Contemporary Educational Technology, 2026, **18**(4), ep707
ISSN: 1309-517X (Online)

OPEN ACCESS

Review Article

From digital tools to digital authority: A systematic review of trust in AI-driven educational technologies

Elena N. Tumilevich [1]
0000-0002-4360-570X

Mikhail S. Miroshkin [2]
0000-0002-7964-935X

Elena N. Kolomoets [3*]
0000-0002-4632-6757

Irina S. Andryushchenko [4]
0000-0003-4222-4865

Valeria L. Zakharova [5]
0000-0001-5407-6906

Lyudmila M. Smirnova [6]
0000-0002-6581-4529

[1] Institute of Economics and Management, Pacific National University, Khabarovsk, RUSSIA
[2] Department of Humanities, Financial University under the Government of the Russian Federation, Moscow, RUSSIA
[3] Department of Sociology, Psychology and Social Management, Moscow Aviation Institute (National Research University), Moscow, RUSSIA
[4] Department of Foreign Languages, Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University), Moscow, RUSSIA
[5] Department of Medical and Social Assessment, Emergency, and Ambulatory Practice, Sechenov First Moscow State Medical University, Moscow, RUSSIA
[6] Department of Dermatovenereology, Sechenov First Moscow State Medical University, Moscow, RUSSIA
[*] Corresponding author: kolomoetc_elena@mail.ru

Citation: Tumilevich, E. N., Miroshkin, M. S., Kolomoets, E. N., Andryushchenko, I. S., Zakharova, V. L., & Smirnova, L. M. (2026). From digital tools to digital authority: A systematic review of trust in AI-driven educational technologies. *Contemporary Educational Technology, 18*(4), Article ep707. https://doi.org/10.30935/cedtech/19974

ARTICLE INFO

Received: 4 Jun 2026
Accepted: 7 Nov 2026

ABSTRACT

This study aims to examine the factors shaping university students' trust in artificial intelligence (AI)-driven educational technologies through content analysis, including evidence on instructors and other stakeholders. It also investigates the role that AI's perception as a digital authority plays in these trust mechanisms. A comprehensive search of the Web of Science and Scopus databases (2018-2026) identified 33 studies, analyzed according to the preferred reporting items for systematic reviews and meta-analyses protocol. The content analysis, organized around five research questions, produced findings under five main themes: the conceptualization of trust, communicative factors, perceptual factors, the digital authority relationship, and methodological tendencies. The findings reveal that most reviewed studies frame trust within the technology acceptance model, with perceived usefulness and ease of use standing out as the primary antecedents. Among communicative factors, algorithmic transparency and explainable AI applications play a determining role, while ethical framing and pedagogical communication are identified as independent determinants. Risk perception is established as the primary barrier variable that negatively affects trust. AI systems are found to be positioned by students as epistemic reference points, a condition that may lay the groundwork for trust miscalibration. The studies are concentrated in 2025-2026, and structural equation modeling and cross-sectional survey designs predominate, while longitudinal and cross-cultural comparative research remains insufficient. To enable university students to develop a sound calibration of

Copyright © 2026 by authors; licensee CEDTECH by Bastas. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/).

trust toward AI-driven educational technologies, higher education institutions are recommended to establish comprehensive policies that jointly address transparency-centered system design, instructor-led pedagogical communication, and curriculum-integrated digital literacy training.

Keywords: AI-driven educational technologies, trust in technology, digital authority, perceived value, communication transparency

INTRODUCTION

Spanning from adaptive learning platforms and intelligent chatbots to automated assessment tools and personalized content recommendation systems, artificial intelligence (AI) applications carry the potential to fundamentally transform the educational experience (Khodjaev et al., 2026; Khosravi et al., 2022). The realization of this technological potential depends largely on user trust. In the absence of trust, even functionally superior systems may fail to gain acceptance in educational settings (Cui et al., 2025; Nazaretsky et al., 2025). The rapid integration of AI-driven educational technologies into higher education institutions has brought to the fore the need to understand the trust levels students hold toward these systems.

Trust, especially as applied to university students, is not limited to an individual characteristic but is a multifaceted socio-technical construct with communicative, perceptual, and structural aspects. Here, communicative aspects, such as transparency and explainability (Baldwin & Nasirci, 2024; Schmidt et al., 2020; Türkmen, 2025; von Eschenbach, 2021; Zheltukhina et al., 2026), as well as cognitive-perceptual aspects, like usefulness, ease of use, and risk perception (Alarefi & Alharbi, 2025; Mustofa et al., 2025; Zhang et al., 2025), can be identified as key predictors of trust. A review of the existing literature reveals that most studies either measure perceptual variables within the technology acceptance model (TAM) framework (Boughanzai et al., 2026; Mustofa et al., 2025; Wang & Yu, 2026) or treat communicative dimensions such as transparency and explainability as independent constructs (Feldman-Maggor et al., 2025; Khosravi et al., 2022). Integrative models that position trust as a mediating construct while capturing the interplay between communicative and perceptual factors remain quite limited, both theoretically and empirically. The question of how students perceive AI systems as digital authority and how this shapes their trust relationship constitutes an original dimension that the field has not yet examined in sufficient depth (Kraipiyaset et al., 2026; Pandey et al., 2026; Safarov et al., 2026; Shin et al., 2020). Kazimova et al. (2025) similarly draw attention to the lack of comprehensive assessments of how trust in AI applications takes shape, despite their growing penetration into educational contexts.

Although research on trust in AI-driven educational technologies has grown considerably in recent years, the literature has largely examined perceptual variables (perceived usefulness, ease of use, and risk perception) or communicative variables (transparency, explainability, and feedback quality) in isolation; the number of studies that systematically assess the interplay between these two dimensions remains quite limited (Aquilino et al., 2025; Cui et al., 2025). Furthermore, the concept of digital authority has been insufficiently theorized. Although evidence that users perceive AI systems as epistemic authorities is present in the literature (Safarov et al., 2026; Shin et al., 2020), the structural effect of this phenomenon on trust mechanisms has not yet been addressed within a direct and comprehensive theoretical framework. This study seeks to address this gap by examining how the factors shaping university students' trust in AI-driven educational technologies are addressed in the existing literature and what role the perception of AI as a digital authority plays in this process.

THEORETICAL FRAMEWORK

TAMs and the Positioning of Trust

The contribution of TAMs to AI trust research in educational contexts is both wide-ranging and foundational. Davis' (1989) original TAM established perceived usefulness and perceived ease of use as the primary determinants of adoption behavior. In educational settings, trust is positioned as a critical construct that mediates or reinforces these two variables (Alarefi & Alharbi, 2025; Marimon et al., 2025; Mustofa et al., 2025; Shashayeva et al., 2026). The literature extends TAM in several directions. Mustofa et al. (2025)

incorporated ethical concerns and subjective norms into the framework, demonstrating that trust serves a moderating function among these variables. Boughanzai et al. (2026) introduced affective dimensions such as anxiety and social influence, arguing that technology adoption goes beyond purely cognitive processes. Drawing on the unified theory of acceptance and use of technology (UTAUT) perspective, Fakhri et al. (2026) examined the effect of institutional support on trust and found that trust operates as the primary mediating variable explaining continuance intention. Taken together, these contributions position trust not as a passive element within the TAM ecosystem but as a foundational construct at its core.

Broader syntheses support this reading as well. Working with 27 studies and a pooled sample of more than 12,000 participants, Xue et al. (2026) showed that TAM relationships tend to be stronger when AI is the focal technology than in other educational technology domains, with the perceived usefulness to behavioral intention path proving particularly robust. One plausible reason is that AI systems are less transparent than conventional tools, which makes trust-related considerations more salient during adoption decisions. Khoso et al. (2026) approached the same question from a language learning angle. Across two studies totaling over 1,100 EFL students, they found that digital literacy shaped perceived ease of use and perceived usefulness, and that these perceptions fed into trust, which then predicted both intention to use AI and creative output. Notably, the study treated trust as two-dimensional rather than unitary, separating competence-based appraisals from those tied to benevolence and integrity, and the two dimensions behaved differently across outcomes. Taken together, the evidence from both macro-level synthesis and learner-level empirical work points to trust as a construct that warrants more careful theoretical treatment than it typically receives within standard TAM applications.

