RAISE the Standard: A Framework for Transparent Reporting of Artificial Intelligence Studies in Education

Abstract:

The rapid integration of artificial intelligence (AI) into educational research and practice has highlighted the need for clear and consistent reporting standards. The RAISE (Reporting AI Studies in Education) framework offers a structured checklist of 30 items across ten thematic domains, designed to guide authors in transparently documenting AI interventions, study design, learner context, data collection, outcomes, and findings. This editorial introduces RAISE, explains its rationale, and provides practical guidance for its application, including illustrative “Bad” and “Good” reporting examples. By promoting comprehensive and replicable reporting, RAISE aims to enhance the interpretability, reproducibility, and scholarly impact of AI-focused educational research, while supporting authors in meeting emerging expectations for transparency in the field.

Keywords:

RAISE framework, AI in education, AI interventions, reporting standards, transparency, reproducibility

In the last few years, the research landscape on educational computing has shifted rapidly and irrevocably. The proliferation of artificial intelligence (AI) tools, particularly large language models, generative systems, and adaptive learning technologies, has reshaped how educational interventions are designed, delivered, and evaluated (Allison, et al., 2025; Guizani et al., 2025). These transformations have been widely embraced in classrooms and research labs alike, bringing both genuine innovation and new kinds of complexity. For journals like the Journal of Educational Computing Research, this presents both an opportunity and a challenge.

While AI has long been used to power adaptive learning systems (Wang et al., 2024), recommendation engines (Deschênes, 2020), and predictive analytics in learning contexts (Liz-Domínguez et al., 2019), what distinguishes today's wave of AI integration is its accessibility, opacity, and generativity. Tools such as ChatGPT, Claude, and Gemini, as well as custom AI classifiers, adaptive feedback engines, and multimodal platforms, are now routinely deployed by researchers without detailed technical description, theoretical grounding, or pedagogical rationale. Too often, submitted manuscripts provide little clarity on what AI tools were used, how they were configured, what data they processed, or what educational logic they followed. Reviewers are left asking:

- What exactly was the AI doing?
- Was it necessary?
- Is this replicable?
- Are the learning claims credible?

Editors must assess submissions for rigour, ethics, and contribution, often without the information necessary to do so transparently or consistently. In response to these concerns, we are introducing the Reporting Artificial Intelligence Studies in Education (RAISE) framework, as a new set of guidelines for authors submitting work which includes AI as a key contributing factor to the pedagogical intervention or approach, and as a broader contribution to the educational research community.

This editorial explains why RAISE is needed, how it works, and how it supports authors, reviewers, and editors alike. Our goal is to foster transparency, reflexivity, and shared standards in an area where novelty has too often outpaced scrutiny.

Why We Need a Reporting Standard for AI in Education

Research involving AI is not simply about introducing new tools; it fundamentally entails shifts in the conceptual frameworks and epistemological foundations underpinning educational inquiry (Rismanchian & Doroudi, 2025). Every AI system reflects choices about data, design, power, and pedagogy. These systems generate feedback, scaffold learning, model cognition, or even produce curriculum. Yet without consistent reporting, it is impossible to understand how the technology interacts with learners, educators, or knowledge itself.

As editors, we routinely observe manuscripts that use AI without stating even basic information about:

- Which model was used (e.g., GPT-4, Claude, or a custom-built algorithm).
- How the system was configured (e.g., prompts, fine-tuning parameters).
- What role the AI played in the learning experience (e.g., feedback generator, co-author, tutor, evaluator).
- Whether human actors designed or reviewed the outputs.
- How limitations, bias, or ethical risks were addressed.

This is not just a matter of transparency, it is a matter of scholarly credibility. AI is not neutral; its effects depend on how it is situated, deployed, and understood. A lack of standardised reporting creates problems for reproducibility, peer review, meta-analysis, and ultimately, impact. In this context, the RAISE framework offers a shared vocabulary and minimal reporting standard to elevate the quality and consistency of published research.

Introducing the RAISE Framework

RAISE stands for Reporting AI Studies in Education. It is a structured reporting checklist comprising 30 core items. See Figure 1 for the full RAISE Checklist. An editable version is also available as supplementary material for usage by authors, reviewers, and editors alike.

