratic processes (Cruz, 2024). The potential for widespread deception is amplified in societies where digital literacy remains limited and fact-checking infrastructures are underdeveloped.

In Pakistan, the problem is particularly acute. The country has witnessed a surge in internet penetration and social media engagement, with approximately 66.9 million active social media users as of early 2025 (DataReportal, 2025). This rapid digital growth, combined with low media literacy and political polarization, creates an environment where misinformation thrives. Deepfakes have already been weaponized within the country's political sphere; an AI-generated appearance of former Prime Minister Imran Khan in a virtual rally attracted massive online engagement, illustrating the persuasive power of synthetic media (Dawn, 2024). Similarly, female political leaders such as Uzma Bokhari and Maryam Nawaz Sharif have been targeted through manipulated videos that exploit conservative cultural sensitivities, resulting in reputational and emotional harm (Dawn, 2024).

From a technical standpoint, while AI detection tools and forensic analysis can identify inconsistencies in pixelation, lighting, and facial motion, such methods are often inaccessible to general users. Hence, scholars emphasize the necessity for human-centered skills verification, contextual reasoning, and cross-referencing as essential competencies to evaluate media authenticity in everyday digital environments (Qureshi et al., 2024). These abilities are central to the concept of *media literacy*, which encompasses the cognitive and evaluative skills needed to critically assess digital content.

Recent studies emphasize that media literacy not only enables individuals to recognize manipulated media but also strengthens their cognitive resistance to misinformation (Jones-Jang et al., 2021; Moore & Hancock, 2022; Hameleers, 2023). For young audiences, particularly in developing nations, media literacy includes understanding platform dynamics, identifying production cues, and recognizing inconsistencies in narrative coherence (Lao, Hirvonen, & Larsson, 2025). Furthermore, socio-cultural context shapes individuals' ability to detect manipulation factors such as exposure habits, familiarity with public figures, and social media behavior influence how deepfakes are perceived (Turós et al., 2024; Nas & de Kleijn, 2024).

However, despite a growing global body of research, Pakistan remains underrepresented in scholarly discussions on media literacy and deepfake awareness. There is limited empirical understanding of how young Pakistanis, who constitute the majority of the nation's digital population, interpret and respond to manipulated media. This knowledge gap hinders the development of effective educational interventions and public awareness initiatives designed to mitigate the risks associated with synthetic media.

2.2. The Impact of Deepfakes on Perception and Credibility

Deepfakes not only distort factual reality but also profoundly influence audience perception, trust, and credibility in digital communication. Psychological research suggests that even when viewers suspect a video or image may be fabricated, exposure to deepfakes can still induce uncertainty and weaken overall confidence in media authenticity (Vaccari & Chadwick, 2020). This erosion of trust has far-reaching implications, as audiences may begin to doubt legitimate content as well, leading to a generalized cynicism toward news, journalism, and political messaging.

Empirical evidence supports the idea that repeated exposure to manipulated media desensitizes individuals, making them more susceptible to both false information and confirmation bias (Qian et al., 2023; Shukla et al., 2024). Moreover, deepfakes exploit emotional triggers fear, anger, and outrage, which can override rational judgment, thereby facilitating the spread of misinformation through social sharing. Such effects are intensified on algorithm-driven platforms like TikTok, YouTube, and X (formerly Twitter), where engagement rather than accuracy dictates visibility.

In the Pakistani context, the consequences are amplified by low digital literacy and intense political partisanship. Deepfakes have become tools of political warfare, used both to defame opponents and to manipulate public opinion. As Qureshi et al. (2024) note, the spread of AI-generated misinformation in Pakistan not only distorts electoral narratives but also deepens existing divisions along ideological and gender lines. Consequently, citizens, particularly youth, find themselves navigating an information environment where truth is negotiable and credibility is continuously contested.

