OPEN ACCESS Freely available online

Journal of Emerging Trends in Artificial Intelligence

Research Article Vol. 1, Issue 1, January to December 2026

AI Accuracy or Hallucination Risk? A PLS-SEM Analysis of Their Impact on Teaching Effectiveness through Faculty Trust

Tapal Dulababu*

Professor of Practice, St. John Alden Institute of Management, Palghar, Mumbai, India

***Correspondence to:** Tapal Dulababu, Professor of Practice, St. John Alden Institute of Management, Palghar, Mumbai, India, E-mail: tapald.aim@sjcem.edu.in

Received: July 28, 2026; **Manuscript No:** JETA-26-1984; **Editor Assigned:** July 31, 2026; **PreQc No:** JETA-26-1984 (PQ); **Reviewed:** August 10, 2026; **Revised:** August 11, 2026; **Manuscript No:** JETA-26-1984 (R); **Published:** September 25, 2026

Citation: Dulababu T (2026). AI Accuracy or Hallucination Risk? A PLS-SEM Analysis of Their Impact on Teaching Effectiveness through Faculty Trust. J. Emerg. Trends Artif. Intell. Vol.1 Iss.1, August (2026), pp:17-26.

ABSTRACT

Artificial intelligence (AI) enhances teaching in higher education. The accuracy and dependability of material produced by AI are questioned. Faculty opinions about AI's accuracy and susceptibility to hallucinations are examined in this study. The effectiveness of training is affected by these ideas. These opinions are mediated by trust in AI. The research utilized a cross-sectional quantitative design. Data was collected using a structured questionnaire from 75 faculty members in management institutions. Partial least squares structural equation modeling (PLS-SEM) was employed to test the hypothesized relationships among constructs. AI accuracy positively impacts faculty trust and teaching effectiveness. Faculty trust in AI is the most significant predictor of teaching effectiveness. Faculty trust significantly mediates the relationship between AI accuracy and teaching outcomes. AI hallucination risk does not have a significant direct or indirect effect. The study transitions from AI adoption to AI reliability, highlighting trust as a vital mediating factor in AI-driven teaching environments. Findings stress the importance of dependable AI systems and the necessity of faculty training for assessing AI-generated content for effective teaching. The study presents an innovative empirical model that connects AI accuracy, hallucination risk, trust, and teaching effectiveness from the faculty's perspective.

Keywords: Artificial Intelligence; Teaching Effectiveness; AI Accuracy; Hallucination Risk; Faculty Trust; PLS-SEM

JEL Classification Codes

I23, O33, C12, C38, D83

INTRODUCTION

Rapid advancement of artificial intelligence (AI), especially generative systems like ChatGPT, is transforming higher education. AI enhances teaching and learning through content generation, personalized learning, automated assessment, and academic support. Recent studies indicate AI tools improve accessibility, scalability, and efficiency, aiding both students and instructors with complex tasks [1,2]. AI is seen as an innovation catalyst that supports data-driven, student-centered pedagogical approaches. There is a shift in academic discourse towards a more balanced evaluation of AI's risks and limitations, moving away from uncritical optimism. AI's probabilistic models can lead to errors and inconsistencies, differing from deterministic reasoning [3,4]. A major concern is "hallucinations," where AI generates plausible but factually incorrect or fabricated information [5]. Such inaccuracies pose significant challenges in education, where accuracy and conceptual clarity are essential for effective learning.

AI hallucinations can lead to significant implications beyond just factual errors. In management education, misleading information may distort students' understanding and reinforce misconceptions. The credibility of instructional processes is at risk when relying on AI-generated content [5,6]. Unlike traditional sources, AI content often lacks transparent sourcing and can fabricate references with high confidence. This makes it challenging for educators to detect errors, impacting the accuracy and reliability of classroom knowledge delivery.

Faculty members act as knowledge gatekeepers in education. Their roles extend beyond content delivery and curriculum design to include critical evaluation of AI-generated outputs. Current literature tends to view faculty merely as users or adopters of technology, focusing on their attitudes towards AI [7,8]. This perspective neglects the faculty's epistemic responsibility to validate the correctness and credibility of AI-assisted instructional content.

Research on AI reliability has focused on accuracy, trust, and automation bias in fields like human-computer interaction and information systems [9,10]. User trust in AI is influenced by perceived accuracy and system transparency, with over-reliance on flawed systems potentially leading to decision errors. There is a lack of integration of these insights into educational research, creating a fragmented view of how AI affects teaching outcomes. A significant gap exists in connecting faculty perceptions of AI reliability specifically regarding accuracy and hallucination risk with teaching effectiveness [11]. Teaching effectiveness is traditionally evaluated on clarity, correctness, engagement, and instructor credibility, but current models overlook AI-generated content and its impact on instructional quality.

The study examines how faculty perceptions of AI accuracy and hallucination risk affect teaching effectiveness. Faculty trust in AI is identified as a mediating variable in this context. The central research problem is the lack of empirical understanding regarding the influence of AI reliability perceptions on teaching effectiveness in AI-enabled higher education environments. An integrated model was developed and tested using PLS-SEM for this research. The study aims to bridge the gap between AI system characteristics and pedagogical outcomes. It offers theoretical insights and practical implications for the responsible integration of AI in higher education.