Section and Topic	Item	Checklist Item	Location in Manuscript
EDUCATIONAL JUSTIFICATION AND THEORETICAL GROUNDING			
Educational Problem	1	Clearly articulate the educational problem or learning objective that the AI system is intended to address.	
AI Justification	2	Justify the use of AI in this context and articulate relevant educational or learning theory.	
AI SYSTEM SPECIFICATION			
AI Model	3	Specify the AI model used (name, version, provider, static/adaptive/generative).	
AI Access	4	Indicate whether the system is proprietary or open-source, and any limitations on access.	
System/Prompt Design	5	Detail the design process of the AI system and/or prompt engineering (manual, iterative, instructional logic), including representative examples.	
AI Settings	6	Report how AI settings or configurations were determined (e.g., temperature, filters).	
AI ROLE AND INTERACTION			
AI Instructional Role	7	Describe what instructional functions were delegated to the AI (e.g., explaining, recommending, guiding).	
AI Assessment Role	8	Describe what assessment or evaluative functions (e.g., marking, detecting errors) were performed by the AI.	
AI and Student/ Teacher Interaction	9	Specify how students and/or teachers interacted with the AI (e.g., inputs, outputs, dialogue, dashboard).	
AI Output Alignment	10	Explain how AI-generated feedback or outputs were aligned with learner cognitive levels and curricular goals.	
AI Structuring	11	Describe any scaffolding, pacing, or feedback strategies to avoid overload or confusion.	
ACCESSIBILITY AND CULTURAL FIT			
Cultural Adaptation	12	State the language(s) of use and whether localisation or adaptation was needed to ensure cultural relevance (e.g., of the AI system or study design).	
Accessibility	13	Report any steps taken to ensure accessibility.	
EDUCATIONAL SETTING AND PARTICIPANTS			
Educational Context	14	Specify the educational context and delivery mode (e.g., course, platform, subject).	
Learner Context	15	Provide demographics and relevant characteristics of learners (e.g., age, prior knowledge).	
HUMAN INVOLVEMENT			
Intervention Roles	16	Specify the role of teachers, facilitators, or others (e.g., administrators) during the intervention.	
Designer Expertise	17	Clarify who designed the systems/prompts and whether they had pedagogical or AI expertise.	
STUDY DESIGN AND EVALUATION			
Research Design	18	Justify the research design (e.g., quasi-experimental, RCT, qualitative, mixed-methods).	
Outcome Measures	19	Describe the outcome measures used (e.g., tests, engagement).	
Outcome Alignment	20	Explain how the outcomes align with the intended learning objectives.	
Data Collection	21	Describe data collection instruments and procedures (e.g., logs, tests, surveys, interviews).	
Presenting Findings	22	Summarise key findings including learning, engagement, or behavioural outcomes.	
ETHICS AND TRUSTWORTHINESS			
Ethical Approval	23	Confirm ethical approval and informed consent procedures.	
Data Privacy	24	Describe data privacy and participant anonymity safeguards.	
AI Risks	25	Identify known risks of bias, hallucination, or misinformation and how they were mitigated.	
TRANSPARENCY AND REPRODUCIBILITY			
Prompt and Model Transparency	26	Where possible, provide access to representative prompts, interaction examples, model code, or system specifications.	
Replication	27	Describe any constraints or limitations for replicating the system or study.	
LIMITATIONS AND IMPLICATIONS			
Unintended Consequences	28	Reflect on unintended consequences (e.g., over-reliance on AI, loss of agency).	
Future Directions	29	Propose future directions or improvements for AI-enhanced learning.	
Implications	30	Discuss implications for practice, curriculum, or policy.	

Figure 1 - RAISE 2025 Checklist (v1.0): Reporting AI Studies in Education

As shown in Figure 1, the 30 items are grouped into ten thematic domains, where each domain contains specific items that ask authors to clarify what was done, why, how, and with what implications. To support authors in translating the checklist into practice, we also provide a RAISE 2025 Reporting Examples and Explanation document within the supplementary material. This document presents side-by-side “Bad” and “Good” reporting examples for each checklist item, accompanied by explanations that illustrate how to achieve clarity, transparency, and alignment with educational objectives. Below is the inclusion rationale for each of the ten domain areas.