The *Elaboration Likelihood Model* (ELM) offers a theoretical lens for understanding these perceptual effects. The model posits two routes of information processing: the *central route*, involving thoughtful, analytical evaluation, and the *peripheral route*, relying on superficial cues such as attractiveness, authority, or repetition (Petty & Cacioppo, 1986). Deepfakes often exploit peripheral processing by appealing to emotion and visual realism rather than logic or evidence. However, media literacy can encourage central processing by equipping individuals with the skills to question, verify, and assess the validity of information sources (Hameleers, 2023; Jones-Jang et al., 2021).

Several studies show that targeted media literacy interventions such as fact-checking exercises, source verification training, and exposure to manipulated examples can enhance individuals' capacity to detect falsehoods and reduce susceptibility to persuasive manipulation (Moore & Hancock, 2022; Lao et al., 2025). However, the effectiveness of such interventions varies across cultural contexts. In Pakistan, where formal media literacy education is minimal and digital consumption is high, there is a pressing need for localized, theory-driven initiatives that incorporate psychological models like the ELM to explain how individuals process and respond to synthetic media.

Recent surveys highlight this gap: although public concern over deepfakes is increasing globally, with over 90% of respondents in one UK study expressing anxiety about their dangers

(Sippy et al., 2024), awareness of deepfake manipulation techniques remains low. In Pakistan, young users may acknowledge the existence of fake content but still struggle to evaluate its authenticity due to limited analytical engagement. This pattern suggests that media literacy efforts must not only provide factual knowledge but also cultivate reflective and critical thinking habits that encourage central-route processing.

Ultimately, the credibility crisis caused by deepfakes is not merely a technological challenge but a cognitive and social one. As synthetic media becomes increasingly indistinguishable from reality, the ability to discern truth will rely less on detection algorithms and more on human judgment. Therefore, enhancing media literacy among Pakistani youth is essential not only for individual protection but also for preserving collective trust in digital communication and democratic discourse.

3. Methodology

This chapter outlines the research design, population, sampling strategy, data collection procedures, instruments, analytical methods, and ethical considerations adopted in this study. The methodology was structured to ensure alignment with the study's objective of exploring the relationship between media literacy, deepfake awareness, and information processing among Pakistani youth.

Quantitative research design was employed using a survey-based approach, grounded in the *Elaboration Likelihood Model (ELM)*, which distinguishes between central (analytical) and peripheral (cue-based) information processing routes. This approach allowed for the efficient collection of self-reported data on perceptions, attitudes, and behaviors from a broad sample. The target population comprised Pakistani digital media users aged 18–35, recruited through a non-probability snowball sampling method, resulting in 301 valid responses.

Data were collected through a structured online questionnaire distributed via social media platforms such as WhatsApp, Facebook, and Instagram, ensuring accessibility to active digital users. The instrument consisted of demographic questions and five validated scales: the *Media Literacy Scale* ($\alpha = 0.800$), *Deepfake Awareness Scale* ($\alpha = 0.716$), *Perception of Deepfakes Scale* ($\alpha = 0.658$), *Central Route Processing Scale* ($\alpha = 0.613$), and *Peripheral Route Processing Scale* ($\alpha = 0.623$). Items were rated on a 5-point Likert scale ranging from 1 (Strongly Disagree) to 5

(Strongly Agree). These scales were adapted from existing instruments (Koc & Barut, 2016; Kaur et al., 2023; Petty & Cacioppo, 1986) and tailored to the Pakistani context.

Data analysis was conducted using IBM SPSS, employing descriptive statistics, reliability analysis, Pearson correlation, linear regression, and Mann–Whitney U tests. Ethical principles were strictly followed, ensuring voluntary participation, anonymity, and informed consent. While self-reported and snowball sampling data limit generalizability, the design provided valuable insights into youth engagement with deepfakes in Pakistan's digital landscape.

4. Results

This chapter presents the findings of the study, analyzing data collected from 301 respondents through descriptive and inferential statistical methods. The results are organized according to the three research questions and examine the relationships between media literacy, deepfake awareness, perception, and information processing routes (central and peripheral). All analyses were performed using IBM SPSS Statistics, and reliability, correlation, and regression tests were conducted to ensure robustness.

