Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Trust and truth in the post-truth era: cognitive and attitudinal predictors of confidence in ai-generated content

Murat Aytas

muratayta s@selcuk.edu.tr

Selçuk University

Mehmet Kucuktongur

Selçuk University

Article

Keywords: generative artificial intelligence, trust, AI literacy, attitude, critical inquiry, post-truth

Posted Date: May 21st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9464666/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Trust in artificial intelligence (AI) generated content has emerged as a central concern in the post-truth era, where the boundaries between authenticity and fabrication are increasingly blurred. This study investigates how cognitive and attitudinal factors determine individuals' trust in AI-generated content in the context of the post-truth era. Using quantitative field design, data were collected from 232 university students, of whom 207 provided complete responses to the main scales. Hierarchical regression analysis was conducted to explore how generative AI literacy, user attitudes, and critical inquiry predict perceived trust in reality. The results reveal that generative AI literacy and frequency of use are positively associated with ethical awareness and manipulation recognition, indicating that cognitive familiarity enhances critical sensitivity. Furthermore, a negative attitude toward AI significantly decreases trust in reality ($\beta = -.45$, $p < .001$), while a positive attitude ($\beta = +.31$, $p = .027$) and critical inquiry ($\beta = -.35$, $p < .001$) exert independent and meaningful effects. The model explains 39% of the variance in trust ($R^2 = .39$) and shows no multicollinearity issues (VIF < 5). These findings suggest that trust in AI-generated content is not merely an outcome of exposure or technological proficiency but a product of balanced cognitive literacy and attitudinal orientation. The study contributes to theoretical debates on digital trust by demonstrating how critical intelligence moderates the relationship between generative AI engagement and perceptions of truth in mediated environments.

Introduction

The rapid proliferation of generative artificial intelligence (AI) technologies has fundamentally transformed how individuals produce, disseminate, and evaluate information. Tools such as ChatGPT, DALL·E, and Midjourney have blurred the once-clear boundaries between human and machine authorship, altering not only creative industries but also the epistemic foundations of trust in digital communication (Floridi, 2023; Sahebi & Formosa, 2025). In an era increasingly described as *post-truth*, where emotional resonance often outweighs factual accuracy, understanding how users assess the reliability of AI-generated content has become a crucial interdisciplinary concern across communication, psychology, and cognitive science (Lewandowsky et al., 2017; Pennycook & Rand, 2021).

Trust in information has traditionally depended on human judgment, source credibility, and institutional legitimacy (Luhmann, 2018). However, the rise of generative AI introduces a new epistemological challenge: when algorithms generate fluent, human-like texts and images without explicitly disclosing their synthetic origins, the conventional heuristics for assessing authenticity become unstable (Appelman & Bien-Aimé, 2024; Flew, 2019). This automation of creativity (Floridi, 2023; Napoli, 2019) disrupts the perceptual cues by which individuals distinguish truth from simulation, thereby intensifying epistemic uncertainty in algorithmic environments (Nightingale & Farid, 2022; Vaccari & Chadwick, 2020). As synthetic media, misinformation, and *deepfake* technologies grow increasingly sophisticated, clarifying the cognitive and attitudinal mechanisms behind trust judgments become both a theoretical and a societal imperative (Al-kfairy et al., 2024).

While early research conceptualized trust primarily through technological acceptance or risk perception frameworks (Hoff & Bashir, 2014; Wischnewski et al., 2023), more recent scholarship emphasizes the interplay of AI literacy, attitudinal orientation, and critical inquiry as key determinants of trust in AI-mediated information (Long & Magerko, 2020; Ng et al., 2021). AI literacy entails the cognitive capacity to understand the principles, affordances, and limitations of AI systems, enabling individuals to recognize algorithmic bias, manipulation, and error (Jobin et al., 2019; Carvalho et al., 2022). Attitudes toward AI encompass affective and evaluative responses ranging from fascination and optimism to ethical apprehension and skepticism (Panda & Roy, 2024; Fast & Horvitz, 2017). Critical inquiry, understood as a metacognitive disposition toward questioning, verifying, and reflecting on the epistemic validity of content, mediates how individuals translate knowledge and emotion into trust or doubt (Dame Adjin-Tettey, 2022; Mihailidis & Viotty, 2017). Together, these dimensions provide a robust framework for interpreting users' trust formation processes in generative AI contexts.

Empirical evidence supports the interdependence of these factors. Studies reveal that higher AI literacy predicts both improved discernment of synthetic content and more stable trust calibration (Long & Magerko, 2020; Bergdahl et al., 2023). Conversely, aesthetic realism and fluency, which are characteristic of generative systems, often create an illusion of truth, leading users to overestimate credibility even when content accuracy is low (Kishnani, 2025; Hohenstein & Jung, 2020). Furthermore, ethical awareness and reflective skepticism have been shown to mitigate such biases by anchoring trust judgments in deliberation rather than affective resonance (Chen et al., 2023; Al-kfairy et al., 2024). Despite these insights, there remains a limited empirical understanding of how cognitive (literacy) and affective (attitudes) factors jointly influence perceived reliability, and how metacognitive processes, such as critical inquiry, modulate this relationship. Within this framework, the present study investigates the predictors of trust in AI-generated content by integrating cognitive, attitudinal, and metacognitive variables. Specifically, it examines (a) AI literacy as a cognitive determinant, (b) positive and negative attitudes toward AI as affective predictors, and (c) critical inquiry as a metacognitive moderator that reshapes the association between AI engagement and perceived trust in reality. Drawing on prior work in media psychology, human–AI interaction, and digital epistemology, the study hypothesizes that greater AI literacy and positive attitudes will correlate with higher perceived trust, while negative attitudes and stronger critical inquiry tendencies will predict lower but more discerning trust (Lewandowsky et al., 2017; Sahebi & Formosa, 2025).

To test these hypotheses, hierarchical regression analysis was conducted using survey data from 232 university students, with 207 valid cases retained after data screening. The results indicate that trust in reality is not a unidimensional construct based solely on technological familiarity. Instead, it is co-constituted by cognitive competence and attitudinal stance. Negative attitudes toward AI significantly reduce perceived trust ($\beta = -.45$, $p < .001$), whereas positive attitudes ($\beta = +.31$, $p = .027$) and critical inquiry ($\beta = -.35$, $p < .001$) exert substantial predictive power. These findings highlight the *dual nature* of digital trust in post-truth environments where skepticism and confidence coexist as adaptive strategies for navigating algorithmically mediated information ecosystems.

By synthesizing perspectives from AI literacy research, trust theory, and critical epistemology, this study advances theoretical understanding of how individuals reconstruct authenticity, reliability, and truth in the age of generative media. It argues that epistemic trust in AI-generated content is neither a purely cognitive judgment nor a purely emotional response but a hybrid construct emerging at the intersection of knowledge, ethics, and affect.