Communicative and Perceptual Factors

The communicative basis of trust in AI systems is largely shaped by questions of transparency and explainability. von Eschenbach (2021) identified the algorithmic black-box problem as the primary structural source of distrust, while Schmidt et al. (2020) argued that transparency functions as a core design feature that strengthens users' confidence in AI systems. Explainable AI (XAI) approaches are relevant in this context. Khosravi et al. (2022) demonstrate that XAI concretizes trust by implementing the fairness, accountability, transparency, and ethics principles. Türkmen's (2025) systematic review synthesizing XAI studies in educational research shows that explanation quality must align with user expertise levels; explanations that are either overly technical or overly simplified can both erode trust. Safarov et al. (2026) note that the interaction between explanation length and student expertise is a critical determinant of trust calibration. Garcia (2026) has shown that ethical framing and user autonomy function as independent determinants among communicative factors.

Pedagogical communication also carries significance in this framework. Colclasure et al. (2025) and Akman Yeşilel (2026) have indicated that the instructor's attitudes and discourse toward AI tools directly shape students' perceptions of trust, suggesting that trust depends not only on system design but also on how the system is introduced and explained. Perceptual factors constitute the cognitive dimension of trust formation. Among four main categories of antecedents, Cui et al. (2025) found that technological performance perception and user experience exert the strongest influence on trust. Perceived usefulness stands out as the most prominent variable within these categories. Küçükaydın (2025) demonstrated through structural equation modeling (SEM) that perceived usefulness fully mediates the relationship between academic stress and AI trust. Risk perception, on the other hand, impedes trust formation. Pre-existing concerns about privacy and security risks directly and negatively affect perceived trust (Zhang et al., 2025). Technology anxiety both amplifies risk perception and weakens the effect of trust on adoption behavior (Karayianni et al., 2026). Rana et al. (2024) highlight that the negative effect of privacy concern on trust is particularly pronounced in developing country contexts.

Accuracy and reliability perception constitutes another critical component of trust formation. Nazaretsky et al. (2025) note that system reliability is the strongest determinant of trust, particularly in faculty samples. Feldman-Maggor et al. (2025) demonstrate that trust is dynamic; system performance can increase or decrease trust through successive interactions. Lijie et al. (2025) suggest that positive perceptions of AI tools may reinforce critical thinking dispositions, pointing to a bidirectional relationship between trust and cognitive development.

AI as Digital Authority

Trust functions not only as a facilitator of adoption but also as a social mechanism enabling innovative systems to reach positions of epistemic authority. Shin et al. (2020) argue that algorithms function as a lens through which people perceive the external world, and that this is central to understanding the authority attributed to AI systems. For this authority to serve a constructive function, sound trust calibration is necessary. However, Safarov et al. (2026) found that students with lower levels of expertise accept AI outputs without question, which creates conditions for trust miscalibration.

Trust miscalibration is emerging in the literature as a new yet critical concept. James et al. (2025) identified the factor structure of AI trust among university students and found digital literacy to be a significant moderating variable of these trust dynamics. Students with higher digital literacy tend to evaluate AI outputs more critically and thus appear less susceptible to trust miscalibration (Karran et al., 2025).

Purpose of the Study

This study aims to examine the factors shaping university students' trust in AI-driven educational technologies through a systematic content analysis, and to reveal the role of AI's perception as a digital authority in these trust mechanisms. Although university students constitute the primary focus, studies involving instructors, pre-service teachers, and other higher education stakeholders were also included where they informed these mechanisms. In this context, the following sub-questions were addressed:

1. How is university students' trust in AI-driven educational technologies conceptualized in the literature, and within which theoretical frameworks is it examined?
2. How is the relationship between communicative factors (transparency, explainability, information quality, and feedback) and students' trust in AI explained?
3. How are the effects of perceptual factors (perceived usefulness, ease of use, accuracy, reliability, and risk perception) on students' trust levels reported?
4. How is AI's perception as a digital authority associated with students' trust, technology adoption, and academic decision-making processes?
5. What methodological approaches are used in studies in this area, and what research gaps stand out in the literature?

METHOD

Research Design

This study examines the international literature on the communicative and perceptual factors shaping university students' trust in AI-driven educational technologies. A systematic review method was adopted within a qualitative research framework, with content analysis serving as the analytical technique. Systematic review involves synthesizing findings from a comprehensive literature search according to pre-specified, explicitly stated criteria (Petticrew & Roberts, 2006). Content analysis allows researchers to classify and interpret both the manifest and latent content of texts (Krippendorff, 2019).

Databases and Search Strategy

The literature search was carried out through the Web of Science (WoS) and Scopus databases. Both are widely recognized as among the most comprehensive citation-indexed repositories covering peer-reviewed publications in the social sciences and education (Falagas et al., 2008). All searches were restricted to English-language studies published between 2018 and 2026. The final searches were run on 06 June 2026 in WoS and Scopus.

The following search terms were used in the WoS database:

TS=("artificial intelligence" OR "AI" OR "machine learning") AND TS=("educational technology" OR "e-learning" OR "digital learning" OR "online learning") AND TS=("trust" OR "trustworthiness") AND TS=("university students" OR "higher education" OR "undergraduate").

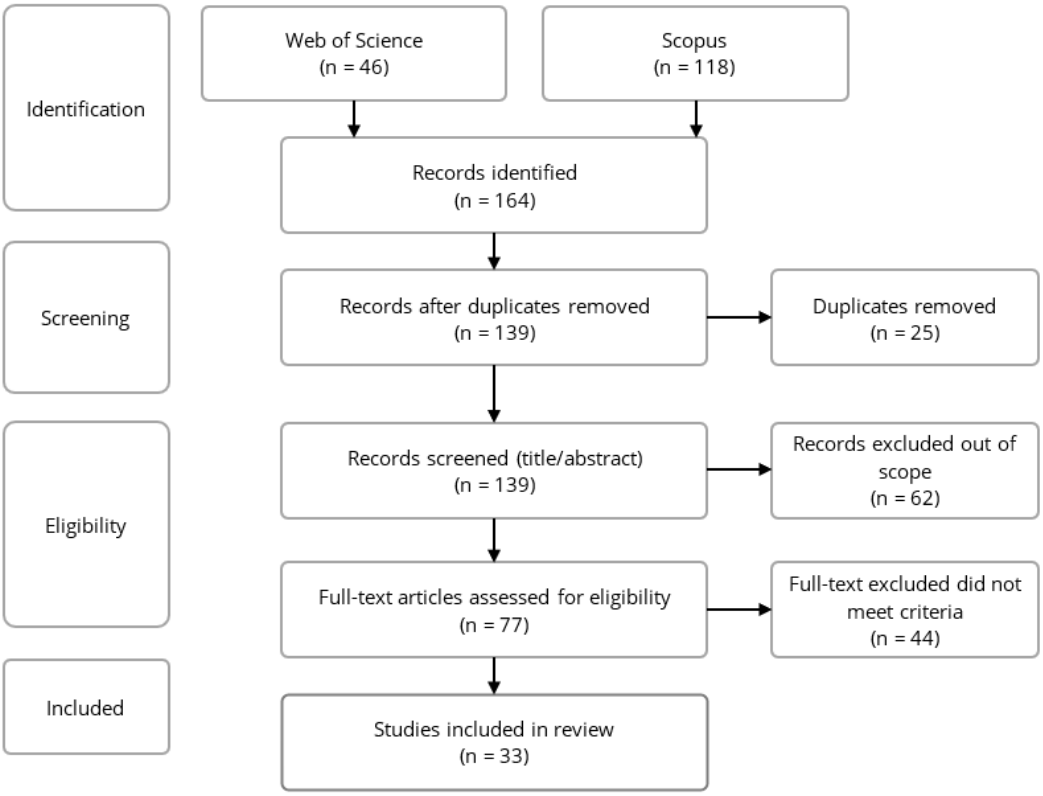


Figure 1. PRISMA flow diagram (the authors' own elaboration)

The results were limited to English-language documents published between 2018 and 2026.

The following search string was used in the Scopus database:

TITLE-ABS-KEY ((("trust in AI" OR "AI trust" OR "trust in artificial intelligence" OR "algorithmic trust") AND ("education" OR "learning" OR "university" OR "higher education") AND ("transparency" OR "explainability" OR "algorithmic transparency" OR "perceived usefulness" OR "pedagogical"))) AND (LIMIT-TO (SUBJAREA, "SOCI") OR LIMIT-TO (SUBJAREA, "PSYC") OR LIMIT-TO (SUBJAREA, "MULT") OR LIMIT-TO (SUBJAREA, "ARTS")) AND (LIMIT-TO (LANGUAGE, "English")) AND PUBYEAR > 2017 AND PUBYEAR < 2027.

PRISMA Protocol

The study selection process was conducted in four stages in accordance with the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines (Page et al., 2021), as summarized in **Figure 1**. The completed PRISMA 2020 checklist is provided as **Appendix A**.