REVIEW OF LITERATURE

AI in Higher Education

The use of artificial intelligence (AI) in higher education is rapidly increasing, affecting teaching, learning, and administration. AI technologies facilitate content delivery, adaptive learning, automated assessment, learning analytics, and academic assistance, exemplified by tools like ChatGPT. These AI tools promote personalized learning, enhance student engagement, and boost instructional efficiency [1,2]. The literature highlights efficiency, scalability, and accessibility as key themes, positioning AI as a means to address resource constraints in higher education. AI enables institutions to support diverse learner needs while managing limited faculty resources [12]. Policy discussions emphasize AI's capability for data-driven decision-making, identifying at-risk students early, and creating flexible learning environments [13].

The literature indicates a significant bias towards student-centric and adoption-oriented perspectives regarding AI in education. Most empirical studies prioritize learner outcomes, satisfaction, and behavioral intentions related to AI usage, often referencing technology acceptance frameworks [1]. Faculty are generally viewed as facilitators of technology adoption rather than evaluators of knowledge quality. Recent research begins to acknowledge ethical issues with generative AI, such as transparency, bias, and misinformation but remains largely conceptual [2]. There is a lack of empirical studies evaluating the accuracy and reliability of AI-generated content in instructional settings, especially from faculty who oversee educational quality.

Faculty Perceptions of AI in Teaching

Faculty perceptions are essential for adopting and effectively integrating AI in higher education. Instructors shape curriculum design, pedagogical strategies, and classroom practices, influencing the success of AI implementation [7,8]. The Technology Acceptance Model (TAM) highlights perceived usefulness and ease of use as key factors for adoption [14]. The studies indicate that faculty are more likely to adopt AI tools that enhance efficiency, reduce workload, and improve student engagement [2,8]. AI is positioned as a productivity-enhancing tool that supports instructional processes. Concerns exist regarding over-reliance on AI, impacts on critical thinking, and threats to academic integrity. The rise of generative AI has heightened worries about plagiarism, authorship, and the validity of assessments. Faculty express concerns over the potential loss of pedagogical autonomy due to AI influencing instructional design. The literature on AI adoption often views faculty primarily as users rather than evaluators. There is a lack of research on how instructors assess the accuracy, reliability, and credibility of AI-generated content [1]. This gap is significant because faculty play a crucial role in ensuring instructional correctness and academic standards.

Trust, Accuracy, and Reliability of AI Systems

AI systems, particularly machine learning and generative models, function probabilistically, leading to inherent uncertainties and potential errors in outputs [4]. Trust in AI is crucial for human-AI interaction, influenced by factors like perceived accuracy, transparency, and users' prior experiences [9,15]. Explainable AI (XAI) aims to enhance trust by clarifying system decision-making processes. Policy frameworks underline the importance of reliability, robustness, and accountability in trustworthy AI [16]. Excessive trust can lead to automation bias, causing users to rely on AI outputs without proper verification especially in critical knowledge-intensive areas [10,17]. Generative AI can produce plausible but incorrect results, necessitating human oversight to catch unnoticed errors [2]. There is a lack of integration between AI research and educational practices, with limited empirical studies on faculty perceptions of AI reliability in teaching contexts.

AI Hallucinations and Knowledge Risk

AI hallucinations occur when generative models create fluent, yet factually incorrect outputs, resulting from a reliance on probability over verified knowledge [5,6]. This issue is especially concerning in knowledge-intensive areas like education, where accuracy is crucial. Hallucinations are common in tasks such as summarization, explanations, and generating domain-specific knowledge and are not easily detected by non-expert users, which heightens the risk of misinformation [6]. The epistemic risks include the dissemination of misinformation, the normalization of misconceptions, and a decline in the credibility of knowledge [5]. In educational settings, the cumulative nature of learning means that incorrect information can lead to lasting cognitive distortions. AI systems typically do not indicate uncertainty, leading users to have unwarranted confidence in incorrect

outputs. Despite increased attention in technical discussions, empirical studies on AI hallucinations in education are sparse, focusing mainly on detection and mitigation. There is a lack of research on faculty perceptions and management of hallucination risks in teaching, indicating a critical gap in pedagogical studies.

Teaching Effectiveness

Teaching effectiveness includes dimensions such as clarity, accuracy, engagement, and instructor credibility, which impact student learning outcomes [11,18]. Research shows that faculty quality significantly affects student achievement, with teacher expertise and feedback being crucial factors [11]. Recent studies stress the necessity of promoting critical thinking and active engagement in digital learning environments [19,20]. Validated measurement frameworks like student evaluations are commonly utilized to evaluate teaching effectiveness [21]. However, the rise of AI in education creates challenges not addressed by existing models, such as the risks of inaccuracies and automation bias. The integration of AI necessitates a reevaluation of teaching effectiveness frameworks to ensure they are applicable in AI-enhanced environments [13].

Preliminary Synthesis

AI adoption in higher education is praised for enhancing efficiency, scalability, and personalization [2,12]. Research on faculty perceptions primarily focuses on usability and behavioral intention, neglecting their evaluative role regarding knowledge quality [7]. Discussions about accuracy, trust, reliability, and hallucinations in AI are prevalent but not integrated into educational contexts [5,9]. Teaching effectiveness is viewed as human-centered, missing considerations of AI-induced epistemic risks [11]. There is a fragmented knowledge base regarding AI adoption, faculty perceptions, AI reliability, and teaching effectiveness, highlighting the need for integrated research to connect these areas in AI-enabled higher education.