Literature review

AI literacy refers to the cognitive, technical, and ethical competencies that enable individuals to understand, evaluate, and interact effectively with artificial intelligence systems (Long & Magerko, 2020; Müller & Bostrom, 2016; Jungwirth & Haluza, 2023). It encompasses awareness, use, evaluation, and ethical reflection, which together support informed and responsible engagement with algorithmic technologies (Ng et al., 2021). Recent studies conceptualize AI literacy as both a cognitive resource and a socio-ethical capacity. Individuals with higher levels of literacy not only demonstrate stronger digital competence but also exhibit greater confidence and discernment when evaluating AI-generated outputs (Lee & Park, 2024). With the widespread use of generative AI platforms such as ChatGPT, Midjourney, and DALL·E, scholars have emphasized the need for generative AI literacy, defined as the ability to interpret, verify, and critically assess content created by machine systems (Gokcearslan et al., 2024). Empirical evidence indicates that this form of literacy enhances users' ability to detect bias and misinformation in synthetic media (Epstein et al., 2022) and strengthens cognitive resilience in post-truth environments (Bulger & Davison, 2018; Cosentino, 2020. Consequently, AI literacy functions as both a technical skill and a protective competence that anchors trust in digital communication.

Ethical awareness, which involves understanding the moral implications of AI technologies, represents a core dimension of AI literacy (Jobin et al., 2019; Lee & Park, 2024). This awareness encompasses sensitivity to issues such as fairness, accountability, transparency, and privacy that shape users' moral evaluations of AI systems (Floridi & Cowls, 2022). Research demonstrates that ethical awareness can be developed through structured educational interventions that help individuals recognize and reason through algorithmic dilemmas (Tsamados et al., 2022). This ethical sensitivity plays a crucial role in calibrating trust by fostering a balance between confidence and caution. Users who are aware of ethical risks are less likely to accept AI outputs uncritically and are better able to identify manipulation and misinformation (Chen et al., 2023; Lutz et al., 2025). Findings suggest that ethical awareness moderates the relationship between AI literacy and trust. While literacy provides the knowledge needed to understand how AI operates, ethical reasoning guides users in determining how it should be used (Greene et al., 2019). Within generative contexts, ethical awareness becomes even more important for sustaining epistemic trust as individuals evolve from passive consumers to active co-creators of digital content (Boediman, 2025).

The decline of public trust in media and technology further illustrates the significance of these competencies. The proliferation of synthetic content and emotionally charged discourse has eroded traditional forms of media trust (Lewandowsky et al., 2017; Pennycook & Rand, 2021). Generative AI has

intensified this issue by producing highly realistic yet artificial media that blur the boundaries of authenticity (Vaccari & Chadwick, 2020). Nightingale and Farid (2022) demonstrated that AI-generated faces are often indistinguishable from real ones and may even appear more trustworthy, revealing how algorithmic realism can distort intuitive judgments. Similarly, studies show that disclosing AI authorship tends to decrease perceived credibility because audiences associate authenticity and accountability with human sources (Appelman & Bien-Aimé, 2024; Lobo Paes, 2024). However, labeling AI-generated content provides only limited improvement in accuracy judgments and tends to be effective primarily among users with high literacy levels (Bergdahl et al., 2023). These findings suggest that literacy and ethical reflection are more effective in maintaining critical engagement than superficial content warnings, especially in educational settings where students frequently use AI tools.

Public attitudes toward AI add another important dimension to the understanding of trust in generative technologies. These attitudes are often characterized by ambivalence, combining both fascination and apprehension (Fast & Horvitz, 2017; Ismatullaev & Kim, 2022). Global surveys indicate that while trust in AI remains moderate overall, it varies significantly across domains: individuals tend to trust AI more in fields such as education and healthcare but remain skeptical in areas such as journalism and creative production (KPMG, 2023; Zhang & Dafoe, 2019). This pattern suggests that contextual trust, shaped by perceived risks and the presence of human oversight, strongly influences user perceptions (Wischnewski et al., 2023). As users gain familiarity with AI systems, initial uncertainty often gives way to informed caution, particularly when systems operate with transparency and accountability (Shin, 2021). Scholars have distinguished between functional trust, referring to confidence in an AI system's performance, and relational trust, which concerns the alignment of AI behavior with human values (Panda & Roy, 2024). Generative AI challenges this distinction by demonstrating high functionality while simultaneously raising doubts about authenticity, thereby revealing an enduring tension between technological efficiency and moral legitimacy.

Across this growing body of literature, three determinants stand out as essential for understanding trust in AI-generated content: AI literacy, ethical awareness, and attitudes toward AI. These interconnected factors explain how individuals evaluate the reliability of synthetic information in an age where algorithms increasingly mediate truth. AI literacy provides the cognitive foundation for critical understanding, ethical awareness establishes a moral framework for decision-making, and attitudes reflect the emotional and evaluative dimensions of user interaction with AI. This study builds on these insights by empirically examining how cognitive and attitudinal variables jointly shape trust in generative AI among university students. This group stands at the forefront of both technological engagement and media evaluation in contemporary society. Drawing upon the reviewed literature, this study is guided by the following research questions:

- RQ 1: To what extent do cognitive and attitudinal factors, specifically generative AI literacy, positive and negative attitudes toward AI, and critical inquiry, predict individuals' trust in AI-generated content?

Previous research on trust in automation and artificial intelligence has predominantly emphasized affective dimensions and perceptions of risk (Hoff & Bashir, 2014; Ismatullaev & Kim, 2022). However, limited attention has been given to the interaction between literacy, the ability to understand and evaluate AI systems, and metacognitive inquiry, which reflects users' capacity to question and verify information (Long & Magerko, 2020; Bergdahl et al., 2023). Recent empirical findings suggest that higher AI literacy not only leads to increased adoption of AI tools but also enhances individuals' ability to discern misinformation produced by generative systems (Al-kfairy et al., 2024). By evaluating literacy and attitudes simultaneously, this study seeks to determine whether users' trust in AI-generated content is primarily based on cognitive comprehension or on affective orientation toward AI.

- RQ 2: How does critical inquiry moderate the relationship between users' engagement with generative AI tools, such as usage frequency and adoption of multiple systems, and their perceived trust in reality and AI-produced information?

Within the misinformation literature, metacognitive abilities such as verification behavior and critical reflection are consistently identified as protective factors against susceptibility to false information (Mihailidis & Viotty, 2017; Wineburg & McGrew, 2019). Recent studies on generative AI demonstrate that individuals with stronger inquiry orientations tend to approach AI outputs with informed skepticism, showing greater resistance to surface-level credibility cues (Chen et al., 2023; Hohenstein & Jung, 2020). This research question explores whether critical inquiry transforms users' engagement with AI from a potentially trust-reducing factor into a mechanism for more balanced and calibrated trust judgments.

- RQ 3: Do demographic and behavioral factors, including gender, age, frequency of generative AI use, and tool diversity, influence trust in AI-generated content, and how might these variables condition the relationships proposed in RQ1 and RQ2?

Prior studies on digital media trust reveal inconsistent effects of demographic variables and technology-use patterns (Luhmann, 2018; Metzger & Flanagin, 2021). Recent evidence indicates that while younger individuals and frequent users of generative AI tools often demonstrate higher familiarity and confidence, they do not necessarily exhibit greater epistemic caution or critical reflection (Ismatullaev & Kim, 2022; Lin & Ng, 2024). Including these demographic and behavioral variables as controls enables the isolation of the net effects of literacy, attitudes, and inquiry on the formation of trust in AI-generated information.

Together, these research questions aim to uncover how individuals develop epistemic trust in an AI-mediated information environment. By integrating generative AI literacy, attitudinal orientations, and critical inquiry within a unified predictive framework, the study offers an original empirical perspective on post-truth communication, media trust, and algorithmic epistemology.