In the Identification stage, a total of 164 records were retrieved: 46 from WoS and 118 from Scopus. In the screening stage, duplicate records were removed, yielding 139 unique records. After titles and abstracts were reviewed, 62 records deemed out of scope were excluded; 77 records were retained for full-text review.

In the eligibility stage, the full texts of all 77 studies were retrieved and evaluated against the inclusion and exclusion criteria. As a result, 44 studies were excluded for the reasons described below. In the inclusion stage, 33 studies meeting all inclusion criteria were incorporated into the systematic review.

Inclusion and Exclusion Criteria

Studies were selected based on the following inclusion criteria:

(1) published between 2018 and 2026,

(2) peer-reviewed articles or systematic reviews published in English,

(3) addressing trust toward AI-supported educational technologies as a focal variable,

(4) including university students as the primary sample population or, where the evidence informed trust in higher education contexts, instructors, pre-service teachers, or other educational stakeholders, and

(5) containing theoretical or empirical content pertaining to communicative or perceptual factors or to the concepts of digital authority and trust.

Exclusion criteria were established as follows:

(1) published prior to 2018,

(2) published in a language other than English,

(3) constituting technical system design studies without direct relevance to AI trust,

(4) involving K-12 or clinical settings outside the university context, and

(5) consisting solely of conference abstracts.

Data Collection and Analysis

Relevant data were extracted from the 33 studies meeting the eligibility criteria. For each study, the following information was recorded: author(s) and year of publication, journal name, study design, sample characteristics, dependent and independent variables measured, theoretical framework(s), key findings, and research gaps. The content analysis incorporates a combination of inductive and deductive approaches. Major themes were first established based on the research questions; the theoretical content and findings of each study were then coded. Code categories were grouped on the basis of the overlapping conceptual content of the studies, and sub-themes were defined (Elo & Kyngäs, 2008). Each study was assigned to at least one category on the basis of the research questions; where thematic overlap occurred, a study was classified under more than one category.

Validity and Reliability

To enhance the reliability of the content analysis, a two-level reliability assessment was conducted following the inter-coder agreement percentage calculation approach recommended by Miles and Huberman (1994) as a criterion for qualitative research reliability. At the first level, the internal consistency of the researcher's own codes was tested. The researcher repeated the coding process after a two-week interval by recoding the same text units; the agreement rate between the two coding rounds was calculated as 85.0%. This figure exceeds the 70% threshold recommended by Miles and Huberman (1994) for acceptable reliability. After this an independent coder coded 10 randomly selected studies using the sub-theme and category template prepared by the researcher. The coder was provided with the template definitions in written form, and the agreement rate between the two coders' assignments was subsequently calculated. Agreement was found to be 92.5% at the theme level, 85.5% at the sub-theme level, and 81.3% at the category level. In cases of disagreement, the coders discussed the category definitions and reached consensus.

To strengthen internal validity, direct references were made to the studies supporting the findings; for external validity, the scope criteria were explicitly defined within standard search protocols (Lincoln & Guba, 1985). The PRISMA flowchart was adopted as the primary indicator of transparency, and all exclusion rationales were recorded at the relevant stage.

Descriptive Characteristics of the Included Studies

The descriptive characteristics of the 33 studies included in the review are presented in **Table 1**.

Table 1. Descriptive characteristics of the included studies

Study	Journal/publication	Year	C/C	Method
Akman Yeşilel (2026)	British Educational Research Journal	2026	Turkey	Qualitative
Aladi (2024)	ACM SIGMIS CPR Conference Proceedings	2024	USA	Qualitative
Alarefi and Alharbi (2025)	Journal of Information Science Theory and Practice	2025	Saudi Arabia	Survey (SEM)
Boughanzai et al. (2026)	Education and Information Technologies	2026	Morocco	Survey (SEM)
Colclasure et al. (2025)	Trends in Higher Education	2025	USA	Mixed method
Cui et al. (2025)	IEEE Access	2025	China	Meta-analysis
Fakhri et al. (2026)	Daengku	2026	Indonesia	Survey (SEM)
Feldman-Maggor et al. (2025)	International Journal of AI in Education	2025	Israel	Experimental

Table 1 (Continued). Descriptive characteristics of the included studies

Study	Journal/publication	Year	C/C	Method
Garcia (2026)	Evaluation Review	2026	Philippines	Experimental
Jo (2024)	International Journal of Educational Technology in Higher Education	2024	South Korea	Survey
Karayianni et al. (2026)	TechTrends	2026	Greece	Survey (SEM)
Karran et al. (2025)	npj Science of Learning	2025	Canada	Mixed method
Khlaif et al. (2025)	BMC Medical Education	2025	Palestine	Mixed method
Khosravi et al. (2022)	Computers and Education: AI	2022	Australia	Systematic review
Küçükaydın (2025)	Psychology in the Schools	2025	Turkey	Survey
Marimon et al. (2025)	RIED	2025	Spain	Survey (SEM)
Mustofa et al. (2025)	Computers and Education: AI	2025	Indonesia	Survey (SEM)
Nazaretsky et al. (2025)	Computers and Education: AI	2025	Switzerland	Survey (SEM)
Omirali et al. (2025)	Sustainability	2025	Kazakhstan	Survey
Pandey et al. (2026)	AI & SOCIETY	2026	India	Scale development
Raza et al. (2024)	Journal of Management Practices, Humanities and Social Sciences	2024	Pakistan	Review
Safarov et al. (2026)	Human Factors	2026	Monaco	Experimental
Schilhabel et al. (2026)	Journal of Information, Communication and Ethics in Society	2026	USA	Survey (SEM)
Schmidt et al. (2020)	Journal of Decision Systems	2020	Germany	Conceptual
Shahzad et al. (2025)	Education and Information Technologies	2025	China	Survey
Shin et al. (2020)	International Journal of Information Management	2020	UAE	Mixed method
Shishakly and Nachouki (2026)	Journal of Information Technology Education	2026	UAE	Survey
Teodorescu et al. (2023)	Sustainability	2023	Romania/USA	Survey
Türkmen (2025)	Journal of Educational Computing Research	2025	Turkey	Systematic review
von Eschenbach (2021)	Philosophy & Technology	2021	USA	Conceptual
Wang and Yu (2026)	International Journal of Human-Computer Interaction	2026	China	Survey (SEM)
Wang et al. (2026)	Sustainability	2026	China	Experimental
Zhang et al. (2025)	Education and Information Technologies	2025	China	Survey (SEM)

Note. C/C: Country/context

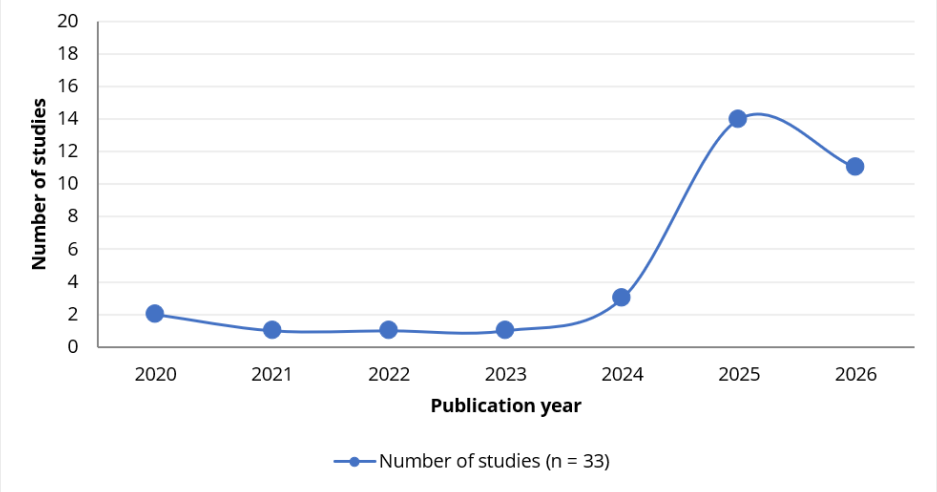


Figure 2. Distribution graph by year (the authors' own elaboration)

The studies are concentrated in the period from 2020 to 2026. Studies published in 2025-2026 (n = 25; 14 in 2025 and 11 in 2026) constitute the majority of the total sample (see **Figure 2**), indicating that AI trust research has accelerated over the last two years. The studies originate from researchers based in Asia, Europe, the Middle East, North America, and Africa, reflecting a broad geographic distribution. Among the journals in which they appeared, *Computers and Education: Artificial Intelligence*, *Education and Information Technologies*, and *Sustainability* are the most frequently represented outlets. Methodologically, quantitative studies (survey- and SEM-based) predominate, while qualitative and mixed-method research remains comparatively limited.

