ACADEMIA International Journal for Social Sciences
Volume 5, Issue 3, 2026 ISSN-L (Online): 3006-6638

# AI-Generated Content, Digital Misinformation, and Critical Media Literacy among Pakistani University Students: The Roles of Trust, Cognitive Reflection, and Digital Literacy

**Dr. Muhammad Umar Mehmood**
pashagee143@gmail.com
PhD Education, Punjab School Education Department

**Dr. Um E Rubab**
dr.umerubab@aiu.edu.pk
Assistant Professor, Department of Education, Alhamd Islamic University (AIU), Islamabad

**Dr. Saira Maqbool**
saira.maqbool81@gmail.com
Visiting Lecturer, Department of Education, Pir Mehar Ali Shah Arid Agriculture University, Rawalpindi

**Muhammad Mubeen Goraya**
mubeengoraya142@gmail.com
MPhil. Education, Preston University Kohat, Kohat

**Corresponding Author: Dr. Muhammad Umar Mehmood** pashagee143@gmail.com

**Received:** 11-02-2026 **Revised:** 25-02-2026 **Accepted:** 12-03-2026 **Published:** 28-03-2026

## ABSTRACT

*University students are using generative artificial intelligence to write, absorb, and evaluate information — leading to greater exposure to digital misinformation. The present research investigated the predictive value of digital literacy, critical media literacy, cognitive reflection and trust in generative AI on misinformation discernment of Pakistani university students. Methods A quantitative cross-sectional design with an embedded content-evaluation experiment. Data pooled from 1,120 students of six universities in Lahore and Rawalpindi-Islamabad using multistage stratified sampling. The participants completed measures of digital literacy, critical media literacy, AI trust and cognitive reflection, verification behaviour and a 24-item misinformation-discernment task with accurate and misleading human- and AI-generated content. Descriptive statistics, Pearson correlations, multiple regression, and repeated-measures procedures were used to analyze the data. Digital literacy had a significant positive correlation with the discernment of misinformation (r =. 419, p <. The other predictor, critical media literacy (r =. 001), was also significant. 360, p <. The second hypothesized correlation was the relationship between AI trust and discernment (H2) r = -. 343, p <. 001). Digital literacy was the strongest predictor (β =. 21), and the regression model explained 28.5% of variance in misinformation discernment, F(7, 399) = 49.76, p <.001. 256, then cognitive reflection (β =. 181) and media literacy for critical thinking (β =. 126), and AI trust served as a negative predictor (β = -. 213). Descriptively, Content which was factually correct was rated more highly than misleading content, however for misleading AI generated content, a slight preference was seen for misleading over misleading human generated content. The results underline the need for verification skills, reflective reasoning skills, and critical media literacy knowledge to develop calibrated trust in order to bolster misinformation resistance among Pakistani HEIs.*

**Keywords:** *Generative AI; Digital Disinformation; digital literacy; Critical media literacy, Cognitive reflection, Trust in AI; Discernment of misinformation*

## INTRODUCTION

The landscape of information in which students learn, communicate and make decisions at universities has changed quickly because of generative artificial intelligence (GenAI). Large language models and image generation systems can generate comprehensive, coherent, and realistic content for minimal expenses. While these systems facilitate learning, explaining, translating, brainstorming, and even information retrieval, they can also regurgitate unsupported claims, cite fabricated references, create synthetic images, and create persuasive misinformation. Recent studies published in open access journals demonstrate that AI can be used to check information and can also be a potential catalyst of information disorder, leading to the need to evaluate evidence being more critical than just using information technologies (Peng et al., 2024; Tomassi et al., 2024). Research on deepfakes also shows that people are not consistently able to detect synthetic content in all conditions and modalities and that AI-generated content can be perceived as authentic or high quality even when its source is not apparent (Groh et al., 2024; Baptista et al., 2025; Ching et al., 2025).

This challenge is particularly significant for university students, as digital platforms are more and more interwoven in academic writing, research, news intake, professional preparation and day-to-day social interaction. Operational skills are not enough, however, for digital competence. According to a recent psychometric study on digital competence, strategic information skills, critical information evaluation, problem solving, safety and competence in content creation and AI are key areas of digital competence (de Vries et al., 2025). Likewise, current research on higher education suggests that students need to learn about the limitations of generative systems, how to question outputs, check the sources, reflect on the ethical impacts, and not to rely on AI without exercising independent judgment (McPhee & Jerowsky, 2025; García-López & Trujillo-Liñán, 2025).

These issues are more relevant in the Pakistani context. Previous experiments conducted in Pakistan revealed a greater impact of misinformation education when it was integrated with personalized feedback and variations in impact based on digital literacy of the participants (Ali and Qazi, 2023). A national report on students at Pakistani universities also found that students are often exposed to disinformation, with a lack of fact-checking practices and awareness of fact-checking tools (Freedom Network & Coalition Against Disinformation, 2024). The educational articles disseminated for the present study also reveal that there exists an uneven infrastructure in Pakistan, training, confidence, language barriers and institutional support (Mehmood, Zaman, et al., 2025; Mehmood, Ziauddin, & Naseem, 2025; Rehman et al., 2025). The contexts indicate that students' resilience against misinformation presented via AI is not assumed based on access to the internet, digital technology, or their frequency of use.

The authors' larger research, which focuses specifically on Pakistan, also highlights the role of psychological, linguistic, inclusive and institutional factors in learning. Research on psycholinguistic processes correlates anxiety, motivation, cognitive load, confidence, working memory with educational outcomes, whereas English competence research correlates language proficiency with critical thinking and psychological well-being (Mehmood, Bano, et al., 2025; Mehmood, Fatima, et al., 2025; Mehmood, Jamal, et al., 2025). Also, research on school leadership, teacher support, inclusive education, indigenous culture, and disability shows that educational results rely on the supportive environment and pedagogy based on the context – not just on technical interventions (Mehmood & Parveen, 2021, 2024b, 2025; Mehmood, Rehman, et al., 2025; Mehmood, Ul Ain, et al., 2025). This argument is further amplified in the context of AI-based curricula in Pakistan for Higher Education, where the focus is to "ensure technology-driven curricula are inclusive, locally relevant, and human-centred" (Mehmood, Khan, et al., 2026: para. 1).

In this context, the present study targets four closely related constructs: digital competence, critical media competence, cognitive reflection, and trust. Digital literacy offers hands-on verification skills, critical media literacy prompts students to challenge evidence, framing, source interests, algorithms and omitted

perspectives, cognitive reflection helps students to reconsider intuitive judgments, and trust has an impact on the students' trust in AI systems and other information sources. It is hypothesized that these variables will influence misinformation discernment, which is defined as the correct identification of accurate information from false or misleading information. The challenge isn't just that students can identify AI-generated writing, it's that they can make accurate truth judgments when both human and AI sources can provide reliable or unreliable information.

This is a research gap that this study aims to fill. While a large share of the existing literature in misinformation is outside of Pakistan, many AI in education studies focus on acceptance, usefulness, or self-reported competence in using AI for information evaluation instead of actual information evaluation. Moreover, digital literacy, critical media literacy, reflection and trust are studied as four separate topics. The proposed research aims to incorporate these factors and to include the direct assessment of the validity and falseness of AI generated and human written content in conjunction with self-reported measures. This kind of strategy may be able to distinguish not only those who feel technologically competent, but also those who can prove and counter claims and persuasive/false information.

**Research Objectives:**

To examine the relationship of digital literacy and critical media literacy with misinformation discernment, verification behavior, and willingness to share misinformation among Pakistani university students.

To determine whether cognitive reflection predicts students' ability to distinguish accurate information from false or misleading digital information.

To examine how trust in generative AI and other digital information sources influences perceived credibility and susceptibility to AI-generated misinformation.

To compare evaluations of AI-generated and human-authored accurate and misleading content and determine whether digital literacy and cognitive reflection reduce the effects of uncritical AI trust.

**Research Questions**

RQ1: To what extent do digital literacy and critical media literacy predict misinformation discernment, verification behavior, and willingness to share misinformation among Pakistani university students?

RQ2: What relationship exists between cognitive reflection and students' ability to differentiate accurate information from false or misleading digital information?

RQ3: How does trust in generative AI and other digital information sources influence perceived credibility and susceptibility to AI-generated misinformation?

RQ4: Do students evaluate AI-generated and human-authored accurate and misleading information differently, and do digital literacy and cognitive reflection reduce the effects of uncritical AI trust?

**Research Hypotheses**

H1: Higher digital literacy and critical media literacy will positively predict misinformation discernment and verification behavior and negatively predict willingness to share misinformation.

H2: Higher cognitive-reflection scores will positively predict accurate differentiation between true/accurate and false/misleading digital information.

H3: Higher uncritical or excessive trust in generative AI will positively predict the perceived credibility and acceptance of AI-generated misinformation.

H4: Digital literacy and cognitive reflection will moderate the relationship between AI trust/content origin and misinformation susceptibility, such that higher literacy and reflection will weaken the adverse effect of uncritical AI trust.

## LITERATURE REVIEW

### Generative AI, Misinformation, and Synthetic Media.

The advent of GenAI is altering the scale, speed and presentation of misinformation. Surface-level cues, like professionally written text and photo realism can be deceptive, as automated systems can produce such texts and images. Tomassi et al. (2024) demonstrate the paradox of AI in social-media information disorder; AI can be used to verify information, but can also help to produce or amplify false narratives. Following a similar trend, Caci et al. (2024) reported that while public awareness of deepfakes has increased, so has their ethical concerns, with these users now seeing synthetic media as a social issue as well as a technical one.