4.1 Research Question 1:

What is the Level of Awareness of Deepfake Technology among Digital Users Aged 18 to 35 in Pakistan?

The first research question explored the overall level of awareness and familiarity with deepfake technology among Pakistani digital users. Descriptive analyses were used to summarize participants' demographic characteristics, social media usage patterns, and awareness scores.

4.1.1 Demographic Profile of Participants

The sample consisted of 301 respondents aged between 18 and 35 years. The majority of participants were between the ages of 23–27 years (48.5%), followed by those aged 18–22 years (42.5%), and only 9% were between 28–35 years. Male respondents made up 56.8% of the sample, while females accounted for 43.2%. Most participants held undergraduate (46.8%) or graduate (34.9%) degrees.

Table 1: Participant Demographics (N = 301)

Variable	Category	Frequency (n)	Percentage (%)
Age	18–22	128	42.5
	23–27	146	48.5
	28–35	27	9.0
Gender	Male	171	56.8
	Female	130	43.2
Education	Undergraduate	141	46.8
	Graduate	105	34.9
	Postgraduate	39	13.0
	Other	16	5.3

Participants were active social media users: 69.4% reported using social media daily, with 41.2% spending 1–3 hours online per day and 23.3% using it for more than 6 hours daily. This highlights that the sample primarily represents digitally active youth who are regularly exposed to online multimedia content.

4.1.2 Awareness of Deepfake Technology

The Deepfake Awareness Scale exhibited acceptable reliability (Cronbach's $\alpha = .716$). Respondents demonstrated moderate familiarity and recognition of deepfakes, with a mean awareness score of 3.28 (SD = 1.13)

Table 2: Descriptive Statistics of Awareness Items (N = 301)

Item	Mean	SD
Familiarity with deepfakes	2.94	1.43
Confidence in identifying deepfakes	2.83	1.35
General deepfake knowledge	4.02	1.57
Encountered deepfakes online	3.33	1.75

Respondents reported moderate awareness of deepfakes. While most participants had heard of the technology and occasionally encountered manipulated videos, only a minority expressed confidence in detecting such content. This indicates that awareness exists primarily at a conceptual level rather than through practical experience.

A regression analysis revealed that media literacy significantly predicted deepfake awareness, accounting for 13.6% of variance ($R^2 = .136$, $F(1, 299) = 47.17$, $p < .001$). Each unit increase in media literacy corresponded to a 0.43-unit increase in awareness ($\beta = .369$, $p < .001$).

Table 3: Regression Analysis: Media Literacy Predicting Awareness

Predictor	B	SE	β	t	p
Constant	2.09	0.18	—	11.32	< .001
Media Literacy	0.43	0.06	0.369	6.87	< .001

This analysis confirms that individuals with higher media literacy exhibit significantly higher awareness of deepfake technology. In other words, critical evaluation skills and familiarity with online verification behaviors enhance recognition of manipulated content.

4.2 Research Question 2

What Are the Dominant Perceptions of Pakistanis Regarding the Societal Impact of Deepfake Technology?

This section presents findings related to participants' attitudes toward deepfake technology and its perceived societal consequences, including misinformation, privacy violations, and political manipulation.

4.2.1 Perception of Deepfakes

The Perception Scale (α = .658) showed moderate reliability. Respondents scored a mean of 3.08 (SD = 0.89), suggesting a balanced view of deepfakes, acknowledging their influence but not perceiving them as overwhelmingly threatening.

Table 4: Descriptive Statistics of Perception Scale (N = 301)

Item	Mean	SD
Deepfakes cause reconsideration of beliefs	3.09	1.03
Deepfakes influence personal perceptions	3.08	1.03

Respondents demonstrated moderate agreement that exposure to deepfake videos can alter attitudes and beliefs, supporting the notion that synthetic media can influence public perception through emotional or persuasive cues.