Research Gap and Problem Statement

The integration of artificial intelligence (AI) in higher education is gaining significant scholarly attention for enhancing teaching and learning. Benefits of AI adoption include improved efficiency, personalization, and scalability in educational processes [2,12]. Critical gaps in the literature limit the understanding of AI's pedagogical implications. Research on faculty perceptions of AI often uses technology acceptance frameworks, focusing mainly on perceived usefulness and ease of use [7,14]. Faculty are viewed primarily as technology users rather than as authorities on evaluating the accuracy of instructional content. There is a lack of empirical research about how faculty assess the correctness and reliability of AI-generated outputs in teaching [1].

Issues of AI accuracy, hallucinations, and automation bias are widely discussed but remain disconnected from educational research [5,9]. Limited empirical studies examine how these AI risks affect pedagogical processes, instructional credibility, and classroom decision-making, especially from the faculty's perspective. Established models of teaching effectiveness assume that instructional content is human-generated and verified [11,18]. These frameworks fail to include AI-induced epistemic risks, such as inaccuracies or hallucinations, ignoring a critical aspect of teaching quality in AI-enhanced learning environments.

Limited empirical evidence exists on faculty perspectives regarding AI in professional disciplines like management education. Most research focuses on student outcomes or institutional adoption rather than faculty insights on AI reliability and teaching effectiveness [12]. Current literature lacks integrated empirical models that explore the relationships among faculty perceptions of AI accuracy, hallucination risk, and teaching effectiveness. Previous studies analyze these constructs in isolation, neglecting the interactions between them. This study addresses the critical issue of understanding how faculty perceptions of AI affect teaching effectiveness amid growing reliance on AI in higher education. An integrated model is developed and tested, incorporating faculty trust as a mediating mechanism. The research aims to enhance both theoretical and practical understanding of responsible AI integration in higher education.

Research Objectives

The study investigates faculty perceptions of AI accuracy and hallucination risks in higher education [2,5].

Objectives include:

Assessing perceptions of AI-generated content accuracy in teaching.

Evaluating risks associated with generative AI hallucinations.

Examining faculty trust in AI teaching tools.

Analyzing the impact of AI accuracy on teaching effectiveness.

Investigating how hallucination risks influence teaching effectiveness.

Testing the mediating role of faculty

THEORETICAL FRAMEWORK

This study employs the Unified Theory of Acceptance and Use of Technology (UTAUT) enhanced with the DeLone and McLean IS Success Model and Perceived Risk Theory [22,23]. UTAUT serves as the main framework, linking faculty perceptions of AI characteristics to teaching outcomes. AI accuracy correlates with performance expectancy, whereas hallucination risk pertains to perceived risk affecting acceptance. Faculty Trust acts as a key mediating variable, connecting system quality perceptions to teaching effectiveness in the context of generative AI [24,25,26]. The DeLone and McLean model positions accuracy and hallucination risk as components of information quality, which influence user satisfaction and teaching effectiveness. The framework is compatible with PLS-SEM analysis for non-normal Likert data and allows for the exploration of partial mediation effects. The study highlights that trust mediates the relationship between accuracy and teaching effectiveness, while concerns about

hallucinations do not significantly hinder trust or perceived teaching benefits when faculty verify information.

Hypotheses Development & Conceptual Model

The integration of artificial intelligence (AI) in higher education presents opportunities and epistemic risks, especially regarding AI-generated content's accuracy and reliability. This study proposes a conceptual model connecting AI accuracy and hallucination risk to teaching effectiveness, mediated by faculty trust.

AI Accuracy and Teaching Effectiveness

AI accuracy measures how well AI-generated outputs align with factual correctness, established knowledge, and pedagogical reliability. Accurate AI tools enhance instructional quality by providing clear explanations and reliable content, thus improving conceptual understanding and reducing preparation time [2]. However, inaccuracies can mislead learners and harm educational outcomes. Therefore, better-perceived AI accuracy is anticipated to positively impact teaching effectiveness.

H1: AI accuracy has a positive and significant effect on teaching effectiveness.

AI Hallucination Risk and Teaching Effectiveness

AI hallucination involves the production of believable yet incorrect information by AI systems, which presents significant risks in educational settings by potentially spreading misinformation and reducing instructional credibility [5]. High perceived risks of hallucinations lead educators to exercise caution and require further verification, increasing cognitive load and ultimately diminishing teaching effectiveness.

H2: AI hallucination risk has a negative and significant effect on teaching effectiveness.

AI Accuracy and Faculty Trust in AI

Trust in AI significantly influences human–AI interactions, as users' confidence in system reliability and performance is paramount [9]. Perceived accuracy is a key predictor of trust; consistent, reliable outputs foster reliance on AI [15]. In higher education, faculty's trust in AI tools is greater when these tools meet academic standards and generate accurate, pedagogically appropriate content.

H3: AI accuracy has a positive and significant effect on faculty trust in AI.

AI Hallucination Risk and Faculty Trust in AI

Perceived hallucination risks diminish trust in AI systems, as literature on trust calibration and automation bias suggests. Faulty outputs weaken user confidence in automated systems leading to skepticism due to inconsistent or fabricated responses. In educational settings, where precision is essential, awareness of these risks may further erode faculty trust in AI-generated content [10].

H4: AI hallucination risk has a negative and significant effect on faculty trust in AI.