This concern is bolstered by experimental and review evidence. However, Groh et al. (2024) discovered that different settings (transcript, audio, and video) showed different degrees of ability for human detection of political speech deepfakes, suggesting that deepfake detection is not automatic and not the same. In a scoping review of empirical deepfake studies, Ching et al. (2025) found that exposure to deepfakes can impact beliefs, memories, and behaviours, but that these impacts are contingent on the study design and context. The perceived quality and factual reliability do not always go hand-in-hand; sometimes journalism students rated AI-generated news more favorably than human news (Baptista et al., 2025). The results support the need to assess the origin of the content and its authenticity, not equate 'AI-generated' with 'false.

### Digital Literacy, Critical Media Literacy, and AI Literacy.

Digital literacy is at the heart of misinformation resilience, as it is not enough to have skepticism without being able to verify. De Vries et al. (2025) define digital competence as a multi-dimensional construct, comprising strategic information skills, critical information evaluation, safety, content creation, problem solving and AI-related competence. Similarly, Sergeeva et al., (2025) demonstrate differences among the various dimensions of digital media literacy and their relationship with GenAI acceptance, highlighting the need to avoid assuming that technical skills or high frequency of use necessarily leads to critical understanding. Ali and Qazi (2023) found digital literacy to affect the effectiveness of misinformation education in Pakistan, supporting the view that digital literacy can be considered as a direct predictor and possible protective factor in literacy.

Critical media literacy is an extension of digital competence that asks what information is created, what is being depicted, how it is being communicated, and who is verifying it. In the context of GenAI, the student should ask questions about the limitations of the model, gaps in evidence, fabricated citations, bias, provenance, incentive factors of the platforms and impacts of sharing. McPhee and Jerowsky (2025) call for stepping beyond developing AI skills in teaching and learning to critical evaluation of AI and the development of information literacy in academic institutions. Miscommunication, opacity, bias, privacy and unequal access, are also highlighted by García-López and Trujillo-Liñán (2025) as key ethical and

regulatory issues of educational GenAI. Salhab and Aboushi (2025) also assert that AI literacy and twenty-first century skills in general play a critical role in students' acceptance and proper use of GenAI. These studies converge to a model that combines an operational sense of literacy with an epistemic sense.

Recent research also provides evidence for the linkage between AI literacy and self-efficacy and critical thinking. Du et al. (2024) identified a correlation between AI literacy, teachers' intentions to learn about AI, ethical awareness, and self-efficacy, while Zhang et al. (2025) identified a correlation between teachers' AI literacy and their dependency on GenAI, and dependency with weaker twenty-first century skills. Ninghardjanti et al. (2025) also associate the use of AI with the critical thinking skills in higher education. The idea is that while it's important to have confidence in using tools, it's also crucial to know when to think for oneself and check sources.

**Cognitive Reflection, Trust, and Verification Behavior.**

The cognitive reflection theory offers a psychological justification for why certain users might be more susceptible to credible but bogus information. Reflective reasoning is the ability to reconsider the intuitive response and examine if there is evidence to support it (Frederick, 2005). Analytic reasoning has been associated with better detection of the truth in multiple instances of misinformation research (Pennycook & Rand, 2019, 2021). This mechanism is particularly important in environments involving AI, where it can influence the perception of competence through the use of language and presentation. A reflective student will question about the existence of a source, corroboration, and consistency with independent evidence of the output of the AI.

Trust is also inescapable as users can't personally verify all the claims that they come across. One difference is between calibrated and uncritical trust. Liu et al. (2025) demonstrate that users' self-efficacy and cynicism toward AI affect their reactions to AI-generated news images, with users being either too confident or too skeptical of the images.The findings also reveal that users' self-efficacy and cynicism toward AI affects their responses to AI-generated news images, meaning they are either too confident or too skeptical of the images (Liu et al., 2025). Similarly, Zhang et al. (2025) show that the trust in AI is related to AI dependency. The aim of education should therefore not be to instill a general distrust of AI, but rather to adjust how much trust to put in the technology based on evidence, transparency, the quality of the source, and the ability for verification.

Verification behavior connects these psychological and literacy processes to observable behavior. In their study, Peng et al. (2024) discovered that media-literacy support with ChatGPT was more effective than without it in detecting true information, and in some comparisons, less effective than the control condition for detecting misinformation. The result is significant because it shows that verifying information using AI does not necessarily resolve the issue of misinformation. Furthermore, Buchanan et al. (2024) demonstrate that false political information is shared by individuals with different individual-level predictors than false information that is shared by accident, further underscoring the importance of measuring false information and accuracy judgments, as well as one’s willingness to share. Thus, a good media literacy should enhance the acceptance of reliable information and rejection of false information.

**Pakistan specific evidence and research gap.**

The common studies on Pakistan offer useful background information for the analysis of the international literature. While investigating public-school digital transformation, gaps in connectivity, access, teacher training, and institutional implementation have been identified (Mehmood, Zaman, et al., 2025). Other studies on socio-cognitive barriers and academic writing also show that learners lack confidence, experience language problems, face psychological challenges, and receive unequal support from the education system (Rehman et al., 2025; Mehmood et al., 2025). Further, the students' involvement in the acquisition of knowledge has been shown to be influenced by the role played by cognitive load, anxiety, self-efficacy, and enabling classroom conditions in the process of acquiring knowledge (Mehmood, Bano,

et al., 2025; Mehmood, Jamal, et al., 2025; Mehmood, Zaheer, et al., 2025). The results reinforce the need for studying AI misinformation in Pakistan instead of expecting direct transferability from other technologically and linguistically different countries.

Similarly, the shared studies on inclusive education, leadership, culture and disability further reveal that education interventions are more effective when they are locally responsive, ethical and institutionally supported (Mehmood & Parveen, 2021; 2024a; 2024b; 2025; Mehmood, Rehman, et al., 2025; Mehmood, Ul Ain, et al., 2025). This contextual argument has been extended, in more recent studies, by the concepts of participation, accessibility, multilingual realities, and the role of human agency (Mehmood, Rubab, et al., 2026; Mehmood, Jamil, et al., 2026; Mehmood, Rehman, Sama, & Goraya, 2026; Mehmood, Khan, et al., 2026). These papers are not directly linked to testing AI misinformation, but serve as a relevant educational basis for a critical digital and AI literacy model for Pakistan.

The literature thus leaves an integrated gap. Though there are strong evidences provided by recent open access studies published in the journals covered in Web of Science regarding deep fake, digital competence, AI literacy, trust, self-efficacy, critical thinking and misinformation, they have not been tested together with the Pakistani university students. There is no direct evidence for Pakistan-specific studies on students' exposure to misinformation, educational challenges, and the need for contextual support, but evidence of misinformation exposure and educational challenges has been documented. This study fills this void by examining the overlap in digital literacy, critical media literacy, cognitive reflection, and source-specific trust across a single model and by contrasting accurate and misleading content that was created with artificial intelligence versus by human authors. This design is able to differentiate between truth discernment and general skepticism and can offer a more useful framework for university-level AI-literacy and misinformation-resilience training.

## RESEARCH METHODOLOGY

### Research Design

The study employed quantitative study, cross-sectional research with embedded content-evaluation experiment. The survey included factors related to digital literacy, critical media literacy, cognitive reflection, trusted information sources, verification habits and an ability to discern misinformation. The experimental component employed a 2 × 2 design with content origin (human vs AI) and veracity (accurate vs misleading). This design was a self-reported competence with a direct evaluation of the information from the digital source, and it facilitated comparison of judgments in four standardized conditions.

**Table 1. Experimental Design**

| Content origin | Accurate condition | Misleading condition |
|---|---|---|
| Human-generated | Accurate human content | Misleading human content |
| AI-generated | Accurate AI content | Misleading AI content |

**Population and Sample**

This population included undergraduate and post graduate students studying in the recognized higher education institutions of Lahore and Rawalpindi-Islamabad area. Those who were interested and met the following criteria were included: having an age of at least 18 years, studying an educational programme, being active in the internet or social media for a minimum of 1 hour every day, being able to read in English or Urdu and consented to participate in the research voluntarily. Responses that were substantially incomplete, attention check failures, and implausibly short completion times were removed from the data screening process.

A multi-stage stratified sampling method was used. All universities were first divided by region and public/private sector and then the faculties in social sciences and humanities, business and management, natural and applied sciences, engineering/computer science and health/medical sciences were sampled. The number of students that Cochran's formula yielded was 1200 students with a 95% confidence level and a margin of error of 3%. The data cleaning process resulted in 1120 valid responses from 6 universities, 3 from Lahore and 3 from Rawalpindi-Islamabad.