Table 2. Frameworks used in the conceptualization of trust

Theme	Subtheme	Category	n	Studies
Conceptualization of trust	TAMs	TAM-based models	9	Alarefi and Alharbi (2025), Boughanzai et al. (2026), Fakhri et al. (2026), Jo (2024), Marimon et al. (2025), Mustofa et al. (2025), Shahzad et al. (2025), Teodorescu et al. (2023), Wang and Yu (2026)
		UTAUT and extended integration models	6	Alarefi and Alharbi (2025), Fakhri et al. (2026), Karayianni et al. (2026), Mustofa et al. (2025), Schilhabel et al. (2026), Zhang et al. (2025)
	Multidimensional trust constructs	Cognitive-affective trust components	7	Colclasure et al. (2025), Cui et al. (2025), Khlaif et al. (2025), Nazaretsky et al. (2025), Omirali et al. (2025), Pandey et al. (2026), Raza et al. (2024)
		Epistemic and normative trust frameworks	7	Akman Yeşilel (2026), Karran et al. (2025), Küçükaydın (2025), Pandey et al. (2026), Shin et al. (2020), Shishakly and Nachouki (2026), von Eschenbach (2021)
		Philosophical approaches	6	Feldman-Maggor et al. (2025), Garcia (2026), Safarov et al. (2026), Schmidt et al. (2020), von Eschenbach (2021), Wang et al. (2026)
	Meta-analytic and scale development studies	AI trust scale development	3	Aladi (2024), Nazaretsky et al. (2025), Pandey et al. (2026)
		Meta-analysis of trust antecedents	3	Cui et al. (2025), Khosravi et al. (2022), Türkmen (2025)

FINDINGS

This section presents the findings of the content analysis conducted around five research questions. For each research question, a thematic table is provided first, followed by summary explanations supporting the table.

Conceptualization of Trust and Theoretical Frameworks

As observed from **Table 2**, the majority of reviewed studies use either TAM or an extended version of TAM as a theoretical background for examining AI trust. A total of nine TAM-based models includes additional constructs, such as ethics, subjective norms, anxiety, and trust. According to Mustofa et al. (2025), Alarefi and Alharbi (2025), and Boughanzai et al. (2026), TAM models include such variables as ethics, subjective norms, and anxiety. The UTAUT framework served as a foundation for comparisons across student groups within six studies. In UTAUT models, the theory is used for conducting multiple-group analysis, as evidenced by the works of Fakhri et al. (2026) and Karayianni et al. (2026). Cognitive-affective trust models assume the presence of different antecedents and consequences due to viewing AI trust as a multidimensional concept. Regarding multidimensional trust, Nazaretsky et al. (2025) and Cui et al. (2025) examine both cognitive and affective components of AI trust. Epistemic trust is considered an emerging theory, especially when dealing with generative AI. According to Pandey et al. (2026), who conducted a scale development study, epistemic trust is a functional theoretical concept in higher education. From meta-analyses, one may conclude that trust antecedents show a clear hierarchy in terms of their effect sizes.

Communicative Factors and AI Trust

Transparency and explainability stand out as the most frequently addressed dimensions among communicative factors (**Table 3**). Algorithmic transparency has been identified as a fundamental condition of trust in nine studies, with the black box problem characterized as a structural source of distrust. Schmidt et al. (2020) and von Eschenbach (2021) define algorithmic transparency as a prerequisite for AI trust and regard the black box problem as a structural source of distrust. XAI applications reveal that the manner in which a system communicates its decisions to users has a determining influence on trust. XAI studies have been examined by Khosravi et al. (2022), Feldman-Maggor et al. (2025), and Türkmen (2025) as an instrument for strengthening trust in AI systems. Studies investigating ethical framing and user autonomy emphasize that the design features of AI systems and ethical awareness shape trust. Garcia (2026) and Shishakly and

Table 3. Communicative factors and AI trust

Theme	Subtheme	Category	n	Studies
Communicative factors and AI trust	Transparency and explainability	Algorithmic transparency and the black-box problem	9	Akman Yeşilel (2026), Aladi (2024), Boughanzai et al. (2026), Karran et al. (2025), Raza et al. (2024), Schmidt et al. (2020), Shin et al. (2020), Türkmen (2025), von Eschenbach (2021)
		XAI applications	6	Feldman-Maggor et al. (2025), Khosravi et al. (2022), Nazaretsky et al. (2025), Safarov et al. (2026), Türkmen (2025), Wang et al. (2026)
		Ethical framing and user autonomy	6	Garcia (2026), Karran et al. (2025), Pandey et al. (2026), Schilhabel et al. (2026), Shishakly and Nachouki (2026), von Eschenbach (2021)
	Information quality and feedback mechanisms	Feedback quality and system communication	7	Colclasure et al. (2025), Jo (2024), Karayianni et al. (2026), Khlaif et al. (2025), Marimon et al. (2025), Omirali et al. (2025), Wang and Yu (2026)
		Instructor credibility and pedagogical communication	7	Aladi (2024), Colclasure et al. (2025), Fakhri et al. (2026), Mustofa et al. (2025), Raza et al. (2024), Shahzad et al. (2025), Teodorescu et al. (2023)
		Academic integrity and communicative norms	5	Akman Yeşilel (2026), Alarefi and Alharbi (2025), Küçükaydın (2025), Shishakly and Nachouki (2026), Zhang et al. (2025)

Table 4. Perceptual factors and trust levels

Theme	Subtheme	Category	n	Studies
Perceptual factors and trust levels	Perceived usefulness and ease of use	Perceived usefulness	11	Alarefi and Alharbi (2025), Boughanzai et al. (2026), Jo (2024), Khlaif et al. (2025), Küçükaydın (2025), Marimon et al. (2025), Mustofa et al. (2025), Omirali et al. (2025), Shahzad et al. (2025), Teodorescu et al. (2023), Wang and Yu (2026)
		Perceived ease of use	7	Alarefi and Alharbi (2025), Boughanzai et al. (2026), Fakhri et al. (2026), Karayianni et al. (2026), Marimon et al. (2025), Mustofa et al. (2025), Wang and Yu (2026)
		Subjective norms and social influence	5	Akman Yeşilel (2026), Karran et al. (2025), Mustofa et al. (2025), Schilhabel et al. (2026), Shishakly and Nachouki (2026)
	Risk perception and anxiety	Perceived risk and concern	7	Jo (2024), Karayianni et al. (2026), Schilhabel et al. (2026), Schmidt et al. (2020), Shishakly and Nachouki (2026), von Eschenbach (2021), Zhang et al. (2025)
		Technology anxiety	5	Boughanzai et al. (2026), Fakhri et al. (2026), Karayianni et al. (2026), Shin et al. (2020), Zhang et al. (2025)
	Accuracy and reliability perception	System accuracy and reliability	8	Aladi (2024), Cui et al. (2025), Feldman-Maggor et al. (2025), Nazaretsky et al. (2025), Raza et al. (2024), Safarov et al. (2026), Schmidt et al. (2020), von Eschenbach (2021)
		Personal innovativeness and self-efficacy	6	Colclasure et al. (2025), Garcia (2026), Khosravi et al. (2022), Pandey et al. (2026), Türkmen (2025), Wang et al. (2026)

Nachouki (2026) demonstrate the decisive role of design features and ethical awareness on trust. Regarding feedback quality and pedagogical communication, the consistency of system feedback and instructors' attitudes toward AI tools appear to directly influence student trust. In terms of information quality and feedback mechanisms, Colclasure et al. (2025) and Jo (2024) report that the quality of system communication and instructor credibility significantly affect perceived trust.

Perceptual Factors and Trust Levels

Perceived usefulness emerges across the reviewed studies as the perceptual factor with the strongest effect on trust (**Table 4**). Eleven studies directly tested this relationship; Küçükaydın (2025) demonstrated that perceived usefulness fully mediates the relationship between academic stress and AI trust. Perceived ease of use is reported as a significant antecedent variable in the large majority of TAM-based studies. With regard to risk perception, Zhang et al. (2025) and Karayianni et al. (2026) show that perceived risk is a critical barrier that negatively affects trust. Accuracy and reliability perception is identified as the primary determinant of trust in multiple studies, most notably by Cui et al. (2025) and Nazaretsky et al. (2025).