4.2.2 Perceived Societal Risks of Deepfakes

Participants identified misinformation as the primary societal risk of deepfakes (54.7%), followed by privacy violations (41.9%) and cybercrime (38.9%). Other concerns included reputation damage (36.5%) and political manipulation (28.2%).

Table 5: Perceived Societal Risks of Deepfakes (N = 301)

Risk Type	Frequency (n)	Percentage (%)
Misinformation	164	54.7
Privacy violation	126	41.9
Cybercrime	117	38.9
Reputation damage	110	36.5
Political manipulation	85	28.2
Not sure	37	12.3

The findings highlight that respondents are most concerned about false information and data misuse, aligning with global studies (Vaccari & Chadwick, 2020; Kumar et al., 2022) that view deepfakes as a critical threat to information integrity and personal security.

4.2.3 Public Concern and Regulation

Participants expressed **moderate concern** about the use of deepfakes in political campaigns: 31.6% remained neutral, while 31.5% reported being somewhat or very concerned.

Table 6: Concern About Deepfakes in Political Campaigns (N = 301)

Level of Concern	Frequency (n)	Percentage (%)
Not concerned at all	63	20.9
Slightly concerned	48	15.9
Neutral	95	31.6
Somewhat concerned	51	16.9
Very concerned	44	14.6

When asked who should manage deepfake content, most participants favored social media platforms (45.8%) and government regulations (35.2%), indicating limited trust in individuals' ability to manage the issue.

Table 7: Preferred Oversight of Deepfake Regulation (N = 301)

Responsible Entity	Frequency (n)	Percentage (%)
Social media companies	138	45.8
Government regulations	106	35.2
Individual users	48	15.9
No regulation needed	9	3.0

Respondents largely favored institutional accountability over individual regulation, revealing a public preference for top-down management of misinformation. However, responses regarding regulation were polarized 22.3% strongly supporting and 22.3% strongly opposing strict regulation, indicating an uncertain public stance.

4.3 Research Question 3

To What Extent Does Media Literacy Influence the Ability to Critically Evaluate Deepfake Content Through the Central Route?

This section examines how media literacy affects individuals' analytical and heuristic processing when evaluating digital content, using both correlation and regression analyses grounded in the **Elaboration Likelihood Model (ELM)**.

4.3.1 Correlation Analysis

Pearson's correlation coefficients were computed to examine relationships between Media Literacy, Awareness, Perception, and the two processing routes (Central and Peripheral).

Table 8: Correlation Matrix Among Main Variables (N = 301)

Variable	1	2	3	4	5
1. Media Literacy	—	.369**	.160**	.254**	.012
2. Awareness		—	.141*	.369	0.00
3. Perception			—	.160	.180**
4. Central Route				—	.18
5. Peripheral Route					—

Note. $p < .05$; $p < .01$ (two-tailed).

Results reveal a moderate positive correlation between media literacy and awareness ($r = .369$, $p < .001$), confirming that higher literacy enhances recognition of deepfakes. A smaller yet significant relationship exists between media literacy and perception ($r = .160$, $p = .005$). Media literacy also correlates positively with central route processing ($r = .254$, $p < .001$), indicating that media-literate users engage in more analytic evaluations when viewing online content.

4.3.2 Regression Analyses

Media literacy significantly predicted awareness, perception, and central route processing, confirming ELM's proposition that cognitive engagement depends on motivation and the ability to process information critically.

Table 9: Regression: Media Literacy Predicting Awareness, Perception, and Processing Routes

Dependent Variable	β	R^2	F (1,299)	p
Awareness	.369	.136	47.17	< .001
Perception	.160	.027	7.82	.005
Central Route	.254	.064	20.58	< .001
Peripheral Route	.150	.022	6.87	.009

Media literacy explained 13.6% of the variance in awareness, 2.7% in perception, and 6.4% in central processing. Notably, even though media literacy slightly increased peripheral reliance, it predominantly enhanced analytic (central) processing, indicating that media-literate individuals are more likely to question, verify, and cross-check online content.