1. **Educational Justification and Theoretical Grounding.** AI interventions in classrooms cannot be regarded as neutral tools; their educational value emerges only when they are explicitly aligned with clear statements of the learning problem, a coherent theoretical rationale, and transparent mapping between intended learning objectives and outcome measures. Recent empirical work illustrates the risks of neglecting this alignment. For instance, studies of ChatGPT-based aids and related AI pedagogies demonstrate that without grounding in frameworks such as self-regulated learning, there is a danger of conflating short-term task performance with lasting learning gains (Wu et al., 2023). Similarly, research on analogy-based AI pedagogy shows that making the pedagogical logic explicit, such as clarifying how analogical reasoning supports conceptual change, not only highlights mechanisms of effect but also enhances the generalisability of findings across diverse learning contexts (Dai et al., 2023).

   These examples underline a broader need for consistent reporting practices in educational technology research involving AI. When authors clearly articulate the educational problem under study, the theoretical framing that motivates their intervention, and the alignment between objectives and outcomes, the resulting work becomes more interpretable and comparable across contexts. Such transparency enables the scholarly community, including editors, reviewers, and readers to judge whether reported effects reflect sound pedagogy rather than the novelty of the technology itself (Topali et al., 2025).

2. **AI System Specification.** Recent educational work emphasises that reporting the model identity, version, provider, access conditions (open vs proprietary), and key configuration choices is essential for interpreting results and enabling replication. Prompt outcomes and learner effects can change substantially with different model versions or settings, so transparent details regarding specifications is not optional but core methodological information. For instance, the systematic review of prompt engineering in higher education by Lee and Palmer (2025) and Federiakin's et al (2024) argument that prompt engineering is an emergent methodological skill both stress that model and setting details materially shape educational outputs, so reporting on these details is crucial for replicability and understanding. For instance, authors could refer to Qian's (2025) systematic review of approaches and educational applications of prompt engineering in education to help clarify what prompting strategies have been used.
3. **AI Role and Interaction.** In educational contexts, AI systems can occupy a range of pedagogical roles such as tutor, feedback agent, or scaffolded conversational partner, and the ways in which students interact with these systems critically shape both learning outcomes and ethical considerations. Recent research illustrates this point clearly. Studies of generative AI chatbots show that well-designed interaction patterns, emphasising explanation, clarification, and integration into learners' existing frameworks, can support knowledge-building, elaboration, and collaborative reasoning (Song et al., 2025). Investigations into ChatGPT in programming education further demonstrate that dynamic, context-aware interaction, including explanation of code, debugging support, and adaptive conversational style, fosters engagement, confidence, and technical

mastery, while also surfacing risks of misuse and overreliance (Guner & Er, 2025). Complementary quasi-experimental work has shown that conversational agents providing both cognitive and metacognitive scaffolding enhance students' engagement and reasoning in collaborative writing tasks (Hu et al., 2025).

Taken together, these studies point to a clear implication for reporting practice: authors should describe not only the AI system's instructional role but also the design of learner–AI interaction in sufficient detail, including prompts, scaffolds, modality, and dialogue structures. Such transparency strengthens interpretability and allows findings to be compared and translated across contexts, enabling the field to distinguish between pedagogical effects and the contingencies of specific interaction designs.

4. **Accessibility and Cultural Fit.** Accessibility and localisation are not peripheral; they determine whether an AI tool is usable by learners with disabilities or by students in different linguistic and cultural contexts. AI systems should be designed to accommodate diverse learning needs (Funa & Gabay, 2025), while recent empirical work documents both the promise of generative AI for accommodations and the real risks of exclusion when accessibility and localisation are not addressed (Zhao et al., 2025). Hence, Stefaniak and Moore (2024) argues that designers must report inclusivity and localisation choices to make findings meaningful across contexts. These findings support explicit items on language, adaptation, and accessibility provisions within the RAISE framework.
5. **Educational Setting and Participants.** Contextual factors such as course type, subject area, delivery mode, and learner characteristics (age, prior knowledge, socio-demographics) fundamentally shape internal and external validity. Both

Topali et al (2025) and Cingillioglu et al (2024) document how a lack of contextual detail impedes interpretation of purported learning gains and masks important subgroup effects. Therefore, standard reporting of these contextual details is important so stakeholders can assess generalisability and equity implications.