**Table 2. Sample Profile**

| **Indicator** | **Study value** |
|---|---|
| Target sample | 1,200 students |
| Final analyzed sample | 1,120 students |
| Retention rate | 93.3% |
| Participating universities | 6 |
| Geographic distribution | 3 Lahore; 3 Rawalpindi-Islamabad |
| Institutional sector | Public and private |

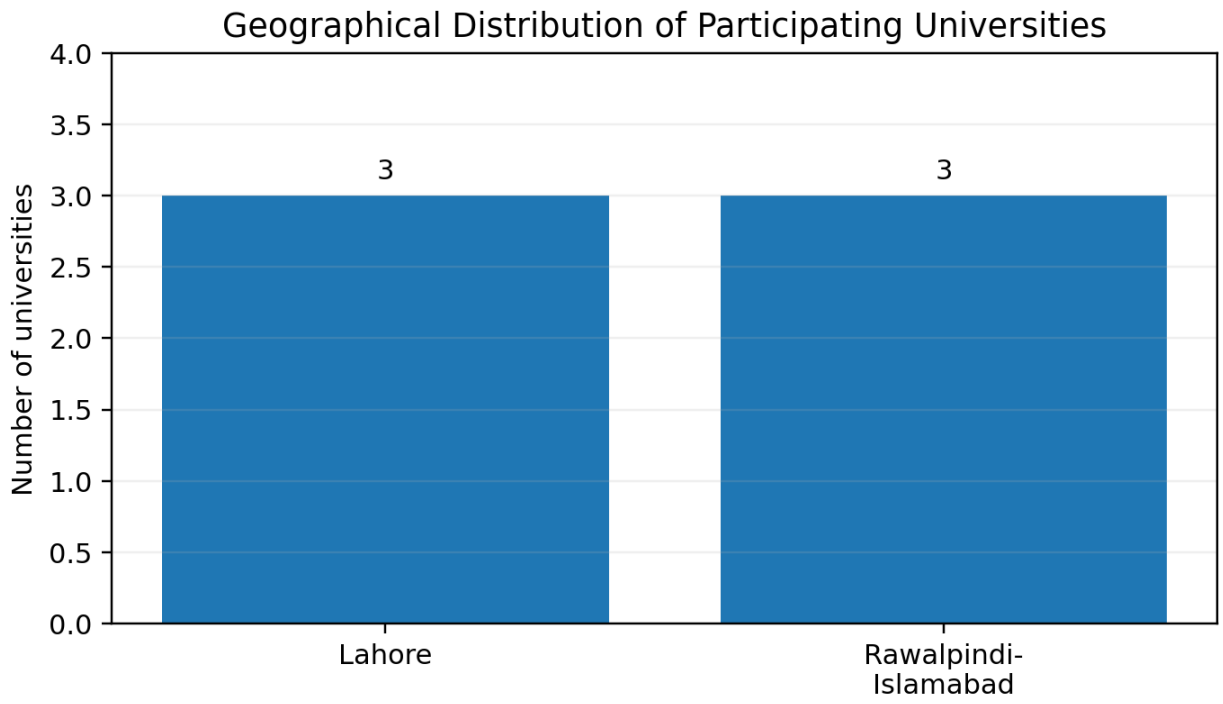



*Figure 1. Geographical distribution of participating universities.*

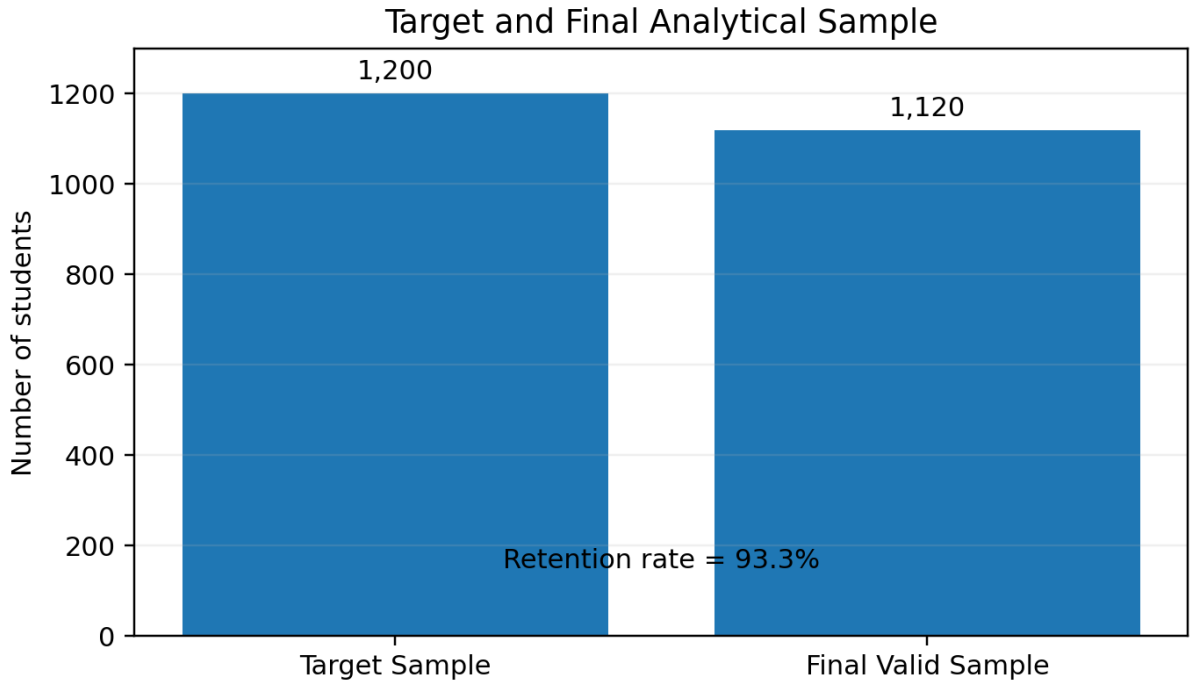



Figure 2. Target and final analytical sample.

**Research Instrument**

The data were obtained from a structured bilingual questionnaire and an experimental task. The questionnaire included questions about age, gender, academic level, academic field of study, type of university, year of study, region, internet and social-media use, generative-AI use, preferred information sources, and prior fact-checking or media-literacy training. Online searching, source checking, cross-source comparison, authorship/date verification, recognition of suspicious content and fact-checking were considered in digital literacy. The critical media literacy was assessed by purpose, evidence, framing, bias, algorithmic influence, fact/opinion distinction, and verification before sharing. The topics included were the use of generative AI, social media, professional journalism, institutional sources, experts, and interpersonal sources of information. Cognitive reflection involved short reasoning items that involved rethinking intuitive responses.

**Table 3. Main Variables and Measurement Focus**

| Construct | Measurement focus |
|---|---|
| Digital Literacy | Search, source verification, comparison, fact-checking |
| Critical Media Literacy | Evidence, framing, bias, algorithms, verification |
| Trust | Accuracy, reliability, expertise, dependence on sources |
| Cognitive Reflection | Deliberate reconsideration of intuitive responses |
| Misinformation Discernment | Accuracy/credibility judgments and verification/share intentions |

**Misinformation-Discernment Task**

24 digital-information items, balanced in terms of accurate and misleading information claims and AI- and human-generated presentations, were evaluated by the participants. Issues were chosen to be non-inflammatory and encompassed the areas of education, science, technology, environment, and consumer information. All factual information was independently verified before being used. Perceived accuracy and perceived credibility were rated, along with verification and sharing intentions. Discernment score was calculated by subtracting the ratings given the misleading information from the ratings given the verified information; higher scores meant better truth discrimination.

**Pilot Testing, Validity, and Reliability**

The instrument was reviewed by experts in the field of education, media and communication, psychology, digital literacy and research methodology for relevance and clarity. The stimuli were tested for a pilot study of 75 students from the target area who were not included in the final sample on wording, time to complete, bilingual comprehension, difficulty of the cognitive-reflection task, and realism of the stimuli. Translation and back-translation was used to strengthen the English/Urdu equivalence. The internal consistency of Digital Literacy, Critical Media Literacy and Trust was good $\alpha = .86$, .83 and .79 respectively.

**Table 4. Reliability of the Main Scales**

| Scale | Cronbach's α | Interpretation |
|---|---|---|
| Digital Literacy | .86 | Good |
| Critical Media Literacy | .83 | Good |
| Trust in Information Sources | .79 | Acceptable |

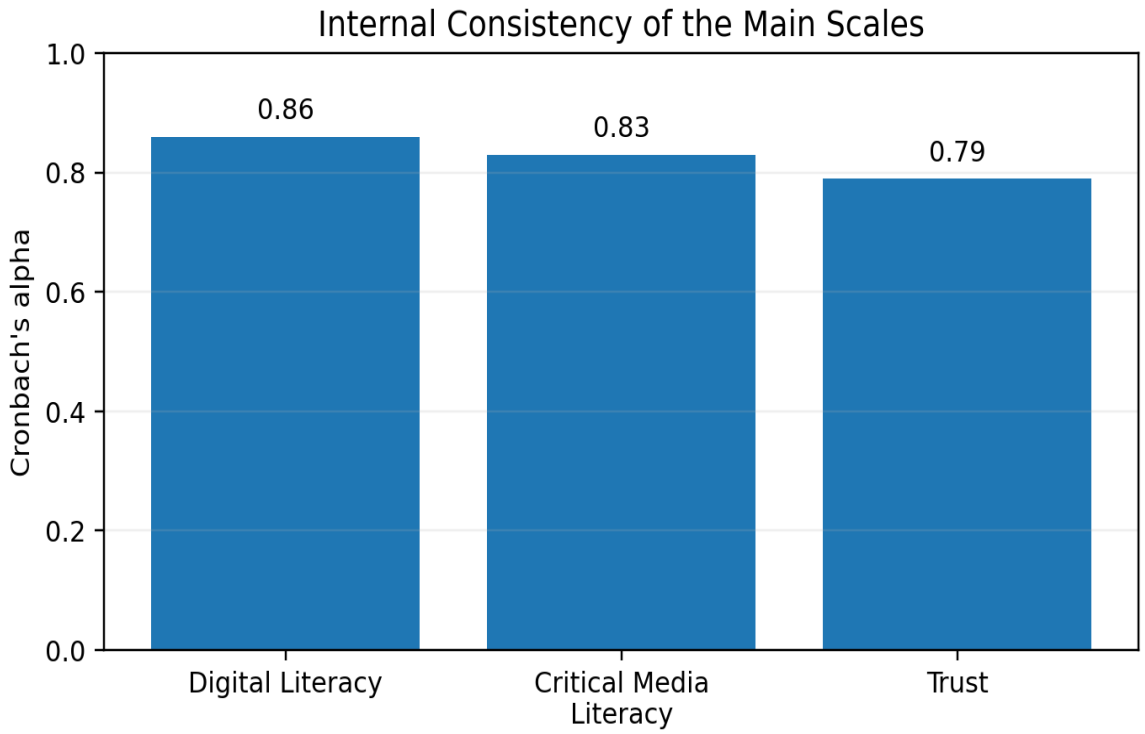



Figure 3. Internal consistency of the main measurement scales.