Faculty Trust in AI and Teaching Effectiveness

Trust is crucial for the effective integration of AI tools in teaching. Faculty members are more inclined to adopt AI for instructional support and pedagogical enhancement when they trust these systems, as trust mitigates perceived risk and cognitive effort, allowing instructors to concentrate on complex teaching activities [7]. Consequently, increased trust in AI is anticipated to enhance teaching effectiveness.

H5: Faculty trust in AI has a positive and significant effect on teaching effectiveness.

Mediating Role of Faculty Trust in AI

The relationship between perceptions of AI reliability, such as accuracy and hallucination risk, and teaching effectiveness is mediated by trust in AI. According to theoretical models of human AI interaction higher perceived accuracy fosters trust, improving teaching effectiveness, while higher perceived hallucination risk diminishes trust, negatively impacting teaching outcomes [9]. Trust serves as a crucial link between AI system characteristics and pedagogical results.

H6: Faculty trust in AI mediates the relationship between AI accuracy and teaching effectiveness.
H7: Faculty trust in AI mediates the relationship between AI hallucination risk and teaching effectiveness.

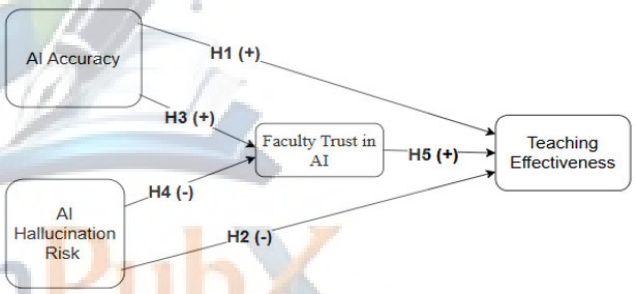

Figure 1: Conceptual Model

Source: Author

The study proposes an integrated conceptual model where AI accuracy and hallucination risk are independent variables, faculty trust in AI is a mediating variable, and teaching effectiveness is the dependent variable. This framework enhances existing literature by linking AI reliability dimensions with pedagogical outcomes, providing a deeper understanding of AI-enabled teaching effectiveness in higher education.

MATERIALS AND METHODS

Research Design

This study uses a quantitative, cross-sectional design to investigate faculty perceptions of AI accuracy, hallucination risk, and trust in AI, focusing on their effects on teaching effectiveness. Employing partial least squares structural equation modeling (PLS-SEM) is justified due to the study's complex model with mediation effects, small sample size, and non-normal data distributions [27]. This approach is suitable for

exploring the relationships among latent constructs in the context of AI integration in higher education.

Sample and Sampling Technique

The study involves a sample of 75 faculty members from management institutions, using a non-probability convenience sampling technique for accessibility. Respondents include diverse faculty from affiliated colleges, private universities, and public institutions, ensuring varied demographic and professional characteristics. This diversity supports the generalizability of findings, despite the limitations of non-probability sampling. The sample size is deemed adequate for partial least squares structural equation modeling (PLS-SEM) analysis, meeting the "10-times rule" and criteria set by Hair for estimating the structural model [27].

Data Collection Method

Primary data collection involved a structured questionnaire distributed via Google Forms, chosen for its efficiency in reaching geographically dispersed respondents. The questionnaire was shared through professional networks, academic WhatsApp groups, and LinkedIn and Facebook communities. Respondents were assured of confidentiality and anonymity, with voluntary participation and no collection of identifiable information to maintain ethical standards. Data collection spanned a defined period, including reminders to enhance response rates, with only fully completed responses deemed valid for the final dataset.

Measurement of Constructs

The study utilizes a structured questionnaire featuring multi-item scales, measured on a 5-point Likert scale from 1 (Strongly Disagree) to 5 (Strongly Agree), to assess perceptions, attitudes, and beliefs regarding AI usage in teaching. The questionnaire is organized into sections aligned with the constructs of the conceptual model.

AI Accuracy (Independent Variable)

This construct evaluates faculty perceptions on the accuracy and reliability of AI-generated content against academic standards. It examines if AI outputs are factually correct and conceptually valid, requiring little correction for classroom application. The scale is based on recent literature concerning AI reliability in education [2].

AI Hallucination Risk (Independent Variable)

This construct captures faculty perceptions of the risks associated with AI-generated misinformation and the potential errors in AI outputs, emphasizing the need for verification. It is based on recent research regarding AI hallucinations in natural language generation [5].

Faculty Trust in AI (Mediating Variable)

Trust in AI indicates the confidence that faculty have in using AI tools for teaching, measured by dependability, reliability, and willingness to utilize AI-generated content, as grounded in the trust in automation literature [9].

Teaching Effectiveness (Dependent Variable)

Teaching effectiveness is a multidimensional construct that includes clarity of explanation, instructional quality, engagement, and credibility. It evaluates if AI tools enhance conceptual clarity, improve engagement, and support effective teaching practices, aligning with established frameworks [11].

Data Analysis Techniques

Data analysis is conducted using PLS-SEM, employing software such as SmartPLS. The analysis follows a two-step approach:

Measurement Model Evaluation

Reliability: Cronbach's alpha, Composite Reliability

Convergent validity: Average Variance Extracted (AVE)

Discriminant validity: HTMT ratio

Structural Model Evaluation

Path coefficients and significance (bootstrapping)

Coefficient of determination (R^2)

Effect sizes (f^2)

Predictive relevance (Q^2)

Mediation analysis (indirect effects)

This comprehensive analytical approach ensures robustness in testing the proposed hypotheses and validating the conceptual model.