6. **Human Involvement.** Outcomes frequently depend on the human ecosystem around the AI. Teacher facilitation, who designed prompts or systems, and the level of human oversight all moderate effectiveness and safety. Recent literature on human–AI collaboration in education stresses that it should be reported who engineered prompts/systems and the pedagogical expertise involved. This is because designer background and expertise influence the classroom environment and generative AI outputs. For instance, Knoth et al (2024) found that individuals with higher-quality prompt engineering skills predict the quality of LLM outputs, while Jacobsen and Weber (2025) conclude that educators must be skilled in prompt engineering and utilising AI tools to achieve optimal results. Therefore, the RAISE framework calls for transparent reporting on both the role and expertise of those involved in designing and utilising AI tools within education.
7. **Study Design and Evaluation.** Credible claims about AI's educational impact depend on transparent justification of research design (Funa & Gabay, 2025), careful selection and description of validated outcome measures, and alignment of those measures with stated learning objectives. A recent systematic review and meta-analysis synthesised experimental studies of ChatGPT and found substantial heterogeneity in designs and measures, which complicates synthesis and inference translation (Deng et al., 2025). The review therefore recommends that authors justify design choices and provide detailed measurement

descriptions to support meta-analytic aggregation and policy translation (Deng et al., 2025).

8. **Ethics and Trustworthiness.** AI deployments raise well-documented ethical concerns including privacy, consent, bias, hallucination/misinformation and fairness that must be reported and mitigated. Systematic reviews and ethics syntheses repeatedly recommend mandatory reporting of ethical review, consent processes, data governance, and explicit description of bias and hallucination mitigation strategies so readers can evaluate potential harms (Fu & Weng, 2024; García-López & Trujillo-Liñán, 2025; Gouseti et al., 2025). Therefore, there needs to be clear reporting of safeguards and review processes of AI related interventions within education.
9. **Transparency and Reproducibility.** Education researchers increasingly recognise that sharing representative prompts, interaction transcripts, model specifications and replication constraints dramatically improves reproducibility and trust. A lack of transparency can lead to the questioning of the validity of an AI intervention, impacting its overall credibility and acceptance within educational contexts (Funa & Gabay, 2025). Hence, it is recommended that authors provide interaction examples (e.g., prompts used) and document limitations to replication (Lee & Palmer, 2025).
10. **Limitations and Implications.** Explicit reflection on unintended consequences (e.g., over-reliance, shifting teacher roles), limitations, and practical/policy implications increases the utility of research for practitioners and policymakers. Recent syntheses recommend that authors go beyond presenting positive outcomes to meaningfully discuss generalisability, constraints on replication, and next steps for safer, equitable deployment (Fu & Weng, 2024). The editorial board

of the Journal of Educational Computing Research have raised concerns regarding the ethical, emotional, and socio-cultural impacts of AI in education, and have recommended greater consideration is placed on exploring the indirect effects and unintended consequences of our work (Allison, et al., 2025). Hence, the RAISE framework explicitly highlights reporting on unintended consequences, future directions, and implications of work utilising AI in education.

The 30 checklist items within the 10 thematic domain areas of the RAISE framework are not arbitrary requirements, as they respond to genuine calls from existing literature and gaps we have observed in AI-related submissions. Authors are therefore encouraged to consider the RAISE checklist when submitting AI-related manuscripts to this journal. The framework does not mandate any particular method, model, or ideology; rather, it insists that whatever choices are made, they are made visible.

As previously mentioned, to support authors in applying the RAISE framework, we have also provided in the supplementary material the RAISE 2025 Reporting Examples and Explanation document, which illustrates realistic “Bad” and “Good” reporting practices for each checklist item. This document offers concrete guidance on how to describe AI systems, learner context, intervention design, data collection, outcomes, and findings with sufficient clarity and transparency. By contrasting insufficiently detailed approaches with those aligned with good practice, the document demonstrates how RAISE can be operationalised in practice, helping authors anticipate potential gaps in reporting and improve the reproducibility, interpretability, and impact of their AI-focused educational research.

Furthermore, in parallel with the proposed 30-item RAISE checklist, we also introduce a supplementary tool: the RAISE Ethics and Risk Matrix, see Figure 2. This is also available as an editable version in the supplementary material which authors can use.