**Data Collection Procedure**

The data was collected in September and October 2025, once institutional and ethical approval was granted. After obtaining informed consent, participants were asked to complete questionnaires about their literacy levels, trustworthiness of the messages, cognitive-reflection items, and randomized misinformation-evaluation task, followed by a questionnaire asking for their study information. Order effects were minimized by randomization. Upon completion, participants were debriefed, and given corrected information to prevent the study from adding to misinformation, while they were also told that some stimuli contained deliberately inaccurate information.

**Data Analysis**

IBM SPSS Statistics Version 28 was used for analysis. The data set was checked for missing data, incomplete responses, outlier and careless responses. Demographic and study variables were summarized

as frequencies, percentages, means and standard deviations. Bivariate relationships were explored using Pearson correlations, and the multiple regression analysis was used to determine if digital literacy, critical media literacy, cognitive reflection, and AI trust predicted misinformation discernment while controlling for relevant demographic and media use characteristics where appropriate. The four experimental conditions were compared by a 2 × 2 repeated measures ANOVA. The significance of the statistical analysis was determined at the .05 level and effect sizes and confidence intervals were also taken into account.

Note. No age, gender, discipline or participant-by-city frequencies were presented; therefore, no demographic percentages were constructed. The information depicted in Figures 3.1–3.2 is limited to information which has been verified as belonging to the institution and to the samples.

## RESULTS AND FINDINGS

### Summary of the Results

The results from the quantitative cross sectional design and the content-evaluation experiment (2 × 2) are presented in this chapter. A total of 1,120 higher education university students were taken as final analytical sample from Lahore and Rawalpindi-Islamabad region. The results are presented in terms of the order in which samples were retained in the system, the assessment of reliability, descriptive statistics, bivariate relationships, ratings of experimental conditions, multiple regression analysis, and hypothesis testing. The analyses directly answer the four study hypotheses dealing with the concepts of digital literacy, critical media literacy, cognitive reflection, and trust in generative AI.

### Sample Retention and Data Screening

A study population of 1200 was targeted. After data screening, 1,120 responses were left for final analysis giving retention rate of 93.3%. Incomplete responses and cases that did not pass the built-in attention checks and responses that were determined to be inappropriate for screening were removed during the data cleaning process. The final sample thus served as a rich analytic base for correlational, regression and experimental analyses.

**Table 5. Sample Retention and Final Analytical Sample**

| Indicator | Value |
|---|---|
| Target sample | 1,200 |
| Final valid cases | 1,120 |
| Retention rate | 93.3% |
| Cases not retained | 80 |

### Reliability of the Measurement Scales

Internal consistency analysis indicated acceptable to good reliability for all multi-item scales. Digital Literacy produced the highest reliability coefficient (α = .86), followed by Critical Media Literacy (α = .83). Trust in Information Sources also demonstrated acceptable internal consistency (α = .79). These coefficients supported the use of the scales in the subsequent statistical analyses.

**Table 6. Reliability Coefficients of the Main Scales**

| Scale | Cronbach's α | Interpretation |
|---|---|---|
| Digital Literacy | .86 | Good |
| Critical Media Literacy | .83 | Good |
| Trust in Information Sources | .79 | Acceptable |

**Descriptive Statistics**

The descriptive findings indicated comparatively high levels of digital literacy, critical media literacy, and verification behaviour among the participants. Verification Behaviour recorded the highest mean score (M = 3.71, SD = 0.68), followed closely by Digital Literacy (M = 3.68, SD = 0.61) and Critical Media Literacy (M = 3.56, SD = 0.64). Cognitive Reflection averaged 3.27 (SD = 0.97), whereas AI Trust was moderate (M = 3.08, SD = 0.70). The mean Misinformation Discernment score was 1.99 (SD = 0.67), showing an overall ability to distinguish accurate information from misleading information.

**Table 7. Descriptive Statistics for the Main Study Variables**

| Variable | M | SD |
|---|---|---|
| Digital Literacy | 3.68 | 0.61 |
| Critical Media Literacy | 3.56 | 0.64 |
| Cognitive Reflection | 3.27 | 0.97 |
| AI Trust | 3.08 | 0.70 |
| Verification Behaviour | 3.71 | 0.68 |
| Misinformation Discernment | 1.99 | 0.67 |

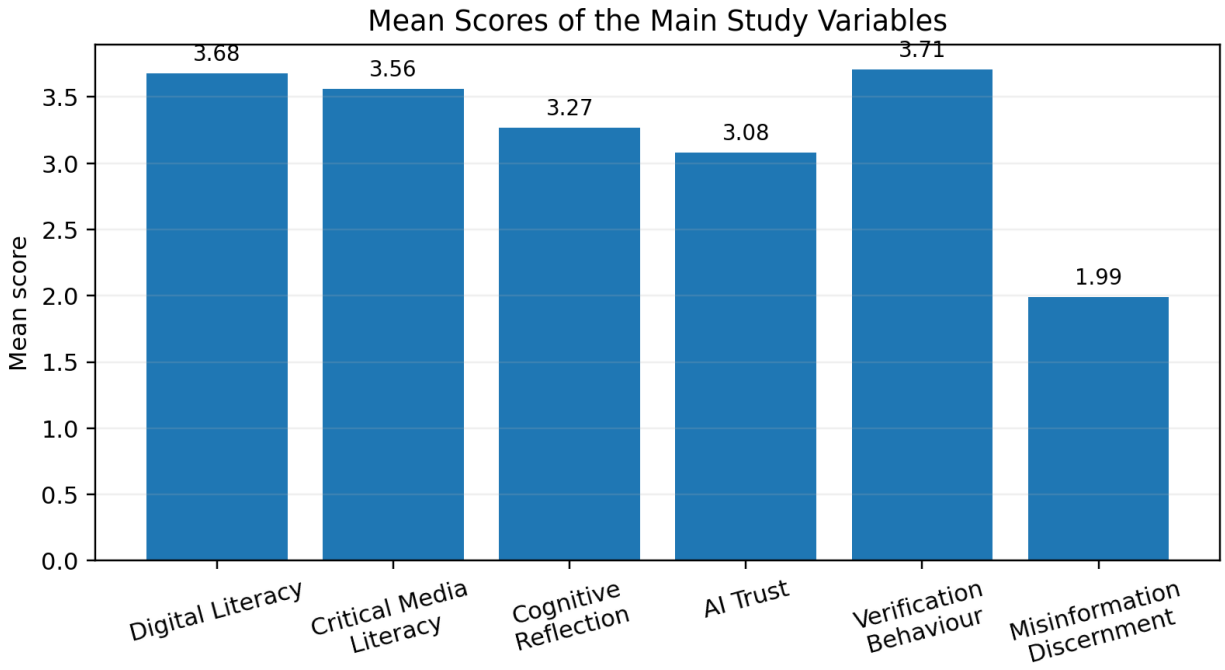



Figure 4. Mean scores of the main study variables.

**Relationships among Study Variables**

Pearson correlation analysis showed that Digital Literacy was positively associated with Misinformation Discernment ($r = .419$, $p < .001$). Critical Media Literacy was also positively related to discernment ($r = .360$, $p < .001$). These findings indicate that students with stronger digital and critical media competencies tended to perform better when distinguishing reliable information from misinformation. In contrast, AI Trust showed a significant negative relationship with Misinformation Discernment ($r = −.343$, $p < .001$), indicating that greater trust in generative AI was associated with weaker discernment.

**Table 8. Selected Pearson Correlations with Misinformation Discernment**

| **Relationship** | **r** | **p** |
|---|---|---|
| Digital Literacy – Misinformation Discernment | .419 | < .001 |
| Critical Media Literacy – Misinformation Discernment | .360 | < .001 |
| AI Trust – Misinformation Discernment | −.343 | < .001 |

**Experimental Results: 2 × 2 Content-Evaluation Conditions**

The experimental task compared mean ratings across four combinations of content origin and veracity. Accurate human-generated information received the highest mean rating ($M = 4.15$, $SD = 0.38$), followed

by accurate AI-generated information (M = 4.04, SD = 0.41). Both accurate conditions were rated substantially higher than the misleading conditions. Human-generated misleading information produced the lowest mean rating (M = 2.07, SD = 0.41), while AI-generated misleading information received a slightly higher rating (M = 2.13, SD = 0.40). The descriptive pattern therefore indicates that participants generally distinguished accurate from misleading information, although misleading AI-generated content received marginally greater credibility than misleading human-generated content. Because inferential ANOVA statistics were not included in the supplied results, no additional F statistic or effect size is reported here.