4.3.3 Cue Recognition and Verification Behaviors

Participants most frequently relied on checking comments (54.5%) and searching for the source (39.9%) when verifying video authenticity.

Table 10: Verification Behaviors Used by Participants (N = 301)

Behavior	Selected (n)	Selected (%)
Checking comments	164	54.5
Searching for the original source	120	39.9
Relying on trusted pages	93	30.9
Looking for inconsistencies	81	26.9
Critically analyzing content	57	18.9

Regarding detection cues, audio–lip mismatch (56.1%) and unnatural facial expressions (54.5%) were most commonly identified.

Table 11: Cues for Detecting Deepfakes (N = 301)

Detection Cue	Frequency (n)	Percentage (%)
Inconsistent audio-lip sync	169	56.1
Unnatural facial expressions	164	54.5
Overly smooth visuals	90	29.9
Background distortions	89	29.6
Unrealistic or dramatic tone	68	22.6

Participants relied mainly on central visual cues, such as facial inconsistencies, rather than peripheral elements like popularity metrics. This suggests moderate technical discernment among Pakistani youth, yet an overreliance on heuristic checks (e.g., comments) reveals incomplete critical engagement.

4.4 Summary of Key Findings

Overall, the results indicate that participants possess a moderate awareness of deepfake technology, with media literacy emerging as a strong predictor of both awareness and analytical evaluation. Although most respondents acknowledge deepfakes' societal risks, especially misinformation and privacy violations, their verification behaviors rely more on quick, heuristic cues than deep analytical scrutiny.

The findings align with the Elaboration Likelihood Model, supporting the idea that individuals with higher cognitive involvement and literacy levels engage more critically through central-route processing. However, the persistence of peripheral reliance underscores a gap between knowledge and practical critical evaluation, highlighting the importance of media literacy education and public policy interventions to enhance resilience against manipulated digital content.

5. Discussion

5.1 Research Contributions

This study provides valuable insights into the relationship between media literacy, deepfake awareness, and perception among young digital users in Pakistan, guided by the Elaboration Likelihood Model (ELM). By investigating how individuals process deepfake content through central and peripheral routes, the study extends theoretical understanding and provides practical guidance for developing digital literacy interventions and policy frameworks in a rapidly evolving media environment.

5.1.1 Theoretical Contributions

Theoretically, this research contributes to the growing body of literature on the Elaboration Likelihood Model and its application in digital misinformation contexts. The findings reinforce the model's assertion that individuals engage in both central and peripheral routes of information processing, depending on their motivation and cognitive engagement (Petty & Cacioppo, 1986). The significant positive correlation between media literacy and deepfake awareness ($r = .369$, $p < .001$) demonstrates that individuals with higher analytical skills are more likely to recognize manipulated content, confirming the central-route prediction of the ELM. This relationship underscores the cognitive dimension of media literacy as a determinant of critical evaluation and supports previous research linking media literacy to enhanced detection of misinformation (Jones-Jang et al., 2021; Hameleers et al., 2024).

Furthermore, the moderate relationship between media literacy and perception ($r = .160$, $p = .005$) highlights that awareness extends beyond technical recognition to encompass attitudinal and evaluative dimensions. These results align with Kahne and Bowyer (2019), who argued that media literacy enables individuals to interpret the social and political implications of media content. The study thus contributes theoretically by showing that media literacy not only shapes analytical evaluation but also influences the affective and cognitive perception of deepfake threats.

Another significant theoretical insight concerns the nuanced coexistence of central and peripheral processing routes. The positive correlation between media literacy and central-route reliance ($r = .254$, $p < .001$) confirms that individuals with higher cognitive engagement are more likely to evaluate information analytically. However, the small yet significant association between media literacy and peripheral cue reliance ($\beta = 0.150$, $p = .009$) challenges the assumption that higher literacy automatically diminishes heuristic processing. Instead, this finding supports a dual-processing interpretation in which users blend analytical and heuristic judgments when evaluating authenticity (Sundar, 2020). This dual-route tendency suggests that even informed audiences occasionally depend on surface cues such as emotional tone, popularity metrics, or source familiarity, especially when cognitive effort is limited or the content is visually sophisticated (Metzger & Flanagin, 2015).