Risk Theme	Risk Category	Description	Observed or Anticipated Impact (Author to specify)	Mitigation/Safeguards (Author to specify)
Pedagogical Integrity and Learning Impact	Loss of Learner Agency	Potential over-reliance on AI that may reduce learner autonomy and critical thinking.		
	Misuse or Overuse of AI	Risk of AI being used beyond intended scope or without adequate oversight.		
	Emotional Impact and Wellbeing	Potential negative emotional effects from AI feedback (e.g., anxiety, discouragement).		
	Data Misinterpretation	Risk that educators or learners may misinterpret AI-generated analytics or feedback.		
	Hallucination / Misinformation	Risk that AI may generate incorrect or misleading information.		
	Transparency and Explainability	Lack of clarity about how AI feedback or decisions are generated.		
Equity, Inclusion, and Accessibility	Bias and Fairness	Risk that AI may perpetuate or amplify biases (e.g., socio-economic, gender, race).		
	Cultural Insensitivity	AI outputs or prompts that may be culturally inappropriate or irrelevant.		
	Accessibility Barriers	Risks of excluding learners with disabilities or diverse needs.		
	Equity in Access to Technology	Unequal access to devices or connectivity impacting participation.		
Data Governance and Learner Rights	Privacy and Data Security	Risks related to unauthorized access, data leakage, or privacy breaches.		
	Informed Consent and Autonomy	Concerns over learner awareness of AI involvement and data use.		
	Data Ownership and Usage Rights	Uncertainties about who owns student-generated data and AI outputs.		
Technical Dependability and System Reliability	Technical Failures / Downtime	Possible interruptions due to AI system malfunctions or outages.		
	Dependence on Proprietary Systems	Concerns about transparency, replicability, and vendor lock-in with closed-source AI.		
	Algorithmic Accountability	Unclear assignment of responsibility for AI-generated outcomes or errors.		

Figure 2 - RAISE 2025 Ethics and Risk Matrix (v1.0)

While ethical approval and data privacy are already addressed within the core RAISE framework (items 23-25), the RAISE 2025 Ethics and Risk Matrix invites authors to go further by critically reflecting on the diverse and often under-reported risks that can emerge when AI is deployed in educational contexts. These include, but are not limited to, risks to learner agency, equity in access, data governance, and algorithmic transparency. Rather than assuming that ethical approval equates to comprehensive ethical consideration, this matrix provides a structured space for unpacking nuanced pedagogical, technical, and social risks that may surface throughout the AI system lifecycle.

Designed as a fillable table that could be incorporated within manuscripts, the RAISE 2025 Ethics and Risk Matrix encourages authors to identify relevant risks according to thematic categories and offer context-specific descriptions of their potential impacts and mitigation strategies. It is intended to support both transparency in reporting and a proactive approach to trustworthiness and inclusivity. While its use is optional, we strongly recommend its inclusion, especially for studies involving generative or adaptive AI tools, complex student interactions, or deployment at scale. By surfacing these considerations, the matrix complements the main RAISE framework and fosters more responsible, context-sensitive scholarship in AI and education research.

Who is RAISE For?

RAISE is intended for all researchers conducting empirical studies that involve AI as part of the educational experience, whether directly or indirectly. This includes, but is not limited to, studies using AI for:

- Personalised tutoring or feedback.
- Automated grading or assessment.
- Curriculum generation or adaptation.
- Predictive analytics or learning analytics.
- Conversational or dialogic learning tools.
- Content creation or co-writing.
- Simulated peer collaboration or role-play.

It applies across methodologies; qualitative, quantitative, mixed-methods, design-based, and across settings, from formal K–12 to higher education and informal learning. We recognise that AI use may be central in some studies and peripheral in others; the framework is designed to be flexible, scalable, and responsive. Importantly, RAISE is also meant for reviewers, providing a clearer sense of what questions to ask of AI-enabled research, and for editors, who must ensure that published work meets standards of ethical accountability and methodological transparency.

A Framework Built on Reflection, Not Compliance

Some may worry that RAISE could be seen as bureaucratic or burdensome. To the contrary, we view it as a prompt for scholarly reflection. AI technologies challenge us to ask fundamental questions: What is teaching? What is learning? Who decides? These questions cannot be answered by code or performance metrics alone. They require researchers to surface assumptions, contextualise findings, and consider the social, cultural, and ethical dimensions of AI use. The RAISE framework includes not just technical detail, but also prompts for reflection on limitations, risks, human involvement, and cultural context. It invites authors to articulate how AI supports or constrains educational values, such as equity, autonomy, and engagement. In this way, RAISE promotes both methodological clarity and moral imagination.