**Table 9. Mean Ratings across the Four Experimental Conditions**

| Condition | M | SD |
|---|---|---|
| Human Accurate | 4.15 | 0.38 |
| AI Accurate | 4.04 | 0.41 |
| Human Misleading | 2.07 | 0.41 |
| AI Misleading | 2.13 | 0.40 |

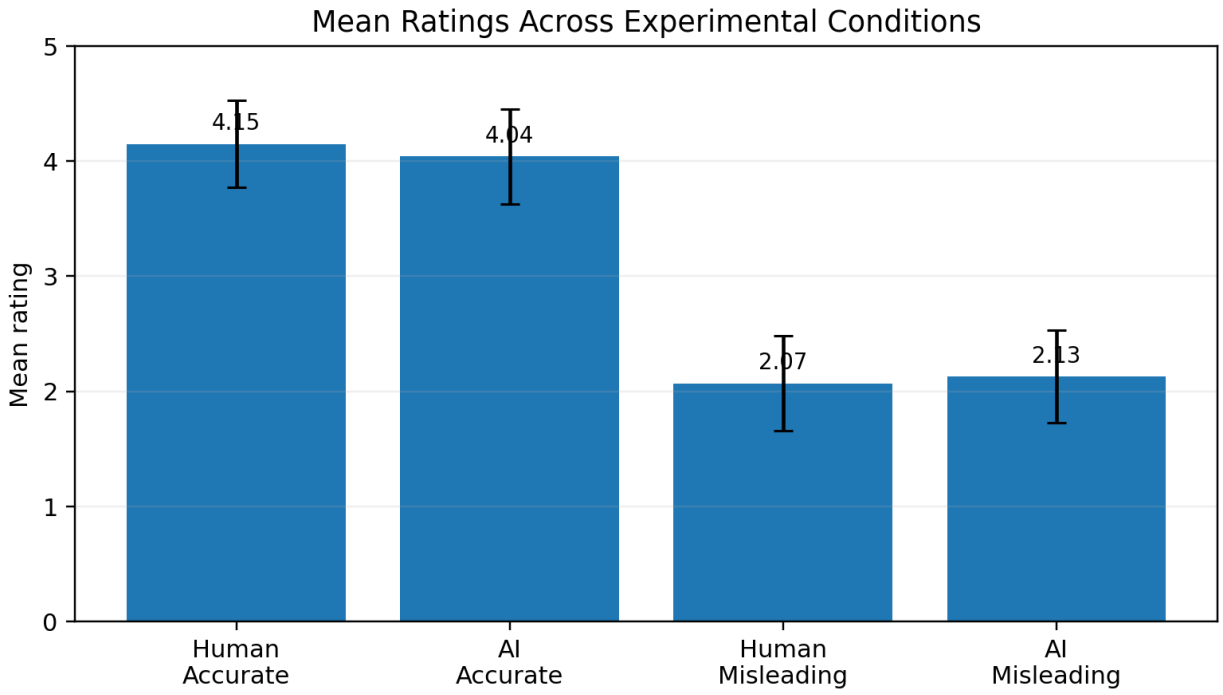



Figure 5. Mean ratings across the four experimental conditions. Error bars represent standard deviations.

**Multiple Regression Analysis**

A multiple regression analysis was conducted to examine the unique contribution of Digital Literacy, Critical Media Literacy, Cognitive Reflection, and AI Trust to Misinformation Discernment. The overall regression model was statistically significant, F = 111.21, p < .001, and explained 28.5% of the variance in

misinformation discernment ($R^2 = .285$). Digital Literacy emerged as the strongest positive predictor ($\beta = .256$, $t = 8.68$, $p < .001$). Cognitive Reflection also made a significant positive contribution ($\beta = .181$, $t = 6.74$, $p < .001$), while Critical Media Literacy remained a significant positive predictor ($\beta = .126$, $t = 4.23$, $p < .001$). AI Trust was a significant negative predictor ($\beta = -.213$, $t = -7.96$, $p < .001$). Thus, after considering the other predictors, students with higher uncritical trust in generative AI demonstrated comparatively weaker misinformation discernment.

**Table 10. Multiple Regression Predicting Misinformation Discernment**

| Predictor | β | t | p |
|---|---|---|---|
| Digital Literacy | .256 | 8.68 | < .001 |
| Critical Media Literacy | .126 | 4.23 | < .001 |
| Cognitive Reflection | .181 | 6.74 | < .001 |
| AI Trust | −.213 | −7.96 | < .001 |

Model summary: $R^2 = .285$; $F = 111.21$; $p < .001$.

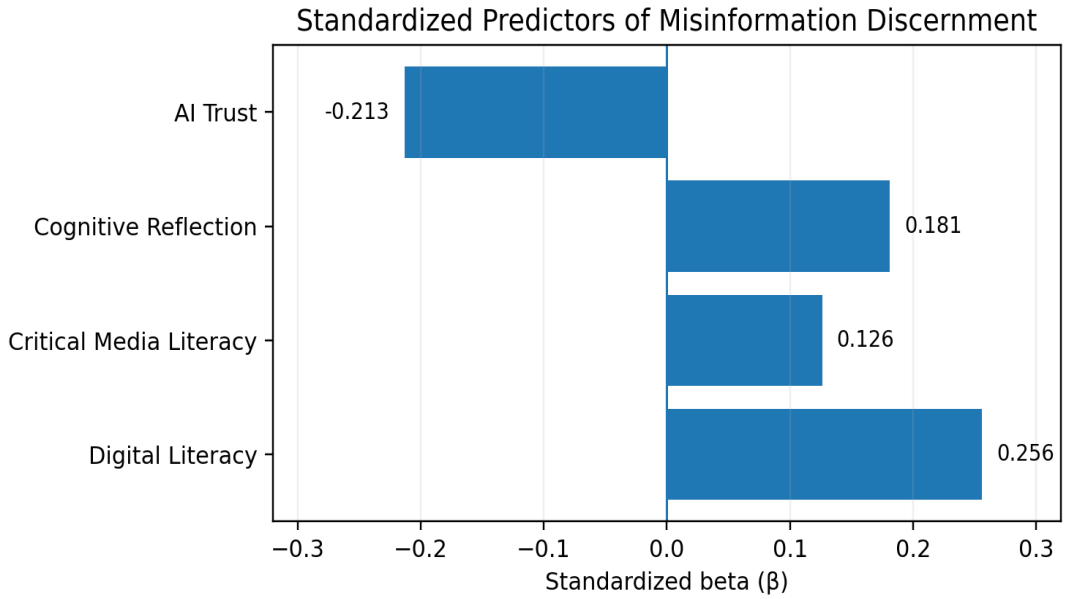



Figure 6. Standardized regression coefficients predicting misinformation discernment.

**Hypothesis Testing Summary**

The regression findings provided statistical support for all four study hypotheses. Digital Literacy and Critical Media Literacy significantly and positively predicted misinformation discernment, while Cognitive Reflection also demonstrated a significant positive predictive effect. AI Trust showed a significant negative association with misinformation discernment, supporting the expectation that greater uncritical trust in generative AI is linked with greater vulnerability to misleading information.

**Table 11. Summary of Hypothesis Testing**

| Hypothesis | Key Statistical Evidence | p | Decision |
|---|---|---|---|
| H1: Digital Literacy positively predicts misinformation discernment | β = .256; t = 8.68 | < .001 | Supported |
| H2: Critical Media Literacy positively predicts misinformation discernment | β = .126; t = 4.23 | < .001 | Supported |
| H3: Cognitive Reflection positively predicts misinformation discernment | β = .181; t = 6.74 | < .001 | Supported |
| H4: Higher AI Trust is associated with weaker misinformation discernment | β = −.213; t = −7.96 | < .001 | Supported |

**Summary of the Findings**

In conclusion, the study finds that misinformation discernment of university students was correlated with informational competencies and psychological factors. Digital Literacy had the highest positive predictive value, and Cognitive Reflection and Critical Media Literacy also had significant predictive value. In contrast, AI Trust ran in the opposite direction – the more people trusted, the less they discriminated. The experimental condition (further) showed that students rated accurate information significantly higher than misleading information, with slightly better ratings for misleading information generated by the AI system than for misleading information generated by the human. The results of these studies lay the groundwork for the discussion in Chapter 5.

## DISCUSSION, CONCLUSION AND RECOMMENDATIONS

**Overview of the Discussion**

This chapter explains the findings of the study in the light of the study objectives, research questions, hypothesis and literature review conducted earlier. This study explored the relationships between digital literacy, critical media literacy, cognitive reflection, and trust on generative AI with misinformation discernment among the University students in Pakistan. It also contrasted students' assessments of accurate and misleading information that was presented as human or AI created. Finally, 1120 students from the higher education institutions of Lahore and Rawalpindi-Islamabad area were the samples.

The overall pattern that the results showed was coherent. The strongest positive predictors of misinformation discernment were digital literacy (β = .256, p < .001), cognitive reflection (β = .181, p < .001), and critical media literacy (β = .126, p < .001). There was a strong negative correlation between AI

Trust and discernment ($\beta = -.213$, $p < .001$). The regression model accounted for 28.5% of the variance in the misinformation discernment. Results from the experimental condition indicated that accurate information had a significantly higher mean rating than misleading information, and that misleading AI-generated content had a slightly higher mean rating than misleading human-generated content. These insights, when combined, indicate that misinformation resilience requires a mix of practical verification skills, critical evaluation, reflective reasoning, and a well-balanced trust in an AI system.

**Digital Literacy, Critical Media Literacy, and Misinformation Discernment**

The most robust finding was in relation to digital literacy. The relationship among digital-literacy with the ability to distinguish accurate information from misinformation was the strongest ($r = .419$, $p < .001$) and digital-literacy was the most robust positive predictor in the regression model. This is a finding that helps to support the argument of Chapter 2 that digital literacy is not just knowing how to use devices or how to use online platforms. It encompasses the skill of finding sources, verifying information on various websites, verifying dates and authorship, comparing information with original sources, locating suspicious information, and employing verification strategies. De Vries et al. (2025) also define digital competence as a multi-dimensional competence that encompasses strategic information use, critical information evaluation, problem solving, information safety and competence related to artificial intelligence.