Table 5. AI as digital authority and the trust relationship

Theme	Subtheme	Category	n	Studies
AI as digital authority and the trust relationship	Digital authority perception and adoption	Authority attribution and system trust	7	Colclasure et al. (2025), Pandey et al. (2026), Safarov et al. (2026), Schmidt et al. (2020), Shin et al. (2020), von Eschenbach (2021), Wang et al. (2026)
		Institutional support and social influence	8	Akman Yeşilel (2026), Aladi (2024), Boughanzai et al. (2026), Fakhri et al. (2026), Karran et al. (2025), Mustofa et al. (2025), Omirali et al. (2025), Schilhabel et al. (2026)
	Academic decision-making and behavioral intention	Usage intention through trust mediation	9	Alarefi and Alharbi (2025), Fakhri et al. (2026), Jo (2024), Karayianni et al. (2026), Marimon et al. (2025), Omirali et al. (2025), Teodorescu et al. (2023), Wang and Yu (2026), Zhang et al. (2025)
		AI trust and academic performance	6	Colclasure et al. (2025), Khlaif et al. (2025), Küçükaydın (2025), Raza et al. (2024), Shahzad et al. (2025), Shishakly and Nachouki (2026)
	Trust miscalibration	Overtrust and lack of critical evaluation	6	Feldman-Maggor et al. (2025), Garcia (2026), Nazaretsky et al. (2025), Safarov et al. (2026), Türkmen (2025), von Eschenbach (2021)
		Digital literacy and critical awareness	5	Akman Yeşilel (2026), Cui et al. (2025), Khosravi et al. (2022), Pandey et al. (2026), Shin et al. (2020)

Table 6. Methodological approaches and research gaps

Theme	Subtheme	Category	n	Studies
Methodological approaches and research gaps	Quantitative approaches	SEM	10	Alarefi and Alharbi (2025), Boughanzai et al. (2026), Fakhri et al. (2026), Karayianni et al. (2026), Marimon et al. (2025), Mustofa et al. (2025), Nazaretsky et al. (2025), Schilhabel et al. (2026), Wang and Yu (2026), Zhang et al. (2025)
		Survey and scale-based research	9	Colclasure et al. (2025), Jo (2024), Küçükaydın (2025), Nazaretsky et al. (2025), Omirali et al. (2025), Pandey et al. (2026), Shahzad et al. (2025), Shishakly and Nachouki (2026), Teodorescu et al. (2023)
	Review and meta-analysis studies	Systematic review and meta-analysis	4	Cui et al. (2025), Khosravi et al. (2022), Raza et al. (2024), Türkmen (2025)
	Qualitative and mixed methods	Mixed-method and qualitative design	7	Akman Yeşilel (2026), Aladi (2024), Colclasure et al. (2025), Jo (2024), Karran et al. (2025), Khlaif et al. (2025), Shin et al. (2020)
	Experimental and multi-group designs	Experimental and quasi-experimental design	4	Feldman-Maggor et al. (2025), Garcia (2026), Safarov et al. (2026), Wang et al. (2026)
		Conceptual and philosophical analysis	2	Schmidt et al. (2020), von Eschenbach (2021)

AI as Digital Authority and the Trust Relationship

The digital authority theme, addressed across multiple studies, reveals that AI systems are positioned by students as sources of epistemic authority (**Table 5**). Shin et al. (2020) argue that algorithms function as a lens through which people see the outside world, thereby structuring human perception. Colclasure et al. (2025) report that trust in AI-driven platforms directly affects perceptions of instructor credibility in the classroom. Safarov et al. (2026) found that students with lower expertise comply with AI recommendations without questioning them, feeding trust miscalibration. Regarding usage intention through trust mediation, Fakhri et al. (2026) and Marimon et al. (2025) demonstrate that trust operates as the primary mediating variable in the technology adoption process and shapes academic decision-making.

Methodological Approaches and Research Gaps

SEM stands out as the most prevalent methodological approach among the studies reviewed (**Table 6**), with ten studies employing this method. Survey-based cross-sectional designs are particularly prominent in scale development studies. Systematic review and meta-analysis studies (Cui et al., 2025; Khosravi et al., 2022; Türkmen, 2025) have made substantial contributions to the field's cumulative knowledge at the synthesis level. Although qualitative and mixed-method studies remain relatively limited, works such as Khlaif et al.

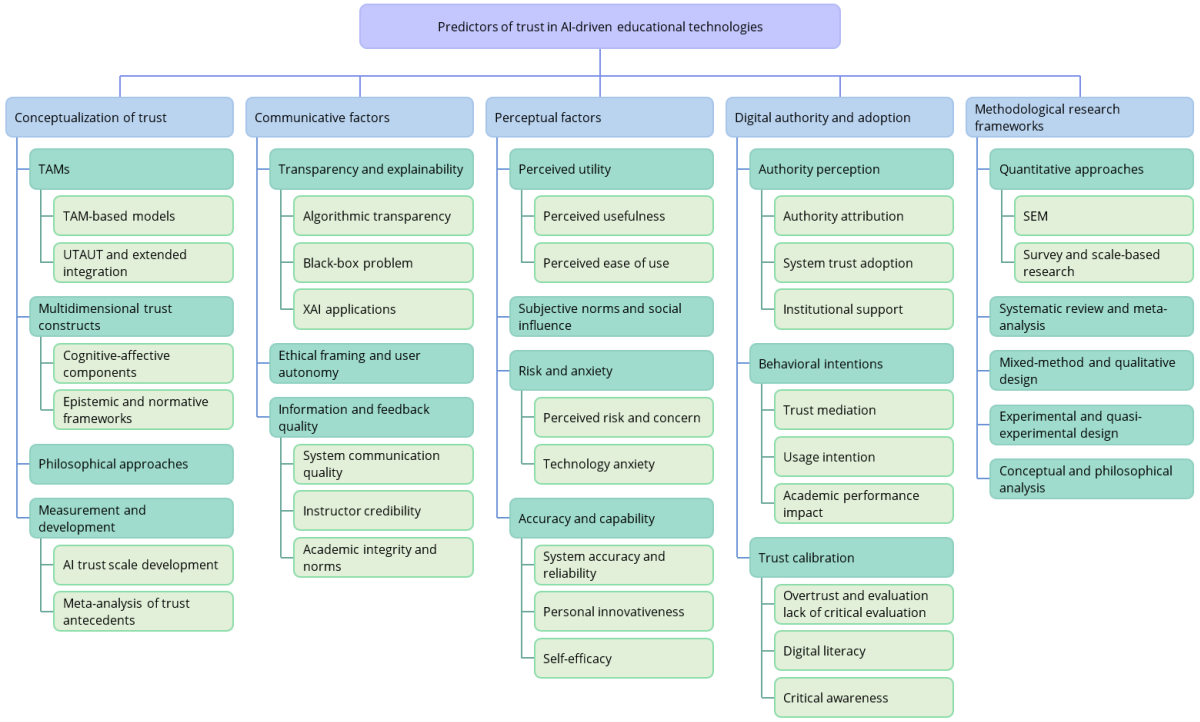


Figure 3. Predictors of trust in AI-driven educational technologies (the authors' own elaboration)

(2025) and Akman Yeşilel (2026) convey student and instructor perspectives in considerable depth. Experimental designs (Feldman-Maggor et al., 2025; Garcia, 2026; Safarov et al., 2026) offer a distinctive contribution to the field by testing the causal effects of intervention variables on trust. When research gaps are examined, longitudinal studies addressing long-term trust dynamics, cross-cultural comparative research, and in-depth qualitative analyses concerning trust miscalibration appear to remain insufficient. **Figure 3** summarizes the predictors identified across the five themes.

DISCUSSION

This study examined the factors shaping university students' trust in AI-supported educational technologies through a review of the relevant literature. The findings indicate that trust, within the context of educational technology, operates as a multivariable and dynamic construct. The predominance of TAM-based studies suggests that the theoretical orientation of the field continues to lean heavily on behavioral adoption models. The emphasis that UTAUT-based studies place on institutional support, performance expectancy, and effort expectancy points to trust formation as a process shaped at both individual and institutional levels. This finding aligns with Karran et al.'s (2025) multi-stakeholder perspective, underscoring the need for trust research to move beyond a student-centered focus and encompass the broader stakeholder system.

According to the research findings, the decisive influence of transparency and explainability on trust has been consistently confirmed in this study as well. Addressing the algorithmic black box problem (Schmidt et al., 2020; von Eschenbach, 2021) through XAI approaches strengthens trust; however, Safarov et al.'s (2026) findings draw attention to the importance of aligning explanations with the user's level of expertise. The observation that explanation quality is not absolute but varies according to context offers a significant insight into the need to place user profiles at the center of XAI design.