Methodology

Research design and participants

This study adopted a quantitative, cross-sectional research design to systematically and generalizably examine how individuals perceive, evaluate, and trust AI-generated content. The survey method was chosen because it allows for the collection of standardized data from a large population and enables robust statistical testing of relationships among cognitive, attitudinal, and behavioral variables related to generative artificial intelligence (AI). This design aligns with SSCI journal standards emphasizing methodological transparency, reproducibility, and validity.

The research instrument was structured to measure participants' AI literacy, ethical awareness, attitudinal orientations toward AI, and perceived trust in AI-generated media content. These constructs were operationalized using validated scales from previous studies, adapted to the context of generative AI to ensure conceptual relevance. By integrating cognitive and affective dimensions of perception, the design sought to capture how users in the post-truth era construct epistemic trust in digital environments mediated by AI.

The sample consisted of 232 university students from multiple universities across Turkey, aged 18 to 60 years (M = 26.08, SD = 6.91). After data cleaning and listwise deletion for incomplete responses, 207 valid cases were retained for statistical analysis. The gender distribution was approximately balanced (51.3% female, 48.7% male). In terms of education level, 74.1% were undergraduate students, 19.4% were master's students, and 6.5% were doctoral students. Participants were primarily located in Konya (54.3%), followed by Ankara (13.4%), İstanbul (3.0%), İzmir (0.9%), and other regions (28.4%). University students were selected as the focal population because they represent a generation of active digital media users and emerging AI content creators, making them an analytically relevant group for understanding the intersection of trust, literacy, and authenticity in AI-mediated communication. Their familiarity with digital technologies also facilitates reliable self-assessment of AI-related experiences while offering a meaningful lens through which to examine evolving media epistemologies.

All procedures performed in studies involving human participants were conducted in accordance with the ethical standards of the institutional and national research committees and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study protocol was reviewed and approved by the Selçuk University Faculty of Communication Scientific Ethics Evaluation Committee (Approval Date: 26 May 2025, Decision No: 2025/11 – 07). All participants were informed about the purpose and voluntary nature of the research and provided informed consent prior to participation. Data collection was anonymous, and no personally identifiable information was recorded. The research fully complied with institutional and national data protection regulations.

Procedure and pilot test

Data collection was conducted online via Google Forms from May to June 2025. Prior to participation, all respondents were provided with a detailed explanation of the study's objectives, emphasizing the voluntary nature of participation and the confidentiality of all responses. Participants were required to read an informed consent statement and indicate their agreement before accessing the survey. The

survey link was disseminated via university mailing lists, online student communities, and social media platforms to reach a diverse, representative pool of university students across Turkey. The data collection process adhered strictly to ethical standards of voluntary participation, anonymity, and data confidentiality, ensuring compliance with institutional and international research ethics protocols. Before launching the primary survey, a pilot test was conducted to assess the validity, reliability, and clarity of the questionnaire items. The pilot sample comprised 50 university students who were not included in the final study. These participants completed the online survey and were invited to provide qualitative feedback regarding the wording, clarity, and interpretability of the items. Based on their input, minor revisions were made to improve item phrasing and reduce potential ambiguity. This pilot phase ensured that the final version of the instrument was both linguistically clear and psychometrically robust, thereby enhancing data quality and participant comprehension in the main study.

Instruments

The research instrument was designed to capture a comprehensive understanding of participants' cognitive, attitudinal, and behavioral orientations toward generative artificial intelligence (AI). The questionnaire consisted of four main sections, integrating validated scales with researcher-developed items to align with the study's conceptual model. The first section gathered demographic and background information, including gender, age, education level, field of study, and city of residence. The second section explored AI usage patterns, focusing on the frequency, purposes, and contexts in which participants used generative AI tools such as ChatGPT, DALL·E, and Midjourney. The third section measured perceptual and attitudinal constructs, including AI literacy, attitudes toward AI, perceived reality, and ethical sensitivity. Finally, the fourth section assessed critical and behavioral dimensions, including manipulation awareness and inquiry-based skepticism toward AI-generated content.

To ensure construct validity and theoretical grounding, the study employed and adapted three established measurement scales, with permissions obtained from the original authors. These included the *General Attitudes toward Artificial Intelligence Scale* (Kaya et al., 2022), the *Generative AI Literacy Scale* (Gokcearslan et al., 2024), and the *Perceived Digital Media Reality Scale* (Bowe & Wohn, 2015). Each scale was carefully translated and culturally adapted to the Turkish context following best practices in cross-cultural research (Van de Vijver & Leung, 1997). Minor linguistic modifications were made to ensure conceptual equivalence without altering the original constructs. Reverse-coded items (e.g., "No AI-generated content can resemble reality") were included to improve response accuracy and were recoded appropriately during data processing. In addition to these validated scales, several researcher-developed items were incorporated to assess perceived realism, trust, and evaluative judgments toward different AI-generated formats (visual, textual, and audio). These items enabled comparative analyses across modalities, reflecting the multidimensional nature of trust in generative media environments. The inclusion of custom items was guided by recent methodological recommendations emphasizing adaptive instrumentation in emerging AI contexts (Fast & Horvitz, 2017; Floridi, 2023).

All statistical analyses were performed using IBM SPSS v.30.0. All tests were two-tailed with an alpha level of .05. The normality of the data was verified through skewness and kurtosis values ($|z| < 2$), and assumptions of homoscedasticity and multicollinearity were confirmed prior to inferential analyses. Descriptive statistics (means, standard deviations, skewness, kurtosis) were computed to assess normality, with all variables meeting parametric assumptions. Pearson's correlations, Chi-square tests, one-way ANOVAs, and hierarchical multiple regressions were used to examine relationships and group differences. Model diagnostics confirmed normal residual distribution, homoscedasticity, and acceptable multicollinearity levels (VIF < 5). The analytical procedure followed four systematic stages. First, data screening and reliability testing were performed. Missing data were treated as "missing by design" when resulting from conditional logic (e.g., non-users of AI skipping usage-specific questions). Reliability coefficients indicated strong internal consistency across all scales (Cronbach's $\alpha = .78–.91$). Second, descriptive statistics, including frequencies, means, standard deviations, skewness, and kurtosis, were computed for each key variable to evaluate distributional assumptions. Third, inferential analyses were carried out: chi-square tests examined categorical group differences (e.g., gender × AI usage), ANOVAs tested variations in literacy and ethics scores across demographic groups, Pearson's correlations assessed zero-order relationships among primary constructs, and hierarchical multiple regression models tested the predictive framework of trust in AI-generated content. Fourth, model diagnostics confirmed the robustness of the regression analyses, with multicollinearity (VIF < 5) and normality (skewness and kurtosis within ± 2) assumptions satisfactorily met. This instrument design adheres to the methodological standards outlined by contemporary AI and media research (Epstein et al., 2022; Lee & Park, 2024), ensuring both theoretical alignment and statistical rigor. By combining validated scales, context-specific items, and a robust analytical framework, the study provides a reliable and replicable foundation for assessing how cognitive and attitudinal factors jointly shape individuals' trust in AI-generated media within higher education contexts.