How RAISE Aligns with Broader Movements

RAISE is not alone. It builds on a growing set of efforts in scholarly communities to improve transparency and reporting standards. Existing frameworks such as CONSORT (Consolidated Standards of Reporting Trials) (Schulz et al., 2010) and PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) (Page et al., 2021) have transformed reporting, particularly in clinical settings. Meanwhile, initiatives like Datasheets for Datasets (Gebru et al., 2021) have emerged to increase transparency and accountability within the machine learning community.

RAISE contributes to this momentum by rooting these ideas in the context of learning sciences, instructional design, and education technology. It acknowledges that AI in education is not merely technical, it is pedagogical, cultural, and situated. Our goal is to ensure that the reporting of such research meets the intellectual and ethical standards expected in this field.

Looking Ahead

The RAISE framework is a living initiative. As AI technologies evolve, so too will the needs and responsibilities of researchers. This editorial serves as a beginning of a community conversation, not its endpoint. To that end, we welcome collaboration and critique, and are particularly interested in how RAISE can be adapted across global contexts, languages, and pedagogical traditions.

The educational research community has a responsibility not only to explore what AI can do, but to document and explain what it is doing, and with what consequences. RAISE is one step toward that goal. It is a scaffold for better reporting, a catalyst for better questions, and a call for collective care in a rapidly transforming field. Let us RAISE the standard, together.

References

Allison, J., Hwang, G.-J., Mayer, R. E., Pellas, N., Karnalim, O., de Freitas, S., Ng, O.-L., Huang, Y.-M., Hooshyar, D., Seidman, R. H., Al-Emran, M., Mikropoulos, T. A., Schroeder, N. L., Roscoe, R. D., & Sanusi, I. (2025). From generative AI to extended reality: Multidisciplinary perspectives on the challenges, opportunities, and future of educational computing. *Journal of Educational Computing Research*, 0(0). https://doi.org/10.1177/07356331251359964

Cingillioglu, I., Gal, U., & Prokhorov, A. (2024). AI-experiments in education: An AI-driven randomized controlled trial for higher education research. *Education and Information Technologies*, 29, 19649–19677. https://doi.org/10.1007/s10639-024-12633-y

Dai, Y., Lin, Z., Liu, A., Dai, D., & Wang, W. (2023). Effect of an analogy-based approach of artificial intelligence pedagogy in upper primary schools. *Journal of Educational Computing Research*, 61(8), 1695–1722. https://doi.org/10.1177/07356331231201342

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 105224. https://doi.org/10.1016/j.compedu.2024.105224

Deschênes, M. (2020). Recommender systems to support learners' agency in a learning context: A systematic review. *International Journal of Educational Technology in Higher Education*, 17(1), 50. https://doi.org/10.1186/s41239-020-00219-w

Federiakin, D., Molerov, D., Zlatkin-Troitschanskaia, O., & Maur, A. (2024). Prompt engineering as a new 21st-century skill. *Frontiers in Education*, 9, Article 1366434. https://doi.org/10.3389/feduc.2024.1366434

Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centred AI practices. *Computers & Education: Artificial Intelligence*, 7, Article 100306. https://doi.org/10.1016/j.caeai.2024.100306

Funa, A. A., & Gabay, R. A. E. (2025). Policy guidelines and recommendations on AI use in teaching and learning: A meta-synthesis study. *Social Sciences & Humanities Open*, 11, 101221. https://doi.org/10.1016/j.ssaho.2024.101221

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

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., III, & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, 64(12), 86–92. https://doi.org/10.1145/345872

Gouseti, A., James, F., Fallin, L., & Burden, K. (2025). The ethics of using AI in K-12 education: A systematic literature review. *Technology, Pedagogy and Education*. Advance online publication. https://doi.org/10.1080/1475939X.2024.2428601

Guizani, S., Mazhar, T., Shahzad, T., Ahmad, W., Bibi, A., & Hamam, H. (2025). A systematic literature review to implement large language models in higher education: Issues and solutions. *Discover Education*, 4(1), 1–25. https://doi.org/10.1007/s44217-025-00424-7