By revealing this interaction between cognitive and heuristic processing, the study extends the ELM framework to the context of synthetic media. It demonstrates that digital environments amplify the overlap between central and peripheral cues, where even knowledgeable users employ mixed strategies of scrutiny and intuition. These results echo previous findings that cognitive ability alone may not fully protect individuals from persuasive misinformation (Winter et al., 2016; Luo et al., 2022). Thus, this research provides empirical grounding for adapting the ELM to modern digital contexts, emphasizing cognitive–affective interdependence in deepfake evaluation.

Finally, this study enriches the theoretical understanding of the perceived societal risks of deepfakes. The prominence of misinformation (54.7%) and privacy violations (41.9%) as top concerns mirrors global discourses on the destabilizing effects of synthetic media (Westerlund, 2019; Hameleers et al., 2020). These findings substantiate theoretical claims that digital deception erodes epistemic trust and challenges users' sense of reality. Consequently, this research bridges psychological processing theory with social-level outcomes, illustrating how cognitive responses to deepfakes connect to broader issues of information integrity, political manipulation, and digital ethics (Chesney & Citron, 2019; Diakopoulos & Johnson, 2021).

5.1.2 Practical Contributions

The practical contributions of this study lie in its implications for digital literacy education, social media governance, and public policy. First, the findings highlight the importance of integrating media literacy training into both formal and informal education systems. Since media literacy strongly predicted deepfake awareness and central-route processing, educational initiatives should focus on enhancing users' analytical evaluation skills, particularly their ability to detect visual and auditory inconsistencies in digital media. This aligns with recommendations by Hobbs (2017) and Choi and Lee (2021), who emphasize that comprehensive literacy programs must develop not only technical competence but also reflective and ethical reasoning.

Moreover, the finding that individuals still rely on peripheral cues such as social validation, emotional appeal, and video aesthetics suggests that media literacy programs must incorporate strategies to help users recognize and critically evaluate these influences. Training modules could simulate real-world social media environments where heuristic cues are deliberately manipulated, allowing learners to practice distinguishing between credible and deceptive content (Metzger & Flanagin, 2015; Luo et al., 2022). In this way, education can cultivate both cognitive vigilance and emotional resilience, reducing susceptibility to persuasive manipulations.

From a policy standpoint, this research supports a multi-stakeholder approach involving government regulators, educational institutions, and social media companies. The results indicate that participants overwhelmingly prefer institutional accountability, with 45.8% favoring regulation by social media platforms and 35.2% supporting government involvement. These views align with global calls for stronger digital governance and transparency from platform operators (Napoli, 2019; Gillespie, 2018). Consequently, policymakers should establish frameworks that require platforms to implement robust content authentication and fact-checking systems. Such frameworks must also safeguard user rights and privacy, ensuring that measures against misinformation do not infringe on freedom of expression.

In addition, this study offers insights for technology developers. The persistent influence of peripheral cues suggests a need for interface designs that reduce emphasis on superficial metrics like “likes” and “shares,” which can bias users’ trust judgments. Developers can incorporate authenticity indicators such as verified source labels or AI-generated content disclosures to assist users in making more informed evaluations. Integrating such tools would operationalize the findings of this study into practical solutions for mitigating misinformation risks.

Overall, the practical contribution of this research lies in its comprehensive view of how educational, technological, and policy interventions can collectively enhance resilience against digital deception. Strengthening cognitive literacy, fostering emotional awareness, and promoting institutional accountability can work together to reduce the societal harms of deepfakes while empowering individuals to navigate information critically.