The discovery also corroborates with the evidence from Pakistan. Ali and Qazi (2023) demonstrated that misinformation education worked best in the context of providing personalized feedback and that it was positively impacted by participants' preexisting digital literacy. The current findings build on this argument by demonstrating a direct link between digital literacy and students' ability to discern misinformation. It is important to note that the regular use of the internet or the use of digital technologies is not a justification for concluding that students can determine the accuracy of information on the internet.

Discernment was also greatly aided by critical media literacy. The positive regression coefficient means that students who were more likely to question framing, evidence, motives of the source, bias and algorithmic influence were more immune to misleading content. This aligns with the perspective of McPhee and Jerowsky (2025), which calls on higher education to shift its focus from learning technical skills with AI to how to engage with AI-generated information. Additionally, it resonates with García-López and Trujillo-Liñán (2025), who list misinformation, opacity, bias and unequal access as key issues in the education application of generative AI. The two results can thus be said to complement each other, digital literacy offering the practical tools for verification, while critical media literacy offers the evaluative perspective needed for determining when and why it is needed.

**Cognitive Reflection and Truth Discernment**

Another important positive predictor of misinformation discernment was cognitive reflection. In terms of the reflective-reasoning scores, students who scored higher did better at separating reliable information from misleading information. This is in line with dual-process theories of reasoning, which propose that there are two modes of reasoning: intuitive (fast) and evaluative (slow). Cognitive reflection was introduced by Frederick (2005) as a tendency to revisit a seemingly convincing response, and later misinformation studies have found that analytical reasoning is related to greater truth detection skills (Pennycook & Rand, 2019, 2021).

This is particularly significant in an AI-driven context, as generative systems can provide structured, confident, and fluent responses. This form of communication can give instant credibility to a statement even if the statement is not true. Reflective reasoning students will be more likely to challenge this initial impression by asking if there is evidence to support the information, if there are sources cited, and if there

are independent sources that confirm the information. Therefore, the present finding indicates that misinformation education should not only develop fact-checking skills, but also foster behaviors of critical thinking, skepticism, and evidence-based decision making.

**Trust in Generative AI and Misinformation Vulnerability**

AI Trust was negatively correlated with misinformation discernment ($r = -.343$, $p < .001$), and was a significant negative predictor when the other variables were included. This suggests that learners with more non-reflective uses of generative AI exhibited less discrimination between correct and incorrect information. This finding lends confirmation to the distinction between calibrated trust and uncritical trust in Chapter 2. Educational goal should not be to instill general distrust of AI as a learning, explaining, translating and information retrieval tool. Students should rather be able to vary their faith based on evidence, source clarity, task nature and the potential for independent verification.

This interpretation aligns with the findings of Liu et al. (2025), which demonstrated that self-efficacy and cynicism impact perceptions of AI-generated content, as well as Zhang et al. (2025) who identified correlations between AI literacy, AI trust, dependency, and twenty-first-century skills in general. From the current findings, it can be seen that trust is problematic when fluent AI output is not verified. Confidence, fluency, speed and impression of neutrality should not therefore be taught as replacement for the factual reliability in the university.

**Evaluation of AI-Generated and Human-Generated Content**

The experimental task was used to give another angle on the judgments made by the students. Accurate human generated content ($M = 4.15$, $SD = 0.38$) was rated very similarly to accurate AI generated content ($M = 4.04$, $SD = 0.41$). The ratings for misleading human generated content were significantly lower ($M = 2.07$, $SD = 0.41$), and misleading AI generated content got a slightly higher mean rating ($M = 2.13$, $SD = 0.40$). The general pattern of this descriptive data indicated that participants tended to identify correct from incorrect information, whether from an external source or internal source. Simultaneously, the slightly more favorable ratings for misleading AI-generated content indicate that false content presented by AI technology might have a slight “credibility” advantage.

This is consistent with previous studies which found that users are not confident enough in their ability to recognize synthetic media in all contexts and can even sometimes have a positive attitude toward AI-generated media (Groh et al., 2024; Baptista et al., 2025; Ching et al., 2025). It also reinforces the idea that identifying if content is AI-generated is not the core of the education challenge. Information can be accurate or not, produced by humans or AIs. Therefore, students need to develop source independent truth-evaluation skills involving the use of evidence, corroboration, source identification, and factual consistency.

In the experimental section of Chapter 4 however, only the four condition means and standard deviations are reported, and no inferential statistics are reported for the $2 \times 2$ ANOVA. Thus, differences between the conditions generated by the AI and the human-generated conditions should be viewed as descriptive differences, not as statistically established differences. This distinction should be made for consistency between the results reported, and conclusions drawn from these results.

**Alignment with the Research Questions and Hypotheses**

The results demonstrate good support for the central relationships proposed in the study, but interpretation must be based upon the exact analyses presented in Chapter 4. Research Question 1 was significantly answered as both digital literacy and critical media literacy significantly positively predicted

misinformation discernment. The analysis of descriptive results also revealed a high level of verification behavior, but there was no description of willingness to spread false information as an independent result in the final regression model. The big positive effect of cognitive reflection directly answered Research Question 2. The findings of Research Question 3 were also explored in this study with significant negative correlation between AI Trust and Misinformation Discernment.

The four experimental condition means provided descriptive support for Research Question 4, where the content was clearly differentiated between accurate and misleading information, and there was a small difference between the AI and human-origin misleading information. However, as originally written in H4, there are interactions (AI Trust × Digital Literacy and AI Trust × Cognitive Reflection). This type of interaction statistics is not reported in Chapter 4. Thus, the direct negative influence of AI Trust is supported but the moderation proposition should be regarded as something that needs to be tested further and not as established by the reported regression model.

**Theoretical and Practical Implications**

The results are consistent with a model of misinformation resilience that is integrated. None of the single factors alone accounted for students' performance. Digital literacy helped develop competencies for practical verification, critical media literacy led students to scrutinize the content and sources, and cognitive reflection provided a means to engage in deliberate reasoning, while AI trust had an effect on the acceptance of generated information. These factors accounted for 28.5% of the variance in misinformation discernment, suggesting that these are important factors, but that there is still a significant role for other factors, including prior knowledge, language proficiency, topic familiarity, political attitudes, social influence, and platform-specific cues.

The findings have direct implications in the curriculum for higher education in Pakistan. Critical AI literacy and misinformation verification must be integrated into other professional domains and incorporated into the curriculum of universities. Pupils should be taught practical skills in sourcing, lateral reading, reverse image searching, verifying citations, comparing claims against authoritative sources, spotting fabricated claims and assessing claims created by AI. This should also involve reflective activities such as where students have to explain why a claim is credible (not just say they think it is or not).

Inconsistent infrastructure, training, confidence, language and institutional support are discussed in the reviewed literature related to digital transformation, which is particularly relevant to Pakistan (Mehmood, Zaman, et al., 2025; Mehmood, Ziauddin, & Naseem, 2025; Rehman et al., 2025). As such, AI-literacy interventions should be locally responsive, accessible in English and Urdu, when required, and teacher-supported and supported by institutional policy. Recent research on AI-integrated curricula in Pakistan also underlines the importance of making technological innovations inclusive, contextually relevant, and people-centred (Mehmood, Khan, et al. 2026).

**RECOMMENDATIONS**

Based on the results, universities should integrate structured modules on digital- and AI-literacy, including a critical evaluation and verification practice. The teaching of these modules needs to be explicit on teaching students that generative AI can generate fake citations, unsupported claims, and convincing misinformation despite sounding authoritative in its language.

Second, the teacher needs to incorporate verification based tasks in the regular coursework. Students can be required to compare the answers they find generated by AI to a research article, authoritative database, institutional report, or trusted news source, and record the process for each claim they find. Thirdly,

universities need to encourage reflective thinking and ask students to explain their credibility judgements, alternative interpretations, and reread/rewrite initial conclusions before accepting or sharing information.

Fourth, AI-use policies should foster calibrated instead of unconditional trust. The students should be informed when it is appropriate to use AI to brainstorm or explain and when it is necessary to verify high-stakes claims. Last but not least, universities need to offer multilingual and discipline-specific training, enabling students with diverse academic and linguistic backgrounds to build similar misinformation-resilience skills.

**Limitations and Directions for Future Research**

There are a number of limitations to be considered. The study employed a cross-sectional design and thus only allowed for the identification of associations, but not causal relationships over time between literacy, reflection, trust, and misinformation discernment. Secondly, the sample comprised 1120 students, but was collected from six universities in Lahore and Rawalpindi-Islamabad region, which limits the findings of the study. The results should then be interpreted with prudence and applied to other Pakistani university students, especially to subjects attending the smaller cities, rural areas, and university institutions having different levels of digital access.

Third, some of the survey literacy, trust and verification measures were self-reported, and this may be influenced by social desirability or overestimation of skills. This was mitigated in a recent embedded misinformation task, by incorporating direct evaluations of content; however, future studies should include behavioral verification indicators like opening the content, searching for confirming evidence, or using fact-checking tools. Fourth, the reported experimental results pertained to the mean, but not to the inferential statistics used to determine if observed differences between human and AI were statistically significant.Fourth, the reported experimental results referred to the mean but did not include the inferential statistics required to determine whether or not the observed difference between AI and human was statistically significant.