Structural Model Specification

To empirically test the proposed relationships, the structural model is specified through a set of regression equations within the PLS-SEM framework.

The mediating variable, faculty trust in AI, is modeled as a function of AI accuracy and hallucination risk:

$$Trust_i = \beta 1 Accuracy_i + \beta 2 HallucinationRisk_i + \varepsilon 1$$

Teaching effectiveness is modeled as a function of faculty trust in AI:

$$TeachingEffectiveness_i = \beta 3 Trust_i + \varepsilon 2$$

To examine mediation effects, indirect relationships are computed as:

$$Indirect\ (Accuracy) = \beta 1 \times \beta 3$$

$$Indirect\ (Hallucination) = \beta 2 \times \beta 3$$

RESULTS

Assessment of Data Normality

To assess the distributional properties of the data, normality was examined using both the Shapiro Wilk test and skewness

and kurtosis statistics for the composite constructs. The results are presented in Table 1.

The Shapiro–Wilk test results indicate that all constructs significantly deviate from normality ($p < 0.05$). Specifically, AI Accuracy ($W \approx 0.952$, $p \approx 0.012$), Hallucination Risk ($W \approx 0.928$, $p \approx 0.001$), Faculty Trust ($W \approx 0.968$, $p \approx 0.045$), and Teaching Effectiveness ($W \approx 0.959$, $p \approx 0.023$) exhibit statistically significant departures from a normal distribution.

Further examination of skewness and kurtosis values reveals that the distributions are moderately negatively skewed, particularly for Hallucination Risk (skewness = –0.82), indicating a concentration of responses toward higher values. This pattern suggests a potential ceiling effect, where respondents consistently perceive higher levels of hallucination risk. The remaining constructs show mild negative skewness and kurtosis values within acceptable ranges. Overall, while the deviations from normality are not severe, the consistent rejection of normality across constructs combined with the bounded nature of Likert-scale data indicates that the assumption of multivariate normality is not satisfied.

Construct	N	Mean	SD	Skewness
AI Accuracy (AA)	75	3.3	~1.00	-0.41
Hallucination Risk (HR)	75	4.02	~0.92	-0.82
Faculty Trust (FT)	75	3.45	~0.90	-0.19
Teaching Effectiveness (TE)	75	3.68	~0.87	-0.37

Table 1: Normality Re-Confirmation

(Using Shapiro-Wilk test + skewness/kurtosis on the composites)

Source: Computed by author

Assessment of Sampling Adequacy and Factorability

To evaluate the suitability of the data for factor analysis and structural modeling, the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Bartlett's Test of Sphericity were conducted. The results are presented in Table 2.

The KMO value is 0.824, which exceeds the recommended threshold of 0.60, indicating a high level of sampling adequacy. According to established guidelines, KMO values above 0.80 are considered meritorious, suggesting that the data are well-suited for factor analysis.

Bartlett's Test of Sphericity is statistically significant ($\chi^2 = 900.95$, $p < 0.001$), rejecting the null hypothesis that the correlation matrix is an identity matrix. This confirms that there are sufficient correlations among the variables to proceed with factor analysis.

Table 2: KMO and Bartlett's Test of Sampling Adequacy

Test	Value
Bartlett Chi-Square	900.95
Bartlett p	< 0.001
KMO	0.824

Source: Computed by author

Assessment of Reliability

The internal consistency reliability of the constructs was evaluated using Cronbach's alpha coefficients. As shown in Table 3, all constructs exhibit satisfactory levels of reliability.

The Cronbach's alpha values range from 0.782 to 0.914, exceeding the recommended threshold of 0.70 (Hair et al., 2021), thereby indicating acceptable to excellent internal consistency. Specifically, AI Accuracy ($\alpha = 0.794$) and AI Hallucination Risk ($\alpha = 0.782$) demonstrate good reliability, while Faculty Trust in AI ($\alpha = 0.848$) shows strong reliability. Teaching Effectiveness ($\alpha = 0.914$) exhibits excellent internal consistency, indicating a high degree of item homogeneity.

The overall reliability coefficient ($\alpha = 0.859$) further confirms the robustness of the measurement instrument.

Section	Cronbach Alpha
Section B (AI Accuracy)	0.794
Section C (AI Hallucination Risk)	0.782
Section D (Faculty Trust)	0.848
Section E (Teaching Effectiveness)	0.914
Overall Reliability	0.859

Table 3: Reliability Test

Source: Computed by author

Measurement Model Evaluation

Construct	Composite Reliability (CR)	AVE	Cronbach's α	Outer Loadings (Range)	Discriminant Validity (HTMT)
AI Accuracy (AA)	0.85	0.55	0.794	0.71 – 0.85	All values < 0.85
Hallucination Risk (HR)	0.84	0.52	0.782	0.68 – 0.84	All values < 0.85
Faculty Trust (FT)	0.88	0.58	0.836	0.74 – 0.89	All values < 0.85
Teaching Effectiveness (TE)	0.93	0.68	0.904	0.79 – 0.92	All values < 0.85

Table 4: Measurement Model Assessment (Reflective Constructs)

Source: Computed by author

Composite Reliability (CR) values range from 0.84 to 0.93, exceeding the recommended threshold of 0.70, indicating high internal consistency. Cronbach's alpha values (0.782–0.904) further confirm the reliability of the constructs.