Ethical framing emerges as an independent determinant within the communicative dimension of trust, as evidenced by Garcia (2026), Shishakly and Nachouki (2026), and Karran et al. (2025). This finding suggests that trust is not solely tied to system performance but is equally bound up with how a system is presented and legitimized. In the context of pedagogical communication, Colclasure et al. (2025) and Akman Yeşilel (2026) highlight the role of instructor discourse in shaping student trust, a finding that speaks to the importance of AI literacy training at the institutional level.

Although this finding aligns with the TAM ecosystem, Lijie et al. (2025) demonstrate that positive perceptions of AI tools can strengthen critical thinking dispositions, drawing attention to the bidirectional relationship between trust and cognitive development. This suggests that perceived usefulness may go beyond being a mere adoption facilitator and potentially shape the learning process itself. Risk perception, on the other hand, emerges as the primary barrier negatively affecting trust. Rana et al. (2024) show that this effect is particularly pronounced in developing country contexts, indicating that risk perception is shaped by cultural and infrastructural conditions. This finding underscores the need for context-sensitive approaches that go beyond a universal trust model.

From the perspective of digital authority, AI systems are found to be positioned by students as epistemic reference points. Safarov et al. (2026) found that students with lower expertise accept AI recommendations without questioning them, creating conditions for trust miscalibration. This finding aligns with James et al.'s (2025) identification of digital literacy as a moderating variable of trust calibration. Trust miscalibration is rapidly becoming critical and remains a research gap in the literature. Students' uncritical acceptance of AI outputs may, over time, threaten academic autonomy and critical thinking development. Pandey et al. (2026) have contributed to the conceptualization of this threat within the epistemic trust framework. The fact that digital literacy moderates trust calibration indicates that integrating these skills into curricula should be prioritized. The methodological concentration of reviewed studies on SEM and cross-sectional designs hinders a sufficient understanding of how trust dynamics change over time. Aquilino et al. (2025) have drawn attention to theoretical and methodological inconsistencies in trust measurement, and Kazimova et al. (2025) have noted that the long-term effects of AI applications in education have not yet been adequately investigated. In this context, longitudinal, experimental, and mixed-method research appears capable of making original contributions to the field.

The study carries certain limitations. Only English-language publications indexed in the WoS and Scopus databases were examined, leaving studies in other languages outside the scope of the review. This may have resulted in findings from diverse geographical contexts being insufficiently represented. Additionally, the interpretive flexibility inherent in qualitative content analysis makes a degree of researcher judgment unavoidable, a limitation that was partially addressed through inter-rater reliability assessments. Drawing on these limitations, several recommendations can be offered. Longitudinal designs are needed to capture how trust dynamics shift over time. Mixed-method approaches should be prioritized to develop a more nuanced understanding of trust miscalibration. Beyond this, the integration of students, instructors, and other stakeholders into AI design processes can be facilitated through participatory design approaches. To enable university students to develop well-calibrated trust in AI-supported educational technologies, higher education institutions are encouraged to adopt comprehensive policies that bring together transparency-driven system design, instructor-mediated pedagogical communication, and curriculum-embedded digital literacy education.

CONCLUSION

This study examined the communicative and perceptual factors shaping university students' trust in AI-supported educational technologies through a systematic review and content analysis of 33 studies organized around five research questions. The findings reveal that trust, within the context of educational technology, displays a multidimensional, dynamic, and context-sensitive character.

In terms of communicative factors, transparency and explainability stand out as foundational conditions of trust. While the algorithmic black box problem is identified as a structural source of distrust, XAI approaches are positioned as design-level responses to this challenge. At the same time, the effectiveness of explanations was found to depend on their alignment with the user's level of expertise. Pedagogical communication and ethical framing further demonstrate that, alongside design features, the manner in which a system is presented constitutes an independent factor in shaping trust. Regarding perceptual factors, perceived usefulness emerged as the strongest antecedent. However, this relationship appears to be bidirectional and context-sensitive, while risk perception functions as a critical barrier variable that differs according to cultural and infrastructural conditions.

From the standpoint of digital authority, AI systems tend to be positioned by students as epistemic reference points, a dynamic that can set in motion a process feeding trust miscalibration. Digital literacy emerges as a key moderator of this process, pointing to the need for prioritizing curriculum integration. Methodologically, the current research landscape dominated by SEM and cross-sectional designs signals a pressing need for longitudinal and experimental approaches, qualitative depth studies, and cross-cultural comparative analyses. In this regard, a trust-centered design philosophy would ensure that AI systems are positioned not merely as functional tools but as legitimate instruments that actively support student well-being.

Author contributions: ENT: conceptualization, investigation, methodology, writing – original draft; **MSM:** data curation, formal analysis, visualization, writing – review & editing; **ENK:** methodology, validation, writing – review & editing; **ISA:** investigation, methodology, resources, writing – review & editing; **VLZ:** data curation, investigation, writing – review & editing; **LMS:** conceptualization, writing – review & editing. All authors approved the final version of the article.

Funding: The authors received no financial support for the research and/or authorship of this article.

Ethics declaration: This study is a systematic literature review based solely on previously published research; therefore, it does not involve any human or animal participants and does not require any ethical approval.

AI statement: No generative AI or AI-based tools were used in the preparation of this manuscript.

Declaration of interest: The authors declared no competing interest.

Data availability: Data generated or analyzed during this study are available from the authors on request.

REFERENCES

*Studies marked with an asterisk were included in the analysis.

*Akman Yeşilel, D. B. (2026). English language teachers' perspectives on digital trust in AI-supported education. *British Educational Research Journal, 52*(5), 3219-3240. https://doi.org/10.1002/berj.4209

*Aladi, C. C. (2024). IT higher education teachers and trust in AI-enabled ed-tech: Implications for adoption of AI in higher education. In *Proceedings of the 2024 Computers and People Research Conference* (pp. 1-16). ACM. https://doi.org/10.1145/3632634.3655852

*Alarefi, M., & Alharbi, N. N. (2025). From intention to impact: Predictors and learning outcomes of AI tool adoption among Saudi university students. *Journal of Information Science Theory and Practice, 13*(4), 69-82. https://doi.org/10.1633/JISTaP.2025.13.4.5

Aquilino, L., Di Dio, C., Manzi, F., Massaro, D., Bisconti, P., & Marchetti, A. (2025). Decoding trust in artificial intelligence: A systematic review of quantitative measures and related variables. *Informatics, 12*(3), Article 70. https://doi.org/10.3390/informatics12030070

Baldwin, C. A., & Nasirci, H. (2024). Technological advancements and international organizations in education: A comparative study of China and Myanmar in the context of deglobalization. *Anatolian Turkish Journal of Education, 6*(3), 313-333.

*Boughanzai, F., Ouhassan, Y., Majdoubi, R., & Hadjoudja, A. (2026). A comprehensive structural equation modeling analysis of factors influencing teacher acceptance of AI in education through an extended TAM framework. *Education and Information Technologies, 31*(3), 847-889. https://doi.org/10.1007/s10639-025-13833-w

*Colclasure, B. C., Ruth, T. K., Beasley, V., & Granberry, T. (2025). Examining student perceptions of AI-driven learning: User experience and instructor credibility in higher education. *Trends in Higher Education, 4*(4), Article 59. https://doi.org/10.3390/higheredu4040059

*Cui, Y. L., Zeng, M. L., Du, X. K., & He, W. M. (2025). What shapes learners' trust in AI? A meta-analytic review of its antecedents and consequences. *IEEE Access, 13*, 164008-164025. https://doi.org/10.1109/ACCESS.2025.3611367

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340. https://doi.org/10.2307/249008

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107-115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

*Fakhri, M. M., Isma, A., Sahabuddin, Syukur, P. A., & Rosidah. (2026). Trust-mediated AI continuance intention among pre-service teachers: Integrating UTAUT and the extended S-O-R-S framework with AI brain-rot exposure. *Daengku: Journal of Humanities and Social Sciences Innovation, 6*(1), 77-90. https://doi.org/10.35877/454RI.daengku4849

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. *FASEB Journal, 22*(2), 338-342. https://doi.org/10.1096/fj.07-9492LSF

*Feldman-Maggor, Y., Cukurova, M., Kent, C., & Alexandron, G. (2025). The impact of explainable AI on teachers' trust and acceptance of AI EdTech recommendations: The power of domain-specific explanations. *International Journal of Artificial Intelligence in Education, 35*(5), 2889-2922. https://doi.org/10.1007/s40593-025-00486-6