It is recommended that the whole 2 × 2 experimental model with the reported main and interaction effects be tested in future studies, and moderation relationships proposed in the original H4 be directly tested. In addition, longitudinal studies and intervention studies also would be useful to see if the discernment in misinformation stemming from targeted AI-literacy training is maintained over time. More research should be conducted on students from other provinces, disciplines and degree levels and explore if language or familiarity with the topic, prior experience using AI, or experience doing fact-checking affect the relationship between trust and susceptibility to misinformation.

## CONCLUSION

The study illustrates that the capacity of Pakistani university students to learn how to manage AI-driven misinformation is not only reliant on digital technology, but also on other factors. Digital literacy proved to be the most powerful positive predictor of misinformation discernment, and critical media literacy, cognitive reflection also had significant independent contributions. However, stronger uncritical trust of generative AI was correlated with weaker discernment. The experimental ratings also revealed that most of the students assessed the correctness of the information provided from the AI models, while the information from the misleading AI models was rated slightly higher than the information from the misleading human-generated models.

Overall, the results provide a framework for a "misinformation resilience" model involving verification skills, critical media awareness, reflective reasoning, and calibrated trust. Higher education institutions must now go beyond teaching students how to use generative AI tools, especially as these technologies

become more integrated into learning and information interaction in the university. The more important educational task is teaching students to ask questions and validate and responsibly evaluate information these tools generate. These capabilities are essential not just for universities to enable students to engage with AI-enabled learning, but also to be able to fluently interact with a credible and misleading information landscape.

ACADEMIA International Journal for Social Sciences
Volume 5, Issue 3, 2026 ISSN-L (Online): 3006-6638

# AI-Generated Content, Digital Misinformation, and Critical Media Literacy among Pakistani University Students: The Roles of Trust, Cognitive Reflection, and Digital Literacy

**Dr. Muhammad Umar Mehmood**
*(Corresponding Author)*
PhD Education,
Punjab School Education Department
Email: Pashagee143@gmail.com

**Dr. Um E Rubab**
Assistant Professor,
Department of Education,
Alhamd Islamic University (AIU),
Islamabad
Email: dr.umerubab@aiu.edu.pk

**Dr. Saira Maqbool**
Visiting Lecturer,
Department of Education,
Pir Mehar Ali Shah Arid Agriculture
University Rawalpindi
Email: saira.maqbool81@gmail.com

**Muhammad Mubeen Goraya**
MPhil. Education,
Preston University Kohat, Kohat
Email: mubeengoraya142@gmail.com

## BACKGROUND

Generative AI tools are increasingly used by university students for information seeking and content creation, raising concerns about digital misinformation. Students need more than operational skills; they require digital literacy, critical media literacy, reflective thinking, and calibrated trust to discern reliable information.

## OBJECTIVES

- To determine the relationship of digital literacy with misinformation discernment.
- To examine the relationship of critical media literacy with misinformation discernment.
- To investigate the relationship of cognitive reflection with misinformation discernment.
- To determine the relationship of trust in generative AI with misinformation discernment.

## HYPOTHESES

**H1:** Digital literacy positively predicts misinformation discernment.

**H2:** Critical media literacy positively predicts misinformation discernment.

**H3:** Cognitive reflection positively predicts misinformation discernment.

**H4:** Higher trust in generative AI is associated with weaker misinformation discernment.

## METHODOLOGY

**Design:** Quantitative cross-sectional with embedded 2 × 2 experimental task

**Population:** Undergraduate and postgraduate students in Lahore and Rawalpindi-Islamabad

**Sampling:** Multistage stratified sampling across public & private universities

**Sample:** Target N = 1,200; Final N = 1,120 (Retention rate = 93.3%)

**Instrument:** Structured bilingual questionnaire + 24-item misinformation discernment task

**Analysis:** Descriptive statistics, Pearson correlations, Multiple regression, 2 × 2 repeated-measures ANOVA

**Software:** IBM SPSS Statistics (Version 28)

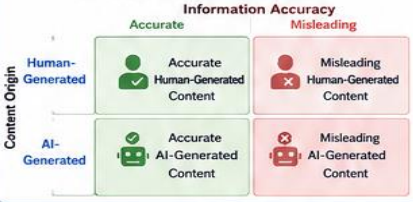
**Data Collection:** September–October 2025

## SAMPLE OVERVIEW

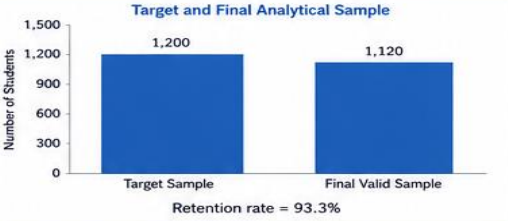
| | |
|---|---|
| Participating Universities | 6 |
| Geographic Distribution | 3 Lahore; 3 Rawalpindi-Islamabad |
| Institutional Sector | Public and Private |
| Target Sample | 1,200 |
| Final Valid Sample | 1,120 |
| Retention Rate | 93.3% |

**Geographical Distribution of Participating Universities**


## 2 × 2 EXPERIMENTAL DESIGN




**Target and Final Analytical Sample**




Retention rate = 93.3%

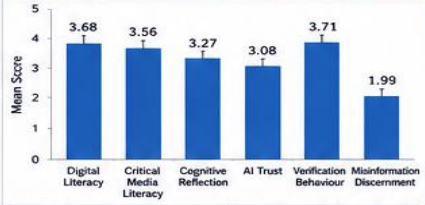
## MEASUREMENT SCALES

| | |
|---|---|
| Digital Literacy | Online search, source checking, cross-source comparison, author/date verification, recognizing suspicious content, fact-checking. |
| Critical Media Literacy | Questioning purpose, source, evidence, framing, bias, algorithmic influence, fact/opinion distinction, verification before sharing. |
| Trust in Information Sources | Trust in generative AI, social media, professional journalism, government/institutions, experts, and interpersonal sources. |
| Cognitive Reflection | Short reasoning items assessing tendency to rethink intuitive responses. |
| Misinformation Discernment Task | 24 items (accurate/misleading × human/AI). Ratings of accuracy, credibility, verification intention, sharing intention, and confidence. |

## RELIABILITY OF SCALES

| Scale | Cronbach's α |
|---|---|
| Digital Literacy | 0.86 |
| Critical Media Literacy | 0.83 |
| Trust in Information Sources | 0.79 |

## MEAN SCORES OF MAIN STUDY VARIABLES




| Variable | M | SD |
|---|---|---|
| Digital Literacy | 3.68 | 0.61 |
| Critical Media Literacy | 3.56 | 0.64 |
| Cognitive Reflection | 3.27 | 0.97 |
| AI Trust | 3.08 | 0.70 |
| Verification Behaviour | 3.71 | 0.68 |
| Misinformation Discernment | 1.99 | 0.67 |

## CORRELATIONS WITH MISINFORMATION DISCERNMENT

| Predictor | *r* | *p* |
|---|---|---|
| Digital Literacy | 0.419 | < .001 |
| Critical Media Literacy | 0.360 | < .001 |
| AI Trust | -0.343 | < .001 |

## MULTIPLE REGRESSION RESULTS

**DV: Misinformation Discernment**

| Predictor | β | *t* | *p* |
|---|---|---|---|
| Digital Literacy | 0.256 | 8.68 | < .001 |
| Critical Media Literacy | 0.126 | 4.23 | < .001 |
| Cognitive Reflection | 0.181 | 6.74 | < .001 |
| AI Trust | -0.213 | -7.96 | < .001 |

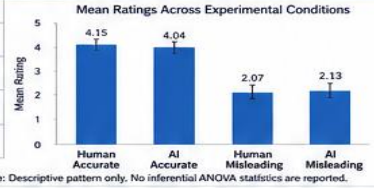
**Model Summary**
$R^2 = 0.285$, $F = 111.21$, $p < .001$

**STANDARDIZED REGRESSION COEFFICIENTS (β)**


## EXPERIMENTAL RESULTS: 2 × 2 CONTENT-EVALUATION CONDITIONS

| Condition | M | SD |
|---|---|---|
| Human Accurate | 4.15 | 0.38 |
| AI Accurate | 4.04 | 0.41 |
| Human Misleading | 2.07 | 0.41 |
| AI Misleading | 2.13 | 0.40 |

Mean Ratings Across Experimental Conditions




Note: Descriptive pattern only. No inferential ANOVA statistics are reported.

## HYPOTHESES TESTING SUMMARY

| Hypothesis | Result | Key Evidence |
|---|---|---|
| H1: Digital literacy positively predicts misinformation discernment. | Supported | β = 0.256, *p* < .001 |
| H2: Critical media literacy positively predicts misinformation discernment. | Supported | β = 0.126, *p* < .001 |
| H3: Cognitive reflection positively predicts misinformation discernment. | Supported | β = 0.181, *p* < .001 |
| H4: Higher trust in generative AI is associated with weaker discernment. | Supported | β = -0.213, *p* < .001 |

## KEY TAKEAWAY

Strengthening students' digital literacy, critical media literacy, and reflective thinking—while fostering calibrated trust in AI—is essential for building resilience against digital misinformation in Pakistani higher education.