5.2 Conclusion

The study set out to explore the relationship between media literacy, awareness, perception, and information processing routes regarding deepfakes among Pakistani youth. The results present a nuanced understanding of how young users engage with manipulated digital content. Participants displayed moderate levels of media literacy and awareness, indicating that while they recognize the existence and risks of deepfakes, they do not consistently employ analytical strategies to verify authenticity.

Consistent with the ELM, higher media literacy predicted stronger awareness and greater reliance on central-route processing. This demonstrates that education and cognitive engagement enhance the ability to detect technical irregularities such as mismatched audio and unnatural expressions. However, even media-literate individuals were found to depend partially on peripheral cues like emotional tone, social validation, and video aesthetics. This dual-processing behavior underscores the complexity of digital cognition, where rational and affective evaluations coexist.

The study also identified widespread societal concerns about misinformation, privacy, and reputational harm caused by deep fakes. These findings reveal that deepfakes are not merely a technological issue but a broader social challenge threatening public trust and democratic integrity. Participants’ preference for regulation by social media companies and government bodies highlights the growing demand for systemic solutions beyond individual responsibility.

In conclusion, the study contributes to both theoretical understanding and practical application. It extends the ELM framework to the digital misinformation landscape, highlighting the coexistence of central and peripheral processing in media evaluation. Practically, it emphasizes the necessity of integrated educational and policy responses to equip individuals with the cognitive and emotional resources required to navigate manipulated media environments. The findings call for a comprehensive approach that balances critical education, ethical technology design, and effective governance to mitigate the multifaceted risks of deepfakes.

5.3 Limitations and Future Recommendations

While this study makes meaningful contributions, several limitations should be considered when interpreting its findings. First, the research employed a non-probability snowball sampling method, which primarily captured young, educated, and digitally active participants. Although this group represents the most frequent social media users, it limits the generalizability of results to broader populations, including rural communities and individuals with lower digital access. Future studies should adopt stratified or random sampling to ensure more representative demographic coverage and explore differences across age, education, and geographic groups (Guess et al., 2020; Hassan et al., 2022).

Second, the study relied on self-reported data collected through online surveys. Such data are susceptible to biases, particularly social desirability bias and self-perception errors, where respondents may overestimate their critical engagement with media. To address this limitation, future research could incorporate experimental or observational methods to measure actual behavior in real-time digital environments. Eye-tracking, response latency, or biometric analyses could yield more objective insights into cognitive and emotional reactions to deepfake content (Podsakoff et al., 2012).

Third, the Perception Scale demonstrated lower internal consistency ($\alpha = .658$), which may have constrained the depth of attitudinal measurement. Enhancing this construct in future research through additional validated items or triangulation with qualitative methods would strengthen reliability. Qualitative approaches such as focus groups and interviews could also uncover richer insights into the motivations, fears, and contextual factors shaping users' perceptions (Khan & Yousaf, 2023).

Another area for future research is the integration of artificial intelligence tools in identifying and combating deepfakes. Given the rapid advancements in AI-based detection, it would be valuable to assess how users in developing contexts interact with such tools, their trust in automated detection systems, and the ethical implications of algorithmic governance (Raza & Ahmad, 2021). Examining the synergy between human literacy and AI assistance could inform the development of hybrid verification systems tailored to local contexts.

Finally, cross-cultural studies could extend these findings by comparing information processing patterns across different sociocultural environments. Such comparative analyses would reveal whether the coexistence of central and peripheral processing routes is universal or shaped by cultural norms and media systems.

In summary, while this study establishes a foundational understanding of how media literacy influences awareness and perception of deepfakes, future research should strive for methodological diversity, broader representation, and interdisciplinary collaboration. Doing so will refine theoretical models like the ELM, enhance educational and technological interventions, and contribute to a more resilient digital society capable of navigating the challenges of synthetic media.