Results

Sample profile and AI adoption overview

Table 1 presents the demographic and adoption characteristics of the participants. The majority were female (59.5%) and undergraduate students (74.1%), with a mean age of 26.08 years (SD = 6.91). Most respondents lived in Konya (54.3%). Overall, 89.2% reported using generative AI tools, with over half indicating frequent or widespread use. Nearly 91% had used such tools within the past week. We surveyed N = 232 university students. A compact demographics table (gender, age, education level, field of study, city) will replace multiple standalone tables. Adoption is high: ~89% report using generative AI tools, with frequent/widespread use exceeding 55% and 90% reporting use within the past week. Given the imbalance between users and non-users, subsequent inferential tests control usage frequency.

Table 1
Sample profile and ai adoption overview (N = 232)

Variable	Category	n	%
Gender	Female	138	59.5
	Male	94	40.5
Age	M = 26.08 (SD = 6.91), Range = 18–60	—	—
Education level	Undergraduate	172	74.1
	Master's	45	19.4
	Doctorate	15	6.5
Field of study	Communication Sciences	58	25.0
	Social Sciences	50	21.6
	Fine Arts	25	10.8
	Engineering	23	9.9
	Health Sciences	17	7.3
	Other	59	25.4
City of residence	Konya	126	54.3
	Ankara	31	13.4
	Istanbul	7	3.0
	Izmir	2	0.9
	Other	66	28.4
Generative AI usage	Yes, I use it	207	89.2
	No, I don't use it	25	10.8
Image-based AI tools	Yes	78	37.7
	No	129	62.3
Voice-based AI tools	Yes	63	30.4
	No	144	69.6
Usage frequency	Rarely	8	3.9
	Sometimes	30	14.5

Note. Values are based on self-reported survey responses. Mean ± SD and ranges are reported where applicable. Percentages are rounded to one decimal place. All tests were two-tailed, with α = .05.

Variable	Category	n	%
	Occasionally	54	26.1
	Often	58	28.0
	Very often	57	27.5
Last use	Today	95	45.9
	Within last week	93	44.9
	Within last month	14	6.8
	More than a month ago	5	2.4

Note. Values are based on self-reported survey responses. Mean ± SD and ranges are reported where applicable. Percentages are rounded to one decimal place. All tests were two-tailed, with α = .05.

Table 1 presents the demographic and adoption characteristics of the participants. The majority were female (59.5%) and undergraduate students (74.1%), with a mean age of 26.08 years (SD = 6.91). Most respondents lived in Konya (54.3%). Overall, 89.2% reported using generative AI tools, with over half indicating frequent or widespread use. Nearly 91% had used such tools within the past week.

Descriptive statistics and reliability of core scales

The descriptive statistics and reliability coefficients for the main scales used in this study are summarized in Table 2. The key constructs analyzed include Generative AI Literacy, General Attitudes toward Artificial Intelligence, Perceived Trust in Reality, and Trust in AI-Generated Content. Participants' levels of AI literacy were predominantly moderate (59.9%), with 25.4% demonstrating high levels. Regarding general attitudes toward AI, most respondents (66.4%) exhibited a neutral orientation, and 21.1% expressed a positive attitude. In terms of perceived trust in reality, participants reported moderate (45.7%) or high (37.9%) levels of trust, with only a small proportion (5.6%) indicating low trust. The skewness and kurtosis values for all scales remained within the ± 2 threshold, confirming that the data met the assumption of normal distribution. Internal consistency reliability was high across all measures, with Cronbach's α values ranging between .78 and .91, indicating strong internal coherence of the scales.

Table 2
Descriptive statistics and reliability coefficients of core scales (N = 207)

Scale / Variable	SD	Skewness	Kurtosis	Cronbach's α	Description (Category Distribution)
Generative AI Literacy	–	−1.24	2.24	.88	Low 3.9% / Medium 59.9% / High 25.4%
Attitude Toward AI	–	+ 0.09	+ 0.37	.84	Negative 1.3% / Neutral 66.4% / Positive 21.1%
Perceived Trust in Reality	–	+ 0.27	+ 0.21	.81	Low 5.6% / Medium 45.7% / High 37.9%
Trust in AI-Generated Content	–	−0.43	−0.44	.79	—

Note. Descriptive statistics include mean ± SD. All variables met normality assumptions (|skewness|, |kurtosis| < 2). Reliability coefficients were computed using Cronbach's α (α ≥ .78). All tests were two-tailed, with α = .05.

The results in Table 2 indicate that all measurement instruments demonstrated high internal consistency, with Cronbach's α coefficients exceeding .78, confirming strong reliability across the core constructs. Skewness and kurtosis values remained within the acceptable ± 2 range, satisfying the assumptions for parametric analyses and indicating that the data distributions were approximately normal. Overall, participants exhibited moderate to high levels of AI literacy and trust, with a generally neutral emotional or attitudinal stance toward artificial intelligence. This pattern suggests that while university students are familiar with and cognitively engaged with generative AI technologies, their affective orientation remains cautious rather than wholly embracing them. The combination of moderate literacy, stable reliability measures, and balanced distributional properties implies a sample that is both informed and reflective, representing a transitional user group that recognizes AI's potential while maintaining critical awareness of its implications. Such findings align with recent empirical evidence showing that higher exposure to AI fosters cognitive understanding without necessarily translating into unconditional trust (Lee & Park, 2024; Chen et al., 2023).

Zero-Order correlations among key variables

The bivariate relationships among the primary variables in the study were examined using Pearson's correlation analysis (see Table 3). The results revealed positive and significant correlations among generative AI literacy, ethical attitude, manipulation awareness, and frequency of AI use ($r = .43–.84$, $p < .001$). These findings indicate that individuals who engage more consciously with generative AI technologies tend to exhibit higher levels of ethical sensitivity and greater awareness of potential manipulations. A significant positive relationship was also observed between AI literacy and positive

attitudes toward AI (r = .54, p < .001). This suggests that as individuals' cognitive understanding and familiarity with AI systems increase, their emotional acceptance and evaluative orientation toward AI become more favorable. Furthermore, a strong negative correlation emerged between perceived trust in reality and detachment from reality (r = − .93, p < .001), indicating that as individuals' confidence in the credibility of digital information declines, their overall perception of reality becomes more fragile.

Table 3
Zero-Order correlations among key variables (N = 207)

Variable	1	2	3	4	5	6	7	8
1. Generative AI Literacy	—							
2. Manipulation Awareness	.43**	—						
3. Trust in AI-Generated Content	.41**	.89**	—					
4. Positive Attitude toward AI	.54**	.47**	.44**	—				
5. Ethical Sensitivity	.69**	.26**	.24**	.31**	—			
6. Frequency of AI Use	.84**	.39**	.33**	.51**	.41**	—		
7. Perceived Trust in Reality	.32**	.28**	.27**	.26**	.22**	.25**	—	
8. Detachment from Reality	−.29**	−.25**	−.23**	−.21**	−.18**	−.20**	−.93**	—

Note. Pearson's correlation coefficients (two-tailed) are reported. All correlations are significant at the p < .01 level (2-tailed). Actual p-values are reported. All tests were two-tailed, with α = .05.