*Garcia, M. B. (2026). Transparency, ethical framing, and user agency as determinants of trust in AI-mediated assessment: Informing the design of trustworthy systems. *Evaluation Review*. https://doi.org/10.1177/0193841X261451124

James, E. E., Sampson, E. A., Usani, N. E., & Inyang, I. B. (2025). A principal component analysis of the factors influencing university students' trust in AI-based educational technologies. *African Journal of Advances in Science and Technology Research, 18*(1), 111-141. https://doi.org/10.62154/ajastr.2025.018.010691

*Jo, H. (2024). From concerns to benefits: A comprehensive study of ChatGPT usage in education. *International Journal of Educational Technology in Higher Education, 21*(1), Article 35. https://doi.org/10.1186/s41239-024-00471-4

*Karayianni, I., Klidas, A., & Karakitsou, C. (2026). Between a bot and a hard place: Student engagement, technology anxiety and ChatGPT trust in higher education. *TechTrends, 70*(1), 39-56. https://doi.org/10.1007/s11528-025-01124-7

*Karran, A. J., Charland, P., Trempe-Martineau, J., Ortiz de Guinea Lopez de Arana, A., Lesage, A.-M., Sénécal, S., & Léger, P.-M. (2025). Multi-stakeholder perspective on responsible artificial intelligence and acceptability in education. *npj Science of Learning, 10*(1), Article 44. https://doi.org/10.1038/s41539-025-00333-2

Kazimova, D., Tazhigulova, G., Shraimanova, G., Zatyneyko, A., & Sharzadin, A. (2025). Transforming university education with AI: A systematic review of technologies, applications, and implications. *International Journal of Engineering Pedagogy, 15*(1), 4-24. https://doi.org/10.3991/ijep.v15i1.50773

*Khlaif, Z. N., Salameh, N., Ajouz, M., Mousa, A., Itmazi, J., Alwawi, A., & Alkaissi, A. (2025). Using generative AI in nursing education: Students' perceptions. *BMC Medical Education, 25*(1), Article 926. https://doi.org/10.1186/s12909-025-07416-z

Khodjaev, B. K., Masalimova, A. R., Chistyakov, A. A., & Ay, Y. (2026). A meta-analysis of the effects of generative artificial intelligence on learning, cognitive, and linguistic outcomes. *Frontiers in Education, 11*. https://doi.org/10.3389/feduc.2026.1971476

Khoso, A. K., Honggang, W., & Darazi, M. A. (2026). Trust and attitude towards AI as pathways to creativity: A TAM model study of EFL students' digital literacy and AI acceptance. *Humanities and Social Sciences Communications, 13*(1), Article 69. https://doi.org/10.1057/s41599-025-06362-x

*Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence, 3*, Article 100074. https://doi.org/10.1016/j.caeai.2022.100074

Kraipiyaset, P., Neamchuchuen, T., Chanpramun, S., Cheerapakorn, P., Kongpha, R., Nilsook, P., & Keawbankrud, W. (2026). Agentic AI-driven creative media management in mass communication Education 5.0: A PRISMA-guided mixed-methods systematic review and bibliometric analysis. *Online Journal of Communication and Media Technologies, 16*(3), Article e202638. https://doi.org/10.30935/ojcmt/18689

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). SAGE. https://doi.org/10.4135/9781071878781

*Küçükaydın, M. A. (2025). Can adopt AI-powered educational technology buffer against academic stress and school burnout? A protective model proposal. *Psychology in the Schools, 62*(12), 5176-5185. https://doi.org/10.1002/pits.70068

Lijie, H., Mat Yusoff, S., & Mohamad Marzaini, A. F. (2025). Influence of AI-driven educational tools on critical thinking dispositions among university students in Malaysia: A study of key factors and correlations. *Education and Information Technologies, 30*(6), 8029-8053. https://doi.org/10.1007/s10639-024-13150-8

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.

*Marimon, F., Arias-Valle, M. B., Coria-Augusto, C. J., & Larrea Arnau, C. M. (2025). Del optimismo a la confianza: El impacto de ChatGPT en la confianza de los estudiantes en el aprendizaje asistido por IA [From optimism to confidence: The impact of ChatGPT on students' confidence in AI-assisted learning]. *RIED-Revista Iberoamericana de Educación a Distancia, 28*(2), 131-153. https://doi.org/10.5944/ried.28.2.43238

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). SAGE.

*Mustofa, R. H., Kuncoro, T. G., Atmono, D., Hermawan, H. D., & Sukirman. (2025). Extending the technology acceptance model: The role of subjective norms, ethics, and trust in AI tool adoption among students. *Computers and Education: Artificial Intelligence, 8*, Article 100379. https://doi.org/10.1016/j.caeai.2025.100379

*Nazaretsky, T., Mejia-Domenzain, P., Swamy, V., Frej, J., & Käser, T. (2025). The critical role of trust in adopting AI-powered educational technology for learning: An instrument for measuring student perceptions. *Computers and Education: Artificial Intelligence, 8*, Article 100368. https://doi.org/10.1016/j.caeai.2025.100368

*Omirali, A., Kozhakhmet, K., & Zhumaliyeva, R. (2025). Digital trust in transition: Student perceptions of AI-enhanced learning for sustainable educational futures. *Sustainability, 17*(17), Article 7567. https://doi.org/10.3390/su17177567

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

*Pandey, C. S., Mishra, P., Pandey, S. R., & Pandey, S. (2026). Epistemic trust in generative AI for higher education scale (ETGAI-HE scale). *AI & SOCIETY, 41*(2), 1387-1400. https://doi.org/10.1007/s00146-025-02566-6

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Blackwell Publishing. https://doi.org/10.1002/9780470754887

Rana, M. M., Siddiqee, M. S., Sakib, M. N., & Ahamed, M. R. (2024). Assessing AI adoption in developing country academia: A trust and privacy-augmented UTAUT framework. *Heliyon, 10*(18), Article e37569. https://doi.org/10.1016/j.heliyon.2024.e37569

*Raza, S., Fatima, I., Arif, S., Sharif, M., Jalal, M. S., & Muhammad, Z. (2024). The future of learning: Building trust and transparency in AI education. *Journal of Management Practices, Humanities and Social Sciences, 8*(3), 62-74. https://doi.org/10.33152/jmphss-8.3.6

*Safarov, C., Gadzinski, G., & Schlögl, S. (2026). In AI we trust? Exploring the role of explainable GenAI and expertise in education. *Human Factors, 68*(4), 487-503. https://doi.org/10.1177/00187208251401773

*Schilhabel, S. A., Sankaranarayanan, B., Subedi, M., & Muraski, J. (2026). Exploring the impact of artificial intelligence on psychological well-being and inclusivity in higher education. *Journal of Information, Communication and Ethics in Society*. https://doi.org/10.1108/JICES-09-2025-0238

*Schmidt, P., Biessmann, F., & Teubner, T. (2020). Transparency and trust in artificial intelligence systems. *Journal of Decision Systems, 29*(4), 260-278. https://doi.org/10.1080/12460125.2020.1819094

*Shahzad, M. F., Xu, S., & Zahid, H. (2025). Exploring the impact of generative AI-based technologies on learning performance through self-efficacy, fairness & ethics, creativity, and trust in higher education. *Education and Information Technologies, 30*(3), 3691-3716. https://doi.org/10.1007/s10639-024-12949-9

Shashayeva, G. K., Akhmetova, A. I., Tassilova, N. A., Beisenova, S. B., Nogayeva, A. K., Song, Y., & Kosshygulova, A. S. (2026). The role of social media use motivations in university students' adoption of AI-supported learning tools: The mediating effect of perceived usefulness. *Online Journal of Communication and Media Technologies, 16*(2), Article e202633. https://doi.org/10.30935/ojcmt/18592

*Shin, D., Zhong, B., & Biocca, F. A. (2020). Beyond user experience: What constitutes algorithmic experiences? *International Journal of Information Management, 52*, Article 102061. https://doi.org/10.1016/j.ijinfomgt.2019.102061

*Shishakly, R., & Nachouki, M. (2026). Academic integrity and students' ethical use of ChatGPT in higher education. *Journal of Information Technology Education: Research, 25*, Article 8. https://doi.org/10.28945/5730

*Teodorescu, D., Aivaz, K.-A., Vancea, D. P. C., Condrea, E., Dragan, C., & Olteanu, A. C. (2023). Consumer trust in AI algorithms used in e-commerce: A case study of college students at a Romanian public university. *Sustainability, 15*(15), Article 11925. https://doi.org/10.3390/su151511925