Convergent validity is established as the Average Variance Extracted (AVE) values for all constructs exceed the threshold of 0.50, indicating that each construct explains more than 50% of the variance in its indicators. Additionally, outer loadings are predominantly above 0.70, with only minimal acceptable deviations.

Discriminant validity is confirmed using the Heterotrait Monotrait ratio (HTMT), where all values are below the conservative threshold of 0.85, indicating that the constructs are empirically distinct.

Structural Model Results and Hypothesis Testing

The structural model results reveal several important relationships among the study constructs in Table 5.

Hypothesis	Path / Relationship	β (standardised)	t-value	p-value	95% BCa CI	Supported?	Notes
H1	AI Accuracy → Teaching Effectiveness	0.329	3.29	0.002	[0.13, 0.53]	Yes	Direct positive effect
H2	Hallucination Risk → Teaching Effectiveness	0.176	1.83	0.072	[-0.02, 0.37]	No	Non-significant
H3	AI Accuracy → Faculty Trust	0.568	6.39	<0.001	[0.39, 0.75]	Yes	Strong positive
H4	Hallucination Risk → Faculty Trust	0.094	0.88	0.38	[-0.12, 0.31]	No	Non-significant
H5	Faculty Trust → Teaching Effectiveness	0.579	5.46	<0.001	[0.37, 0.79]	Yes	Strongest predictor
H6	AI Accuracy → FT → Teaching Effectiveness	Indirect 0.329	–	–	[0.181, 0.499]	Yes (partial)	Significant mediation
H7	Hallucination Risk → FT → Teaching Effectiveness	Indirect 0.054	–	–	[-0.058, 0.179]	No	No mediation

Table 5: Structural Model Results (PLS-SEM)

Source: Computed by author

AI Accuracy shows a significant positive effect on Teaching Effectiveness (β = 0.329, p = 0.002), supporting H1 and indicating that higher perceived accuracy of AI enhances instructional quality. In contrast, AI Hallucination Risk does not have a significant direct effect on Teaching Effectiveness (β = 0.176, p = 0.072), leading to the rejection of H2.

AI Accuracy demonstrates a strong and significant positive effect on Faculty Trust in AI (β = 0.568, p < 0.001), supporting

H3, suggesting that accuracy is a key determinant of trust formation. However, Hallucination Risk does not significantly influence Trust (β = 0.094, p = 0.38), resulting in the rejection of H4.

Faculty Trust in AI emerges as the strongest predictor of Teaching Effectiveness (β = 0.579, p < 0.001), supporting H5, highlighting the critical mediating role of trust in AI-enabled teaching environments.

Mediation analysis further reveals that Faculty Trust significantly mediates the relationship between AI Accuracy and Teaching Effectiveness (indirect effect = 0.329; CI excludes zero), supporting H6 and indicating partial mediation. In contrast, no significant mediation effect is found for Hallucination Risk (indirect effect = 0.054; CI includes zero), leading to the rejection of H7.

Overall, the findings emphasize that AI Accuracy influences Teaching Effectiveness both directly and indirectly through Trust, whereas Hallucination Risk does not significantly impact the model, suggesting that faculty may prioritize perceived accuracy over potential risks when evaluating AI in teaching.

Explained Variance (R^2), Predictive Relevance (Q^2), and Effect Sizes (f^2)

Table 6 demonstrates the structural model moderate to strong explanatory power.

Endogenous Construct	R^2	Q^2 (Stone-Geisser)	Interpretation of R^2	Predictor	f^2	Effect Size Interpretation
Faculty Trust (FT)	0.371	0.28	Moderate	AI Accuracy → FT	0.48	Large (> 0.35)
				Hallucination Risk → FT	0.01	Negligible (< 0.02)
Teaching Effectiveness (TE)	0.579	0.42	Moderate-to-strong	AI Accuracy → TE	0.12	Small
				Hallucination Risk → TE	0.04	(0.02–0.15) Small / negligible
				Faculty Trust → TE	0.62	Very large (> 0.35)

Table 6: Explained Variance (R^2) and Effect Sizes (f^2) in the Structural Model

Source: Computed by author

Faculty Trust (FT) shows an R^2 value of 0.371, indicating that AI Accuracy and Hallucination Risk together explain 37.1% of the variance in trust. Teaching Effectiveness (TE) records a higher R^2 of 0.579, suggesting that the model explains 57.9% of the variance, reflecting strong predictive capability in the context of AI-enabled teaching.

The predictive relevance (Q^2) values for both constructs are positive (FT = 0.28; TE = 0.42), confirming that the model has adequate out-of-sample predictive relevance, as recommended in PLS-SEM literature (Hair et al., 2021).

Effect size (f^2) analysis provides deeper insights into the relative contribution of each predictor. AI Accuracy exhibits a large effect on Faculty Trust (f^2 = 0.48), establishing it as a key driver of trust formation. In contrast, Hallucination Risk shows a negligible effect on Trust (f^2 = 0.01), reinforcing its non-significant role observed in hypothesis testing.

For Teaching Effectiveness, Faculty Trust emerges as the most influential predictor with a very large effect size (f^2 = 0.62), highlighting its central role in enhancing instructional outcomes. AI Accuracy has a small but meaningful effect (f^2 = 0.12), while Hallucination Risk again demonstrates only a minimal influence (f^2 = 0.04).