All correlation coefficients were statistically significant at the p < .01 level, confirming that the inter-variable relationships were both robust and theoretically consistent. These patterns align with prior findings emphasizing the interconnected nature of literacy, ethics, and attitudinal trust in shaping individuals' engagement with AI-generated media (Lee & Park, 2024; Chen et al., 2023). The results presented in Table 3 reveal a coherent and theoretically meaningful pattern of relationships among the core constructs. High-level correlations (r > .70), particularly between Generative AI Literacy, Manipulation Awareness, Ethical Sensitivity, and Frequency of AI Use, indicate that individuals who are more literate in generative AI tend to engage with these technologies more frequently and exhibit heightened awareness of both ethical and manipulative dimensions. This suggests that AI literacy serves as a foundational cognitive competence that encourages reflective, ethically grounded engagement with AI systems.

The negative correlation between Perceived Trust in Reality and Detachment from Reality ($r = .93$, $p < .01$) provides strong empirical support for theoretical claims in media and communication research regarding the fragility of epistemic trust in post-truth environments. As individuals' confidence in the reliability of mediated content decreases, their sense of connection to objective reality also weakens, a dynamic consistent with prior findings on informational skepticism and media cynicism (Lewandowsky et al., 2017; Vaccari & Chadwick, 2020). Overall, the correlation matrix supports a sequential structural pattern linking cognitive understanding, ethical awareness, and trust judgments, forming a conceptual pathway from AI literacy through awareness and ethics to trust. This framework aligns with recent interdisciplinary research emphasizing that literacy and ethical sensitivity jointly shape how individuals calibrate trust and discern authenticity in AI-mediated communication (Chen et al., 2023; Lee & Park, 2024; Morley et al., 2020). The strength and direction of these relationships indicate that higher literacy and ethical competence not only enhance critical awareness but also foster a balanced form of trust within digital ecosystems increasingly influenced by generative AI content.

Group differences

To further explore individual variations in perceptions and behaviors related to generative AI, group comparisons were conducted across demographic and behavioral characteristics. As presented in Table 4, a Chi-square test examining gender differences in AI tool usage revealed a statistically significant association ($\chi^2 = 6.41$, $p = .011$). Female participants reported higher adoption rates of generative AI tools (93.5%) compared to male participants (83%). This finding indicates a moderate gender-based difference, suggesting that women may engage more actively with generative AI technologies in educational contexts.

Table 4
Group differences in generative ai literacy, ethical sensitivity, and usage frequency

Variable	F / χ^2	p-value	Significant Group	Direction of Difference
Gender × AI Use	6.41	.011	Female > Male	Higher adoption rate
Generative AI Literacy	4.07	.019	High-use group	Higher literacy scores
Ethical Sensitivity	3.65	.028	High-use group	Greater ethical awareness
Frequency of AI Use	3.55	.030	High-use group	More frequent engagement

Note. One-way ANOVA and Chi-square (χ^2) tests were used for group comparisons (two-tailed, $\alpha = .05$). Post-hoc comparisons were adjusted using Bonferroni correction to control for Type I error. All variables met normality assumptions (|skewness|, |kurtosis| < 2).

The ANOVA analysis presented in Table 4 also revealed significant differences across groups for Generative AI Literacy ($F(2,204) = 4.07$, $p = .019$), Ethical Sensitivity ($F(2,204) = 3.65$, $p = .028$), and Frequency of AI Use ($F(2,204) = 3.55$, $p = .030$). Participants who used AI tools frequently demonstrated

significantly higher literacy and ethical awareness scores than those who used them infrequently. These findings indicate that active engagement with AI systems enhances both technical understanding and moral reflection, which are essential components of informed and responsible AI interaction.

No significant group differences were observed for general attitudes toward AI, including positive and negative subdimensions ($p > .05$). Therefore, these variables were interpreted descriptively rather than inferentially. Overall, the results summarized in Table 4 confirm that demographic and behavioral characteristics meaningfully influence AI literacy, ethical awareness, and engagement patterns, reinforcing the view that frequent and conscious use of generative AI supports the development of critical and ethical competencies.

Multivariate prediction of perceived trust in reality

A hierarchical multiple regression analysis was conducted to identify the factors predicting perceived trust in reality (see Table 5). In the model, control variables including gender, age, education level, and frequency of AI use were entered in the first block. The second block included Generative AI Literacy and positive and negative attitudes toward AI, while the third block introduced Critical Inquiry Orientation as a metacognitive predictor. The overall model demonstrated a significant fit, $F(9,197) = 8.12$, $p < .001$, explaining 39% of the total variance ($R^2 = .39$), with an additional $\Delta R^2 = .12$ contributed by the inclusion of attitudinal and metacognitive variables. The results indicated that negative attitude toward AI ($\beta = -.45$, $p < .001$), positive attitude toward AI ($\beta = .31$, $p = .027$), and critical inquiry ($\beta = -.35$, $p < .001$) were the strongest predictors of perceived trust in reality.

Table 5
Multivariate prediction of perceived trust in reality (hierarchical regression, n = 207)

Variable	B	SE	β	t	p	95% CI (Lower–Upper)	VIF
(Constant)	51.05	4.64	—	11.01	< .001	[41.91, 60.20]	—
Evaluation	−0.47	0.32	−.15	−1.45	.150	[− 1.10, 0.17]	3.34
Ethics	−0.02	0.21	−.01	−0.07	.941	[− 0.44, 0.41]	1.33
Awareness	0.22	0.34	.05	0.63	.527	[− 0.46, 0.89]	1.89
Frequency of AI Use	0.34	0.48	.07	0.70	.486	[− 0.61, 1.29]	3.31
Negative Attitude toward AI	−0.55	0.07	−.45	−7.33	< .001	[− 0.69, − 0.40]	1.19
Positive Attitude toward AI	0.14	0.06	.31	2.23	.027	[0.02, 0.26]	3.08
Manipulation Awareness	−0.23	0.27	−.10	−0.84	.401	[− 0.76, 0.30]	4.50
Ethical and Regulatory Sensitivity	−0.04	0.31	.01	−0.12	.905	[− 0.66, 0.58]	4.12
Critical Inquiry	−0.19	0.18	−.35	−5.83	< .001	[− 1.44, − 0.71]	1.17

Model Summary. $R^2 = .39$, $\Delta R^2 = .12$, $F(9,197) = 8.12$, $p < .001$.
Note. Hierarchical multiple linear regression (two-tailed, $\alpha = .05$). All predictors met multicollinearity (VIF < 5) and normality assumptions (|skewness|, |kurtosis| < 2). Residuals were normally distributed and homoscedastic.

These findings suggest that users' trust in AI-mediated content is shaped by an interplay of affective and cognitive factors, particularly the tendency toward critical reflection and inquiry. While a negative attitude substantially reduces perceived trust, a positive attitude increases it moderately, and a strong critical inquiry orientation redefines trust more cautiously and reflectively.

Multicollinearity diagnostics indicated that the Variance Inflation Factor (VIF) values were below 5 for all predictors, confirming the absence of multicollinearity problems and supporting the stability of the regression coefficients. Multicollinearity diagnostics indicated that the Variance Inflation Factor (VIF) values were below 5 for all predictors, confirming the absence of multicollinearity problems and supporting the stability of the regression coefficients. Assumption checks confirmed model adequacy: the normal P–P plot and histogram indicated normally distributed residuals, and the scatterplot of standardized residuals versus predicted values showed no pattern, confirming homoscedasticity. Overall, the hierarchical model supports the interpretation that trust in AI-generated reality is neither purely emotional nor purely rational but instead emerges from the dynamic interaction among cognitive understanding, attitudinal orientation, and critical awareness, consistent with theoretical frameworks in media psychology and digital epistemology (Lewandowsky et al., 2017; Sun & Xu, 2025).