**REFERENCES**

Ali, A., & Qazi, I. A. (2023). Countering misinformation on social media through educational interventions: Evidence from a randomized experiment in Pakistan. Journal of Development Economics, 163, 103108. https://doi.org/10.1016/j.jdeveco.2023.103108

Baptista, J. P., Rivas-de-Roca, R., Gradim, A., & Pérez-Curiel, C. (2025). Human-made news vs AI-generated news: A comparison of Portuguese and Spanish journalism students' evaluations. Humanities and Social Sciences Communications, 12, 567. https://doi.org/10.1057/s41599-025-04872-2

Buchanan, T., Perach, R., Husbands, D., Tout, A. F., Kostyuk, E., Kempley, J., & Joyner, L. (2024). Individual differences in sharing false political information on social media: Deliberate and accidental sharing, motivations and positive schizotypy. PLOS ONE, 19(6), e0304855. https://doi.org/10.1371/journal.pone.0304855

Caci, B., Giordano, G., Alesi, M., Gentile, A., Agnello, C., Lo Presti, L., La Cascia, M., Ingoglia, S., Inguglia, C., & Volpes, A. (2024). The public mental representations of deepfake technology: An in-depth qualitative exploration through Quora text data analysis. PLOS ONE, 19(12), e0313605. https://doi.org/10.1371/journal.pone.0313605

Ching, D., Twomey, J., Aylett, M. P., Quayle, M., Linehan, C., & Murphy, G. (2025). Can deepfakes manipulate us? Assessing the evidence via a critical scoping review. PLOS ONE, 20(5), e0320124. https://doi.org/10.1371/journal.pone.0320124

de Vries, D. A., Piotrowski, J. T., & de Vreese, C. (2025). Developing the DigIQ: A measure of digital competence. PLOS ONE, 20(5), e0322995. https://doi.org/10.1371/journal.pone.0322995

Du, H., Sun, Y., Jiang, H., Islam, A. Y. M. A., & Gu, X. (2024). Exploring the effects of AI literacy in teacher learning: An empirical study. Humanities and Social Sciences Communications, 11, 559. https://doi.org/10.1057/s41599-024-03101-6

Frederick, S. (2005). Cognitive reflection and decision making. Journal of Economic Perspectives, 19(4), 25–42. https://doi.org/10.1257/089533005775196732

Freedom Network & Coalition Against Disinformation. (2024). Perceptions and impact of disinformation on Pakistani youth: How university students understand disinformation. Freedom Network.

García-López, I. M., & Trujillo-Liñán, L. (2025). Ethical and regulatory challenges of generative AI in education: A systematic review. Frontiers in Education, 10, 1565938. https://doi.org/10.3389/feduc.2025.1565938

Groh, M., Sankaranarayanan, A., Singh, N., Kim, D. Y., Lippman, A., & Picard, R. (2024). Human detection of political speech deepfakes across transcripts, audio, and video. Nature Communications, 15, 7629. https://doi.org/10.1038/s41467-024-51998-z

Liu, Q., Wang, L., & Luo, M. (2025). When seeing is not believing: Self-efficacy and cynicism in the era of intelligent media. Humanities and Social Sciences Communications, 12, 274. https://doi.org/10.1057/s41599-025-04594-5

McPhee, S. W., & Jerowsky, M. (2025). Beyond technical skills: A pedagogical perspective on fostering critical engagement with generative AI in university classrooms. Frontiers in Education, 10, 1593278. https://doi.org/10.3389/feduc.2025.1593278

Mehmood, M. U., Bano, R., Zubair, A., & Bhatti, A. U. R. (2025). Psychological impact of languages on the human mind: Research on the contribution of psycholinguistics approach to teaching and learning English. Journal of Applied Linguistics and TESOL, 8(3), 1160–1174.

Mehmood, M. U., Fatima, A., Jamal, B., & Bhatti, A. U. R. (2025). The role of English linguistic competence in developing critical thinking, empathy, and psychological well-being through education. Contemporary Journal of Social Science Review, 3(4), 1616–1632.

Mehmood, M. U., Jamil, B., Rubab, U. E., & Bhatti, A. U. R. (2026). Psycholinguistic interventions for enhancing social inclusion of non-verbal students through assistive technology. AL-AASAR Journal, 3(1), 545–563.

Mehmood, M. U., Jamal, B., Dara, D., & Goraya, M. M. (2025). The psychological processes behind English language learning and their role in achieving inclusive and equitable quality education. ACADEMIA International Journal for Social Sciences, 4(4), 6421–6439. https://doi.org/10.63056/academia.4.4.2025.1478

Mehmood, M. U., Khan, S., Rubab, U. E., Sama, A., & Goraya, M. M. (2026). AI-powered curriculum strategies to promote inclusive and quality higher education. Qualitative Research Journal for Social Studies, 3(2), 450–475.

Mehmood, M. U., & Parveen, Z. (2021). Explore areas of needed support, competencies, attitudes, and concerns of elementary school teachers for inclusion of CWDS in elementary schools of Punjab. Global Educational Studies Review, 6(3), 109–120. https://doi.org/10.31703/gesr.2021(VI-III).12

Mehmood, M. U., & Parveen, Z. (2024a). A study to find out organizational stress among special education teachers in Govt. special education centers in Punjab. Al-Qudwah, 2(4), 181–193.

Mehmood, M. U., & Parveen, Z. (2024b). Exploring the Islamic perspective on special persons: A study of the Quran and Hadith. AL-ĪMĀN Research Journal, 2(3), 114–128.

Mehmood, M. U., & Parveen, Z. (2025). Empowering teachers for inclusive education: A support program for children with disabilities in Punjab's elementary schools. International Journal of Social Sciences Bulletin, 3(3), 390–405. https://doi.org/10.5281/zenodo.15010375

Mehmood, M. U., Rehman, M., Sama, A., & Goraya, M. M. (2026). The evolving role of teachers and mentors in inclusive education: From facilitators to champions of diversity and sustainability in line with SDG 4 and SDG 5. ACADEMIA International Journal for Social Sciences, 5(3[s9]), 1–24. https://doi.org/10.63056/academia.5.3(s9).2026.2131

Mehmood, M. U., Rehman, M., Sama, A., & Shaheen, B. (2025). The role of school leadership in promoting psychological safety and inclusion in public schools: Lessons for SDG-4 implementation. American Journal of Psychiatric Rehabilitation, 28(2), 185–195. https://doi.org/10.69980/ajpr.v28i2.840

Mehmood, M. U., Rubab, U. E., Goraya, M. M., & Bhatti, A. U. R. (2026). First-language exclusion and socio-psychological well-being among marginalized children in early childhood education in Pakistan: An explanatory sequential mixed-methods study. Qualitative Research Journal for Social Studies, 3(1), 1279–1303.

Mehmood, M. U., Ul Ain, Q., Jamal, B., & Bhatti, A. U. R. (2025). Integrating indigenous cultural practices into early childhood curricula: Challenges and opportunities. Journal of Political Stability Archive, 3(2), 480–499. https://doi.org/10.63468/jpsa.3.2.28

Mehmood, M. U., Zaheer, K., Jamal, B., & Goraya, M. M. (2025). Motivation, language development, and educational success: A psychological and linguistic approach to English teaching for sustainable learning outcomes (SDGs 4 & 8). Contemporary Journal of Social Science Review, 3(4), 1764–1781.

Mehmood, M. U., Zaman, Q.-u., Iftikhar, N., & Bhatti, A. U. R. (2025). Digital transformation in Pakistan's public schools: Opportunities and challenges for SDG-4. Journal of Applied Linguistics and TESOL, 8(3), 1669–1687.

Mehmood, M. U., Ziauddin, & Naseem, A. (2025). Socio-cognitive barriers in adopting English as a medium of instruction in Pakistani secondary schools: Challenges and stakeholder perspectives. AL-ĪMĀN Research Journal, 3(1), 122–137.

Ninghardjanti, P., Umam, M. C., Subarno, A., Winarno, W., Langgi, N. R., & Widodo, J. (2025). Evaluating the impact of AI on the critical thinking skills among the higher education students by combining the TAM model and critical thinking theory. Frontiers in Education, 10, 1719625. https://doi.org/10.3389/feduc.2025.1719625

Peng, W., Meng, J., & Ling, T.-W. (2024). The media literacy dilemma: Can ChatGPT facilitate the discernment of online health misinformation? Frontiers in Communication, 9, 1487213. https://doi.org/10.3389/fcomm.2024.1487213

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. Cognition, 188, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

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

Rehman, F., Sama, A., Rehman, M., Ahmed, H. M. A. M., & Mehmood, M. U. (2025). Academic writing challenges in English: A linguistic and psychological analysis of university students. Journal of Political Stability Archive, 3(1), 503–525.

Salhab, R., & Aboushi, M. M. (2025). Influence of AI literacy and 21st-century skills on the acceptance of generative artificial intelligence among college students. Frontiers in Education, 10, 1640212. https://doi.org/10.3389/feduc.2025.1640212

Sergeeva, O. V., Masalimova, A. R., Zheltukhina, M. R., Chikileva, L. S., Lutskovskai, L. Y., & Luzin, A. (2025). Impact of digital media literacy on attitude toward generative AI acceptance in higher education. Frontiers in Education, 10, 1563148. https://doi.org/10.3389/feduc.2025.1563148

Tomassi, A., Falegnami, A., & Romano, E. (2024). Mapping automatic social media information disorder: The role of bots and AI in spreading misleading information in society. PLOS ONE, 19(5), e0303183. https://doi.org/10.1371/journal.pone.0303183

Zhang, D., Wijaya, T. T., Wang, Y., Su, M., Li, X., & Damayanti, N. W. (2025). Exploring the relationship between AI literacy, AI trust, AI dependency, and 21st century skills in preservice mathematics teachers. Scientific Reports, 15, 14281. https://doi.org/10.1038/s41598-025-99127-0