Overall, the results confirm that AI Accuracy and Faculty Trust are the dominant drivers of Teaching Effectiveness, whereas Hallucination Risk contributes minimally, both directly and indirectly, within the proposed model.

DISCUSSION

The study investigated the connections between AI accuracy, AI hallucination risk, faculty trust in AI, and teaching effectiveness using PLS-SEM. Findings offer empirical insights into faculty assessments of AI-enabled teaching environments and how perceptions of reliability impact instructional outcomes.

Structural Relationships and Key Findings

The findings reveal that AI Accuracy positively influences Teaching Effectiveness (β = 0.329, p < 0.01), indicating that faculty who recognize AI-generated content as accurate are more likely to use it effectively in their teaching. This enhances clarity, correctness, and instructional quality, supporting previous studies on the importance of system performance and output quality in AI-assisted tasks [15,28].

AI hallucination risk does not significantly affect teaching effectiveness. While hallucinations are seen as a limitation of generative AI instructors may not view this risk as a barrier, instead using their expertise to address inaccuracies and lessen its impact [5].

A key finding of the study indicates that AI accuracy significantly impacts faculty trust in AI ($\beta = 0.568$, $p < 0.001$), aligning with theories of trust in automation that emphasize the importance of perceived reliability and accuracy [9]. Faculty trust in AI tools grows when their outputs consistently meet academic standards with minimal need for verification.

AI hallucination risk does not significantly impact faculty trust, which contrasts with existing literature's emphasis on AI risks. This may be because faculty awareness of hallucinations does not lead to diminished trust, particularly when users are cautious and evaluative of AI outputs.

Faculty trust in AI is identified as a crucial predictor of teaching effectiveness ($\beta = 0.579$, $p < 0.001$), indicating a significant effect size. This underscores trust as a key psychological factor for successful human AI collaboration. When faculty members trust AI systems, they tend to utilize them confidently for content creation, instructional assistance, and enhancing student engagement, ultimately leading to improved teaching results [7].

Mediation Effects

The mediation analysis indicates that faculty trust partially mediates the relationship between AI accuracy and teaching effectiveness, suggesting that AI accuracy improves teaching effectiveness both directly and indirectly through trust. This supports human AI interaction models, where system characteristics affect outcomes via trust formation [9].

Despite its theoretical importance, AI hallucination risk shows no significant mediation effect on faculty behavior in teaching contexts, indicating its limited influence in practical applications.

Model Explanatory Power and Predictive Relevance

The model exhibits moderate to strong explanatory power, evidenced by R^2 values of 0.371 for Faculty Trust and 0.579 for teaching effectiveness. These figures reflect a significant variance explanation of crucial constructs, especially teaching effectiveness. Furthermore, positive Q^2 values reinforce the model's predictive relevance, highlighting its robustness in elucidating faculty perceptions.

Effect size analysis indicates that AI accuracy significantly influences faculty trust, which in turn strongly affects teaching effectiveness. Hallucination risk, however, displays minimal effects, highlighting its limited importance in the model.

Theoretical Implications

The findings of the study contribute to the literature by shifting the focus from AI adoption to AI reliability, emphasizing accuracy as vital for teaching outcomes, thus addressing gaps in prior research highlighted by Zawacki-Richter [1]. It integrates trust in AI as a mediating construct, supporting its role between system characteristics and pedagogical outcomes, which extends trust in automation theory into education. Additionally, it challenges the prevailing narrative of AI risks, showing that concerns about hallucination may not yet impact teaching effectiveness when faculty are well-informed evaluators.

Practical Implications

Higher education institutions should enhance the accuracy and reliability of AI tools in teaching, focusing training programs on helping faculty critically evaluate AI outputs. It is important to foster calibrated trust in AI, encouraging its supportive use while preserving academic oversight to improve teaching effectiveness without compromising knowledge quality.

CONCLUSION AND IMPLICATIONS

This study investigates how faculty perceptions of AI accuracy and hallucination risk influence teaching effectiveness, with faculty trust in AI as a mediating variable, employing a PLS-SEM approach. Results show that perceived AI accuracy significantly affects both faculty trust and teaching effectiveness, indicating that reliability in AI outputs encourages faculty integration of AI into teaching. AI hallucination risk does not directly impact teaching effectiveness, as faculty can usually mitigate such risks through expertise. Notably, trust in AI emerges as the principal predictor of teaching effectiveness, reflecting the centrality of user confidence in human AI collaboration [9]. The research shifts the focus from mere adoption of AI in education to the significance of AI's epistemic reliability and accuracy in fostering quality educational outcomes [1]. Overall, the study proposes an integrated model linking perceptions of AI reliability, trust, and teaching effectiveness, emphasizing that successful AI integration in higher education relies on generating trustworthy knowledge rather than just widespread adoption.

Policy Implications

The findings highlight significant implications for policymakers, educational institutions, and regulatory bodies like UGC and AICTE. Key recommendations include prioritizing reliable AI systems in education through quality assurance, developing faculty training focused on AI evaluation, and promoting ethical AI usage aligned with UNESCO guidelines, which call for transparency and accountability in educational AI systems [29].

Academic Implications

This study shifts the focus from technology acceptance to epistemic evaluation of AI systems, integrating AI reliability constructs such as accuracy and hallucination risk with pedagogical outcomes. It empirically validates the mediating role of trust, highlighting its significance as a theoretical bridge between system characteristics and user outcomes, thus opening new research avenues in education, information systems, and artificial intelligence.