Güner, H., & Er, E. (2025). AI in the classroom: Exploring students' interaction with ChatGPT in programming learning. *Education and Information Technologies*, 30, 12681–12707. https://doi.org/10.1007/s10639-025-13337-7

Hu, W., Tian, J., & Li, Y. (2025). Enhancing student engagement in online collaborative writing through a generative AI-based conversational agent. *The Internet and Higher Education*, 65, 100979. https://doi.org/10.1016/j.iheduc.2024.100979

Jacobsen, L. J., & Weber, K. E. (2025). The promises and pitfalls of large language models as feedback providers: A study of prompt engineering and the quality of AI-driven feedback. *AI*, 6(2), 35. https://doi.org/10.3390/ai6020035

Knoth, N., Tolzin, A., Janson, A., & Leimeister, J. M. (2024). AI literacy and its implications for prompt engineering strategies. *Computers & Education: Artificial Intelligence*, 6, 100225. https://doi.org/10.1016/j.caeai.2024.100225

Lee, D., & Palmer, E. (2025). Prompt engineering in higher education: A systematic review to inform curricula. *International Journal of Educational Technology in Higher Education*, 22, Article 12. https://doi.org/10.1186/s41239-025-00503-7

Liz-Domínguez, M., Caeiro-Rodríguez, M., Llamas-Nistal, M., & Mikic-Fonte, F. A. (2019). Systematic literature review of predictive analysis tools in higher education. *Applied Sciences*, 9(24), 5569. https://doi.org/10.3390/app9245569

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., McGuinness, L. A., Selby, P., Thomas, J., Tricco, A., Welch, V., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, 10, Article 89. https://doi.org/10.1186/s13643-021-01626-4

Qian, Y. (2025). Prompt engineering in education: A systematic review of approaches and educational applications. *Journal of Educational Computing Research*, 0(0). https://doi.org/10.1177/07356331251365189

Rismanchian, S., & Doroudi, S. (2025). The evolution of research on AI and education across four decades: Insights from the AIxEd framework. *International Journal of Artificial Intelligence in Education*, 1–24. https://doi.org/10.1007/s40593-025-00483-9

Schulz, K. F., Altman, D. G., Moher, D., & CONSORT Group. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMC Medicine*, 8, 18. https://doi.org/10.1186/1741-7015-8-18

Song, Y., Huang, L., Zheng, L., Fan, M., & Liu, Z. (2025). Interactions with generative AI chatbots: Unveiling dialogic dynamics, students' perceptions, and practical competencies in creative problem-solving. *International Journal of*

Educational Technology in Higher Education, 22, 12. https://doi.org/10.1186/s41239-025-00508-2

Stefaniak, J. E., & Moore, S. L. (2024). The use of generative AI to support inclusivity and design deliberation for online instruction. *Online Learning*, 28(3), 181–205. https://doi.org/10.24059/olj.v28i3.4458

Topali, P., Haelermans, C., Molenaar, I., & Segers, E. (2025). Pedagogical considerations in the automation era: A systematic literature review of AIEd in K-12 authentic settings. *British Educational Research Journal*, 00, 1–33. https://doi.org/10.1002/berj.4200

Wang, X., Huang, R., Sommer, M., Pei, B., Shidfar, P., Rehman, M. S., Ritzhaupt, A. D., & Martin, F. (2024). The efficacy of artificial intelligence-enabled adaptive learning systems from 2010 to 2022 on learner outcomes: A meta-analysis. *Journal of Educational Computing Research*, 62(6), 1348–1383. https://doi.org/10.1177/07356331241240459

Wu, T.-T., Lee, H.-Y., Li, P.-H., Huang, C.-N., & Huang, Y.-M. (2023). Promoting self-regulation progress and knowledge construction in blended learning via ChatGPT-based learning aid. *Journal of Educational Computing Research*, 61(8), 1539–1567. https://doi.org/10.1177/07356331231191125

Zhao, X., Cox, A., & Chen, X. (2025). The use of generative AI by students with disabilities in higher education. *The Internet and Higher Education*, 66, 101014. https://doi.org/10.1016/j.iheduc.2025.101014