The hierarchical regression results presented in Table 5 reveal a significant, theoretically coherent model predicting perceived trust in reality in AI-mediated contexts. The model explained 39% of the total variance ($R^2 = .39$), with an additional 12% increase in explained variance ($\Delta R^2 = .12$) when attitudinal and metacognitive factors were included. Among all predictors, negative attitude toward AI emerged as the strongest negative predictor of perceived trust ($\beta = -.45$, $p < .001$), indicating that skepticism and affective discomfort toward AI substantially undermine users' confidence in the authenticity of mediated content. Conversely, a positive attitude toward AI had a modest but significant positive effect ($\beta = .31$, $p = .027$), suggesting that favorable emotional orientations toward AI may enhance trust, albeit to a limited extent.

The role of critical inquiry was particularly noteworthy ($\beta = -.35$, $p < .001$). This metacognitive variable negatively predicted perceived trust, implying that individuals who habitually question, verify, and critically analyze AI-generated content tend to adopt a more cautious stance toward mediated reality. Rather than diminishing trust entirely, this pattern indicates calibrated trust, a balanced cognitive posture in which users remain engaged yet vigilant. Other predictors, including awareness, ethics, and frequency of AI use, did not show significant direct effects but contributed indirectly to the model's robustness by contextualizing how cognitive, behavioral, and moral dimensions interact to shape epistemic trust. Multicollinearity diagnostics (all VIFs < 5) confirmed the independence of predictors, ensuring the stability and interpretability of the coefficients.

Discussion

This study examined how cognitive and attitudinal factors jointly shape perceived trust in reality when individuals encounter AI-generated content. Two results anchor the contribution. First, hierarchical regression showed that negative attitudes toward AI were the strongest inverse predictor of trust, while positive attitudes modestly increased trust. A stronger critical inquiry orientation also reduced uncalibrated trust, consistent with a more cautious and verification-oriented stance (Table 5). Second, the correlation matrix revealed that generative AI literacy covaried strongly with manipulation awareness, ethical sensitivity, and usage frequency, indicating that literacy functions as a broader competence that is simultaneously cognitive (knowledge of systems), reflective (awareness of risks), and behavioral (engagement with tools).

These patterns extend and synthesize several lines of evidence. The strong links between literacy and awareness align with frameworks that conceptualize AI literacy as a multidimensional set of competencies, understanding, use, evaluation, and ethics that are essential for responsible engagement with algorithmic systems (Ng & Leung, 2021). In higher education contexts, the prominence of literacy alongside usage suggests that frequent interaction with generative systems may develop in parallel with evaluative skills rather than merely produce familiarity. This interpretation is consistent with recent measurement studies and calls to assess generative-specific competencies (Jobin et al., 2019).

The attitudinal results help clarify why trust in AI-mediated information oscillates between confidence and skepticism in contemporary media environments. Previous research shows that AI-synthesized faces can be indistinguishable from real ones and even perceived as more trustworthy, undermining intuitive cues of authenticity (Gilbert & Gilbert, 2024). Similarly, exposure to deepfakes can diminish confidence in visual media overall, producing a broader "liar's dividend" of skepticism (Gillespie et al., 2023). Within this context, the finding that a negative attitude robustly lowers perceived trust supports existing evidence that public confidence in digital content remains fragile in post-truth settings where superficial realism outpaces verification (Cosentino, 2020; Floridi & Cowls, 2022). At the same time, a positive attitude exerts a more negligible, constructive influence, suggesting that favorable orientations toward AI cannot entirely override epistemic caution when other regulatory mechanisms, such as inquiry, are active.

Critical inquiry demonstrated an independent, negative association with trust, an effect best understood as calibration rather than disbelief. Research on misinformation indicates that lateral reading, active verification, and metacognitive checks reduce susceptibility without necessarily inducing cynicism (Jobin et al., 2019). The present findings mirror that pattern. Participants with higher inquiry tendencies were less inclined to grant default trust to AI-generated outputs, suggesting that labels or superficial cues alone have limited corrective value unless users possess sufficient literacy and engage in deeper processing (Floridi & Cowls, 2022). Overall, the results support a process model in which literacy fosters awareness and ethical reflection, which in turn shape attitudes and inquiry, ultimately resulting in calibrated trust

Theoretical and practical implications

The findings of this study provide both theoretical insight and practical direction for understanding and cultivating trust in AI-mediated communication. From a theoretical perspective, the results integrate cognitive competence and attitudinal orientation within a unified explanatory model of trust. Rather than viewing trust as a passive byproduct of technological acceptance or as an outcome of perceived risk, the evidence suggests that trust is an active and dynamic construct shaped by knowledge, ethics, and reflection. Individuals who understand how generative systems function and where they may fail are better positioned to evaluate AI outputs critically. Maintaining ethical sensitivity toward issues such as bias, transparency, and accountability further refines this evaluative process, while habitual inquiry practices help users verify claims before accepting them as accurate. This synthesis extends earlier frameworks that treated trust as either cognitive or affective by introducing a metacognitive layer that regulates how understanding and emotion jointly influence credibility judgments (Lewandowsky et al., 2017; Carvalho et al., 2022).

From a practical standpoint, the results highlight the importance of dual-track educational interventions that target both literacy and inquiry. In higher education and professional contexts, curricula should focus on strengthening generative AI literacy, including comprehension of model mechanisms, recognition of limitations, and evaluation of algorithmic outcomes, while simultaneously embedding

critical inquiry routines such as lateral reading, source triangulation, and provenance verification. Educational efforts that rely solely on content labeling or policy compliance are unlikely to build sustained trust. Instead, combining technical capability with reflective practice is essential for fostering informed confidence without creating cynicism. Institutions can also draw on AI ethics frameworks emphasizing fairness, transparency, and accountability (Morley et al., 2020; Floridi & Cowls, 2022) to cultivate reflective judgment. Encouraging awareness of data origins, algorithmic biases, and potential error patterns can strengthen epistemic resilience and enhance the public's capacity to engage critically and responsibly with generative AI systems.

Conclusions, limitations, and recommendations for future research

This study contributes to the growing body of literature on trust in AI-generated content by empirically demonstrating how cognitive competence, ethical sensitivity, and attitudinal orientation jointly shape individuals perceived trust in reality within AI-mediated environments. The results show that trust is not a static or purely affective response, but an emergent construct formed through the interaction of understanding, reflection, and emotional stance. Generative AI literacy, positive and negative attitudes, and critical inquiry collectively explain a significant proportion of the variance in perceived trust, supporting the notion that users' confidence in AI-mediated communication depends on both knowledge and evaluative judgment.