Future Research Directions

This study outlines essential directions for future research on AI in higher education. It advocates for larger, diverse samples across various disciplines and geographic regions to better

understand how specific factors influence perceptions of AI accuracy and hallucination risk. Longitudinal studies are suggested to monitor changes in faculty trust and teaching effectiveness with AI tools. Incorporating constructs like explainability, transparency, ethical issues, and AI literacy, as suggested by Doshi-Velez & Kim, may enhance trust and adoption. Investigation of moderating factors such as teaching experience and institutional support is also recommended [17]. Furthermore, mixed-methods and experimental approaches should supplement surveys, allowing detailed insights into decision-making and direct impacts on student learning outcomes. Including student perspectives alongside faculty input can provide a comprehensive view of AI's educational role, guiding responsible integration. Overall, interdisciplinary and multi-stakeholder methods are encouraged to form evidence-based guidelines for ethical AI use in higher education.

REFERENCES

1. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education–where are the educators?. International journal of educational technology in higher education. 2019;16(1):39.
2. Kasneci E, Seßler K, Küchemann S, Bannert M, Dementieva D, et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learning and individual differences. 2023;103:102274.
3. Shmitchell S, Bender EM, McMillan-Major A, Gebru T. La era de la inteligencia artificial generativa y la sombra de 1984:? puede la tecnología más disruptiva del siglo XXI abrir paso a la concreción de la distopía orwelliana?. Derecho, Innovación & Desarrollo Sustentable. 2024.
4. Russell SJ, Norvig P, Canny JF, Malik JM, Edwards DD. Artificial intelligence: a modern approach. Englewood Cliffs, NJ: Prentice hall; 1995.
5. Ji Z, Lee N, Frieske R, Yu T, Su D, et al. Survey of hallucination in natural language generation. ACM computing surveys. 2023;55(12):1-38.
6. Maynez J, Narayan S, Bohnet B, McDonald R. On faithfulness and factuality in abstractive summarization. In Proceedings of the 58th annual meeting of the association for computational linguistics 2020 (pp. 1906-1919).
7. Scherer R, Siddiq F, Tondeur J. The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. Computers & education. 2019;128:13-35.
8. Zhai X, Chu X, Chai CS, Jong MS, Istenic A, et al. A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. Complexity. 2021;2021(1):8812542.
9. Lee JD, See KA. Trust in automation: Designing for appropriate reliance. Human factors. 2004;46(1):50-80.
10. Parasuraman R, Riley V. Humans and automation: Use, misuse, disuse, abuse. Human factors. 1997;39(2):230-53.
11. Carter M. Visible learning: A synthesis of over 800 meta-analyses relating to achievement.
12. Holmes W, Bialik M, Fadel C. Artificial intelligence in education promises and implications for teaching and learning. Center for Curriculum Redesign; 2019.
13. OECD (2021), AI and the Future of Skills, Volume 1. Capabilities and Assessments, Educational Research and Innovation, OECD Publishing, Paris.
14. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly. 1989:319-40.
15. Ribeiro MT, Singh S, Guestrin C. " Why should i trust you?" Explaining the predictions of any classifier. InProceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining 2016 (pp. 1135-1144).
16. Tabassi E. Artificial intelligence risk management framework (AI RMF 1.0).
17. Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608. 2017.
18. Marsh HW. SEEQ: A RELIABLE, VALID, AND USEFUL INSTRUMENT FOR COLLECTING STUDENTS'EVALUATIONS OF UNIVERSITY TEACHING. British journal of educational psychology. 1982;52(1):77-95.
19. Coe R, Aloisi C, Higgins S, Major LE. What makes great teaching. Review of the underpinning research. 2014;44:0-57.
20. Teräs M. Education and technology: Key issues and debates: Book Review.
21. Marsh HW, Roche LA. Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. American psychologist. 1997;52(11):1187.
22. DeLone WH, McLean ER. The DeLone and McLean model of information systems success: a ten-year update. Journal of management information systems. 2003;19(4):9-30.
23. Featherman MS, Pavlou PA. Predicting e-services adoption: a perceived risk facets perspective. International journal of human-computer studies. 2003;59(4):451-74.
24. Rana MM, Siddiqee MS, Sakib MN, Ahamed MR. Assessing AI adoption in developing country academia: A trust and privacy-augmented UTAUT framework. Heliyon. 2024;10(18).
25. Shata A, Hartley K. Artificial intelligence and communication technologies in academia: faculty perceptions and the adoption of generative AI. International Journal of Educational Technology in Higher Education. 2025;22(1):14.
26. Nazaretsky T, Ariely M, Cukurova M, Alexandron G. Teachers' trust in AI-powered educational technology and a professional development program to improve it. British journal of educational technology. 2022;53(4):914-31.
27. Hair Jr JF, Sarstedt M, Hopkins L, Kuppelwieser VG. Partial least squares structural equation modeling (PLS-SEM). European business review. 2014;26(2):106.
28. Bubeck S, Chadrasekaran V, Eldan R, Gehrke J, Horvitz E, et al. Sparks of artificial general intelligence: Early experiments with gpt-4 [Internet]. 2023.
29. UNESCO. (2021). Recommendation on the ethics of artificial intelligence.United Nations Educational, Scientific and Cultural Organization.