References

1. Abbasi, N. A., & Huang, D. (2020). *Digital media literacy: Social media use for news consumption among teenagers in Pakistan. Global Media Journal, 18*(35), 1–7.
2. Ali, A., & Qazi, I. A. (2022). *Digital literacy and vulnerability to misinformation: Evidence from Facebook users in Pakistan. Journal of Digital Society.* https://doi.org/10.51685/jqd.2022.025
3. Chesney, B., & Citron, D. (2019). Deepfakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*(6), 1753–1801.
4. Choi, J., & Lee, J. (2021). The role of media literacy in discerning fake news: A cognitive approach. *Journal of Media Psychology, 33*(2), 74–85. https://doi.org/10.1027/1864-1105/a000272
5. Diakopoulos, N., & Johnson, D. (2021). Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society, 23*(7), 2072–2098. https://doi.org/10.1177/1461444820986734
6. Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media.* Yale University Press.
7. Hameleers, M., Brosius, A., & De Vreese, C. (2024). Media literacy and misinformation: How cognitive engagement influences fake news detection. *Communication Research, 51*(3), 389–412. https://doi.org/10.1177/00936502231113450
8. Hameleers, M., Van Der Meer, T. G. L. A., & Dobber, T. (2022). You won't believe what they just said! The effects of political deepfakes embedded as vox populi on social media. *Social Media + Society, 8*(3), 1–10. https://doi.org/10.1177/20563051221116346
9. Hassan, S., Qureshi, M., & Ali, A. (2022). Digital divide and media literacy in Pakistan: Challenges and opportunities. *Pakistan Journal of Social Sciences, 42*(1), 55–70.
10. Hobbs, R. (2017). *Create to learn: Introduction to digital literacy.* John Wiley & Sons.

11. Jeong, S.-H., Cho, H., & Hwang, Y. (2012). Media literacy interventions: A meta-analytic review. *Journal of Communication, 62*(3), 454–472. https://doi.org/10.1111/j.1460-2466.2012.01643.x
12. Jones-Jang, S. M., Mortensen, T., & Liu, J. (2021). Does media literacy help identification of fake news? Information literacy helps, but other literacies don't. *American Behavioral Scientist, 65*(2), 371–388. https://doi.org/10.1177/0002764219869406
13. Kahne, J., & Bowyer, B. (2019). Can media literacy education increase digital engagement in politics? *Learning, Media and Technology, 44*(2), 211–224. https://doi.org/10.1080/17439884.2019.1601108
14. Luo, Y., Zhao, Y., & Sun, X. (2022). Heuristic cues and cognitive engagement: Understanding fake news sharing on social media. *Computers in Human Behavior, 132*, 107237. https://doi.org/10.1016/j.chb.2022.107237
15. Maras, M. H., & Alexandrou, A. (2019). Determining authenticity of video evidence in the age of artificial intelligence and in the wake of deepfake videos. *The International Journal of Evidence & Proof, 23*(3), 255–262. https://doi.org/10.1177/1365712718807226
16. Metzger, M. J., & Flanagin, A. J. (2015). Psychological approaches to credibility assessment online. In *The handbook of the psychology of communication technology* (pp. 445–466). Wiley.
17. Napoli, P. M. (2019). Social media and the public interest: Governance of news platforms in the realm of individual and algorithmic gatekeepers. *Telecommunications Policy, 39*(9), 751–760. https://doi.org/10.1016/j.telpol.2014.12.003
18. Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences, 25*(5), 388–402. https://doi.org/10.1016/j.tics.2021.02.007
19. Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change.* Springer-Verlag.
20. Petty, R. E., Briñol, P., & Priester, J. R. (2009). Mass media attitude change: Implications of the elaboration likelihood model of persuasion. In J. Bryant & M. B. Oliver (Eds.), *Media effects: Advances in theory and research* (3rd ed., pp. 141–180). Routledge.
21. Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879
22. Sundar, S. S. (2020). Rise of machine agency: A framework for studying the psychology of human–AI interaction (HAII). *Journal of Computer-Mediated Communication, 25*(1), 74–88. https://doi.org/10.1093/jcmc/zmz026

23. Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society, 6*(1), 1–13. https://doi.org/10.1177/2056305120903408
24. Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review, 9*(11), 39–52.