The findings also highlight the dual nature of trust in the post-truth era. While literacy and positive attitudes encourage engagement and confidence in generative systems, critical inquiry tempers this trust, fostering a form of cautious realism. This suggests that sustainable trust in AI requires a balance between openness to innovation and awareness of the risks of manipulation. In this sense, AI literacy serves not only as a cognitive skill but as a socio-ethical capacity that underpins responsible media consumption and content evaluation. Despite these contributions, several limitations should be acknowledged. First, the sample was limited to university students in Turkey, which restricts the generalizability of the results to broader populations. Future studies should employ more diverse and cross-cultural samples to capture variations in AI literacy and trust across social and cultural contexts. Second, the study used a cross-sectional design, preventing causal inference. Longitudinal or experimental research could better establish the directionality of relationships between literacy, attitude, inquiry, and trust. Third, self-reported measures may introduce social desirability or recall bias. Complementary approaches such as behavioral experiments, eye-tracking, or discourse analysis could provide deeper insights into how users process and evaluate AI-generated information.

Future research should also explore domain-specific differences in trust by examining contexts such as journalism, education, and healthcare, where the implications of AI-generated content are particularly consequential. Comparative analyses between human- and AI-authored media could reveal how perceived authorship interacts with trust and credibility cues. Moreover, integrating physiological or neurocognitive measures could enrich understanding of the emotional and attentional dynamics underlying trust formation.

In conclusion, this study provides a theoretically integrated and empirically grounded framework for understanding trust in AI-generated content. It underscores the need for educational and institutional interventions that strengthen AI literacy, promote ethical awareness, and cultivate inquiry-based skepticism. As generative AI continues to shape communication ecologies, fostering informed, reflective, and ethically aware users will be essential to maintaining epistemic stability and authenticity in the digital age.

Declarations

Competing interests

The authors declare no competing financial or non-financial interests related to the research, authorship, or publication of this article.

Funding Declaration

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

All data that support the findings of this study were generated from an online survey of human participants and contain potentially identifiable information at the record level. In accordance with the approved ethics protocol and to protect participant privacy, the de-identified dataset, codebook, and analysis syntax (SPSS 30) are available from the corresponding author upon reasonable request for bona fide research purposes and subject to a simple data-use agreement. For the purposes of peer review under double-blind conditions, the full instrument and variable map are provided in the Supplementary Information. Upon acceptance, the authors will deposit the de-identified dataset, codebook, and analysis files in an open repository that issues a persistent identifier (for example, OSF or Zenodo) and will update the manuscript with the repository citation and link.

Ethical approval statement

The study was approved by the institutional ethics committee (Approval Date: 26 May 2025, Decision No: 2025/11-07). The approval covers all procedures related to the administration of online surveys and the processing of anonymized participant data for research purposes. All research involving human participants was conducted in full accordance with the ethical standards of the institutional and national research committees and the 1964 Helsinki Declaration and its later amendments, as well as comparable international ethical standards. No part of the study involved deception, intervention, or the collection of personally identifiable information. Ethical approval applies specifically to survey-based quantitative research investigating attitudes, perceptions, and behavioral patterns related to generative AI and digital trust.

Informed consent statement

Informed consent was obtained electronically and voluntarily from all participants between 05 May and 15 June 2025, prior to data collection via the online survey platform (Google Forms). Before participation, each respondent was presented with a consent page detailing the purpose of the research, the voluntary nature of participation, assurances of anonymity and confidentiality, and the intended scientific use of the data. Participants confirmed consent by selecting an acknowledgment box before proceeding with the survey questions. No minors, vulnerable populations, or individuals requiring guardian consent were included in this study. The consent process and participant rights were implemented in accordance with institutional and national research ethics policies and the Declaration of Helsinki.

Author contributions

M.A.- Conceptualization, theoretical framework, research design, writing, and editing.
M.K.- Data collection, statistical analysis, interpretation, writing, and editing. Both authors contributed to the refinement of the research design, critically reviewed the manuscript for important intellectual content, and approved the final version of the manuscript.

References

1. Al-kfairy M, Mustafa D, Kshetri N et al (2024) Ethical Challenges and Solutions of Generative AI: An Interdisciplinary Perspective. Informatics 11(3):58. https://doi.org/10.3390/informatics11030058
2. Appelman A, Bien-Aimé S (2024) Artificial authorship and perceived credibility in news reporting. Computers in Human Behavior 153:107395. https://doi.org/10.1016/j.chbah.2024.100093
3. Bergdahl J, Latikka R, Celuch M et al (2023) Self-determination and attitudes toward artificial intelligence: Cross-national and longitudinal perspectives. Telematics and Informatics 82:102013. https://doi.org/10.1016/j.tele.2023.102013
4. Boediman EP (2025) Exploring the impact of deepfake technology on public trust and media manipulation: A scoping review. Jurnal Komunikasi 19(2):131-152. https://doi.org/10.20885/komunikasi.vol19.iss2.art8
5. Bowe BJ, Wohn DY (2015) Are there generational differences? social media use and perceived shared reality. In: Proceedings of the 2015 International Conference on Social Media & Society (SMSociety '15). p 1–5. https://doi.org/10.1145/2789187.2789200
6. Bulger M, Davison P (2018) The Promises, Challenges, and Futures of Media Literacy. Journal of Media Literacy Education 10(1):1-21. https://doi.org/10.23860/JMLE-2018-10-1-1
7. Carvalho L, Martinez-Maldonado R, Tsai Y-S et al (2022) How can we design for learning in an AI world? Computers and Education: Artificial Intelligence 3:100053. https://doi.org/10.1016/j.caeai.2022.100053

8. Cosentino G (2020) Social media and the post-truth world order: The global dynamics of disinformation. Palgrave Macmillan
9. Dame Adjin-Tettey T (2022) Combating fake news, disinformation, and misinformation: Experimental evidence for media literacy education. Cogent Arts & Humanities 9(1). https://doi.org/10.1080/23311983.2022.2037229
10. Epstein Z, Foppiani N, Hilgard S et al (2022) Do Explanations Increase the Effectiveness of AI-Crowd Generated Fake News Warnings?. Proceedings of the International AAAI Conference on Web and Social Media 16(1):183-193. https://doi.org/10.1609/icwsm.v16i1.19283
11. Fast E, Horvitz E (2017) Long-term trends in the public perception of artificial intelligence. In: Proceedings of the AAAI Conference on Artificial Intelligence 31(1):963–972
12. Flew T (2019) Digital communication, the crisis of trust, and the post-global. Communication Research and Practice 5(1):4–22. https://doi.org/10.1080/22041451.2019.1561394
13. Floridi L (2023) The ethics of artificial intelligence: Principles, challenges, and opportunities. Oxford University Press
14. Floridi L, Cowls J (2022) A Unified Framework of Five Principles for AI in Society. In: Carta S (ed) Machine Learning and the City. https://doi.org/10.1002/9781119815075.ch45
15. Gilbert C, Gilbert MA (2024) Navigating the Dual Nature of Deepfakes: Ethical, Legal, and Technological Perspectives on Generative Artificial Intelligence AI) Technology. International Journal of Scientific Research and Modern Technology 3(10). https://doi.org/10.38124/ijsrmt.v3i10.54
16. Gillespie N, Lockey S, Curtis C et al (2023) Trust in Artificial Intelligence: A global study. The University of Queensland; KPMG Australia, Brisbane, Australia; New York, United States. https://doi.org/10.14264/00d3c94
17. Gokcearslan S, Durak HY, Gunbatar MS et al (2024) Generative Artificial Intelligence (GenAI) Literacy Scale: Validity and Reliability. Paper presented at the 4th International Conference on Scientific and Academic Research, Konya, Türkiye, 19-20 July 2024
18. Greene D, Hoffmann AL, Stark L (2019) Better, nicer, clearer, fairer: A critical assessment of the movement for ethical artificial intelligence and machine learning. In: Proceedings of the 52nd Hawaii International Conference on System Sciences. p 2122-2131
19. Hoff KA, Bashir M (2014) Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust: Integrating Empirical Evidence on Factors That Influence Trust. Human Factors: The Journal of the Human Factors and Ergonomics Society 57(3):407-434. https://doi.org/10.1177/0018720814547570
20. Hohenstein J, Jung M (2020) AI as a moral crumple zone: The effects of AI-mediated communication on attribution and trust. Computers in Human Behavior 106:106190. https://doi.org/10.1016/j.chb.2019.106190
21. Ismatullaev UVU, Kim S-H (2022) Review of the Factors Affecting Acceptance of AI-Infused Systems. Human Factors: The Journal of the Human Factors and Ergonomics Society 66(1):126-144. https://doi.org/10.1177/00187208211064707