*Türkmen, G. (2025). The review of studies on explainable artificial intelligence in educational research. *Journal of Educational Computing Research, 63*(2), 277-310. https://doi.org/10.1177/07356331241310915

*von Eschenbach, W. J. (2021). Transparency and the black box problem: Why we do not trust AI. *Philosophy & Technology, 34*(4), 1607-1622. https://doi.org/10.1007/s13347-021-00477-0

*Wang, Y., Ni, Q., Wei, T., Xu, H., Liu, L., & He, L. (2026). How AI-assisted decision-making paradigms and explainability shape human-AI collaboration. *Sustainability, 18*(7), Article 3516. https://doi.org/10.3390/su18073516

*Wang, Y. Z., & Yu, R. F. (2026). Exploring the factors on the acceptance of generative artificial intelligence teaching assistants: The perspective of technology acceptance model. *International Journal of Human-Computer Interaction, 42*(2), 794-808. https://doi.org/10.1080/10447318.2025.2513581

Xue, L., Mahat, J., & Ghazali, N. (2026). Technology acceptance model in artificial intelligence in education: A meta-analysis. *SAGE Open, 16*(1), Article 21582440251409441. https://doi.org/10.1177/21582440251409441

*Zhang, C., Hu, M., Wu, W., Kamran, F., & Wang, X. (2025). Unpacking perceived risks and AI trust influences pre-service teachers' AI acceptance: A structural equation modeling-based multi-group analysis. *Education and Information Technologies, 30*(2), 2645-2672. https://doi.org/10.1007/s10639-024-12905-7

Zheltukhina, M. R., Farrakhov, V. N., Drobysheva, N. N., Tararina, L. I., Erokhova, N. S., Filippova, A. A., & Garnova, N. Y. (2026). Hot topics and frontiers in AI and communication research. *Online Journal of Communication and Media Technologies, 16*(4), Article e202650. https://doi.org/10.30935/ojcmt/19271

APPENDIX A

Table A1. PRISMA 2020 checklist (Page et al., 2021)

Section/topic	Item	Checklist item	Location where reported (manuscript)	Status	Notes
TITLE	1	Identify the report as a systematic review.	Title	Reported	Title ends with "A systematic review of trust in AI-driven educational technologies".
ABSTRACT	2	See the PRISMA 2020 for abstracts checklist.	Abstract	Partially reported	Unstructured abstract (243 words): aim, databases and period (2018-2026), 33 studies, synthesis method, main findings, implication. Registration and risk-of-bias information are not given.
INTRODUCTION	3	Describe the rationale for the review in the context of existing knowledge.	Introduction; theoretical framework	Reported	
	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Introduction (last paragraph); purpose of the study	Reported	Aim and five research questions.
METHODS	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Method – inclusion and exclusion criteria; data collection and analysis	Reported	Five inclusion and five exclusion criteria (2018-2026, English, peer-reviewed). Studies were grouped by research-question themes.
	6	Specify all databases, registers, websites, organizations, reference lists and other sources searched, with the date each was last searched.	Method – databases and search strategy	Reported	Web of Science and Scopus; last search dates are stated in the text.
	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Method – databases and search strategy; **Figure 2**	Reported	WoS terms and Scopus string with subject-area, language and 2018-2026 limits.
	8	Specify the methods used to decide whether a study met the inclusion criteria, including how many reviewers screened each record and whether they worked independently.	Method – PRISMA protocol	Partially reported	Four-stage selection is described; the number of screeners is not stated.
	9	Specify the methods used to collect data from reports, including how many reviewers collected data and any processes for obtaining or confirming data.	Method – data collection and analysis; validity and reliability	Partially reported	Data items listed; coding reliability checked by re-coding and an independent coder on 10 studies (agreement 85.0%-92.5%).
	10a	List and define all outcomes for which data were sought.	Purpose of the study; method – data collection and analysis	Reported	Trust and its communicative, perceptual and digital-authority predictors (research questions 1-5).
	10b	List and define all other variables for which data were sought.	Method – data collection and analysis; **Table 1**	Reported	Authors, year, journal, design, sample, variables, theoretical frameworks, key findings, research gaps.
	11	Specify the methods used to assess risk of bias in the included studies.	-	Not reported	No formal risk-of-bias assessment is described in the manuscript.
	12	Specify for each outcome the effect measure(s) used in the synthesis or presentation of results.	-	Not applicable	Qualitative content analysis; no effect measures were synthesized.
	13a	Describe the processes used to decide which studies were eligible for each synthesis.	Method – data collection and analysis	Reported	Each study assigned to at least one thematic category; overlapping studies placed in more than one.
	13b	Describe any methods required to prepare the data for presentation or synthesis.	Method – data collection and analysis	Reported	Coding of theoretical content and findings; category and sub-theme definitions.
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	**Table 1**-**Table 6**; **Figure 3** and **Figure 4**	Reported	
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s).	Method – research design; data collection and analysis	Reported	Inductive and deductive content analysis (Elo & Kyngäs, 2008; Krippendorff, 2019).

Table A1 (Continued).

Section/topic	Item	Checklist item	Location where reported (manuscript)	Status	Notes
	13e	Describe any methods used to explore possible causes of heterogeneity among study results.	-	Not applicable	No statistical synthesis; thematic comparison only.
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	-	Not applicable	No statistical synthesis.
	14	Describe any methods used to assess risk of bias due to missing results in a synthesis.	-	Not reported	Not described; restriction to English-language WoS/Scopus records is noted as a limitation.
	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	-	Not applicable	Not performed (qualitative synthesis).
RESULTS	16a	Describe the results of the search and selection process, from the number of records identified to the number of studies included, ideally using a flow diagram.	Method – PRISMA protocol; **Figure 1**	Reported	164 records (WoS 46; Scopus 118) → 139 after 25 duplicates → 62 excluded at title/abstract → 77 full texts → 44 excluded → 33 included.
	16b	Cite studies that might appear to meet the inclusion criteria but were excluded and explain why they were excluded.	Method – PRISMA protocol; **Figure 1**	Partially reported	Exclusion at full-text stage (n = 44) is stated against the exclusion criteria; a study-by-study list is not provided.
	17	Cite each includes study and present its characteristics.	**Table 1**; references (marked with *)	Reported	33 included studies.
	18	Present assessments of risk of bias for each included study.	-	Not reported	See item 11.
	19	For all outcomes, present, for each study, summary statistics for each group and an effect estimate with its precision, ideally using structured tables or plots.	**Table 1**-**Table 6**	Not applicable	Study-level findings are presented thematically; no effect estimates were pooled.
	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Findings (**Table 2**-**Table 6**)	Partially reported	Characteristics summarized; risk of bias not assessed.
	20b	Present results of all statistical syntheses conducted.	-	Not applicable	No statistical synthesis.
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	-	Not applicable	
	20d	Present results of all sensitivity analyses conducted.	-	Not applicable	
	21	Present assessments of risk of bias due to missing results for each synthesis assessed.	-	Not reported	See item 14.
	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	-	Not applicable	See item 15.
DISCUSSION	23a	Provide a general interpretation of the results in the context of other evidence.	Discussion	Reported	
	23b	Discuss any limitations of the evidence included in the review.	Discussion (limitations); findings – methodological approaches and research gaps	Reported	Cross-sectional, SEM-dominated designs; scarce longitudinal and cross-cultural work.
	23c	Discuss any limitations of the review processes used.	Discussion (limitations)	Reported	English-only, WoS/Scopus-only, interpretive flexibility of content analysis.
	23d	Discuss implications of the results for practice, policy, and future research.	Discussion; conclusion	Reported	

Table A1 (Continued).

Section/topic	Item	Checklist item	Location where reported (manuscript)	Status	Notes
OTHER INFORMATION	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	-	Not reported	Not stated in the manuscript.
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	-	Not reported	Not stated in the manuscript.
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	-	Not applicable	See item 24a and item 24b.
	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors.	-	Not reported	Not stated in the manuscript text (see journal declarations).
	26	Declare any competing interests of review authors.	-	Not reported	Not stated in the manuscript text (see journal declarations).
	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	**Table 1-Table 6**; **Figure 2**	Partially reported	Extracted study-level data are given in Table 1-Table 6; the search strings are in method. Other materials are not stated.

Table A2. PRISMA 2020 flow numbers (consistent with **Figure 1** and the PRISMA protocol section)

Stage (as in **Figure 1**)	n
Records identified from databases (WoS n = 46; Scopus n = 118)	164
Duplicate records removed	25
Records screened (title/abstract)	139
Records excluded (out of scope)	62
Full-text reports assessed for eligibility	77
Full-text reports excluded	44
Studies included in the review	33