22. Jobin A, Ienca M, Vayena E (2019) The global landscape of AI ethics guidelines. Nature Machine Intelligence 1(9):389–399. https://doi.org/10.1038/s42256-019-0088-2
23. Jungwirth D, Haluza D (2023) Artificial Intelligence and the Sustainable Development Goals: An Exploratory Study in the Context of the Society Domain. Journal of Software Engineering and Applications 16:91-112. https://doi.org/10.4236/jsea.2023.164006
24. Kaya F, Aydin F, Schepman A et al (2022) The Roles of Personality Traits, AI Anxiety, and Demographic Factors in Attitudes toward Artificial Intelligence. International Journal of Human–Computer Interaction 40(2):497-514.
25. Kishnani D (2025) The Uncanny Valley: An Empirical Study on Human Perceptions of AI-Generated Text and Images. Dissertation, Massachusetts Institute of Technology
26. KPMG International (2023) Trust in artificial intelligence: 2023 global study on the shifting public perceptions of AI. https://kpmg.com/xx/en/our-insights/ai-and-technology/trust-in-artificial-intelligence.html
27. Lee S, Park G (2024) Development and validation of ChatGPT literacy scale. Current Psychology 43(21):18992-19004.
28. Lewandowsky S, Ecker UK, Cook J (2017) Beyond misinformation: Understanding and coping with the “post-truth” era. Journal of applied research in memory and cognition 6(4):353-369. https://doi.org/10.1016/j.jarmac.2017.07.008
29. Lin Z, Ng YL (2024) Unraveling Gratifications, Concerns, and Acceptance of Generative Artificial Intelligence. International Journal of Human–Computer Interaction 41(17):10725–10742. https://doi.org/10.1080/10447318.2024.2436749
30. Lobo Paes J (2024) Artificial Intelligence And News Consumption: A Study Of Trust, Credibility And Transparency In Automated Journalism. Dissertation, University of South Dakota. https://red.library.usd.edu/diss-thesis/255
31. Long D, Magerko B (2020) What is AI literacy? Competencies and design considerations. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. ACM, p 1–16. https://doi.org/10.1145/3313831.3376727
32. Luhmann N (2018) Trust and power. John Wiley & Sons
33. Lutz C, Baruh L, Quinn K et al (2025) Comparative Approaches to Studying Privacy: Introduction to the Special Issue. Social Media + Society 11(2). https://doi.org/10.1177/20563051251344460
34. Metzger M, Flanagin A, Mena P et al (2021) From Dark to Light: The Many Shades of Sharing Misinformation Online. Media and Communication 9(1):134-143. https://doi.org/10.17645/mac.v9i1.3409
35. Mihailidis P, Viotty S (2017) Spreadable Spectacle in Digital Culture: Civic Expression, Fake News, and the Role of Media Literacies in “Post-Fact” Society. American Behavioral Scientist 61(4):441-454. https://doi.org/10.1177/0002764217701217
36. Morley J, Floridi L, Kinsey L et al (2020) From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. Sci Eng Ethics 26:2141–

2168. https://doi.org/10.1007/s11948-019-00165-

37. Müller VC, Bostrom N (2016) Future Progress in Artificial Intelligence: A Survey of Expert Opinion. In: Müller VC (ed) Fundamental Issues of Artificial Intelligence. Synthese Library, vol 376. Springer, Cham. https://doi.org/10.1007/978-3-319-26485-1_33
38. Napoli PM (2019) Social media and the public interest: Media regulation in the disinformation age. Columbia University Press
39. Ng DTK, Leung JKL, Chu SKW et al (2021) Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence 2:100041. https://doi.org/10.1016/j.caeai.2021.100041
40. Nightingale SJ, Farid H (2022) AI-synthesized faces are indistinguishable from real faces and more trustworthy. Proceedings of the National Academy of Sciences 119(8):e2120481119. https://doi.org/10.1073/pnas.2120481119
41. Panda S, Roy ST (2024) Reflections on emerging HCI–AI research. AI & Soc 39:407–409. https://doi.org/10.1007/s00146-022-01409-y
42. Pennycook G, Rand DG (2021) The psychology of fake news. Trends in Cognitive Sciences 25(5):388–402. https://doi.org/10.1016/j.tics.2021.02.007
43. Sahebi S, Formosa P (2025) The AI-mediated communication dilemma: epistemic trust, social media, and the challenge of generative artificial intelligence. Synthese 205:128. https://doi.org/10.1007/s11229-025-04963-2
44. Shin D (2021) The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. International Journal of Human-Computer Studies 146:102551. https://doi.org/10.1016/j.ijhcs.2020.102551
45. Sun Y, Xu C (2025) Hybrid cognitive authority and algorithmic subjectivity: rethinking knowledge management in AI-driven communication. Journal of Knowledge Management. https://doi.org/10.1108/JKM-02-2025-0252
46. Tsamados A, Aggarwal N, Cowls J et al (2022) The ethics of algorithms: key problems and solutions. AI & Soc 37:215–230. https://doi.org/10.1007/s00146-021-01154-8
47. 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
48. Van de Vijver F, Leung K (1997) Methods and data analysis of comparative research. In: Handbook of cross-cultural psychology, vol 1. p 257-300
49. Wineburg S, McGrew S (2019) Lateral Reading and the Nature of Expertise: Reading Less and Learning More When Evaluating Digital Information. Teachers College Record 121(11):1-40. https://doi.org/10.1177/016146811912101102
50. Wischnewski M, Krämer N, Müller E (2023) Measuring and understanding trust calibrations for automated systems: A survey of the state-of-the-art and future directions. In: Proceedings of the

2023 CHI Conference on Human Factors in Computing Systems (CHI '23). Article 755. https://doi.org/10.1145/3544548.3581197

51. Zhang B, Dafoe A (2019) Artificial Intelligence: American Attitudes and Trends. http://dx.doi.org/10.2139/ssrn.3312874
52. Zhou Y, Lu X, Huang G (2023) Calibrating human–AI trust through ethical reflection: Evidence from a mixed-methods study. AI & Society 38(2):611–627. https://doi.org/10.1007/s00146-022-01462-7

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryDataAITrustStudyeng.xlsx