The Influence of Artificial Intelligence on Science Education

Anu Lata
Assistant Professor, Department of Education
Dasmesh Girls College, Chak Alla Baksh, Mukerian-144211
doi.org/10.64643/IJIRTV12I6-191332-459

***Abstract*—The integration of artificial intelligence (AI) in science education is reshaping teaching practices, assessment, curriculum design, and learners' epistemic practices. This paper synthesizes recent empirical and review literature (2013-2025) to answer: How does AI influence science learning outcomes, pedagogical practices, and assessment; what challenges and equity risks exist; and what design principles should guide responsible AI integration in science education? We find consistent evidence that AI-driven systems-intelligent tutoring systems (ITS), adaptive learning platforms, educational robotics, and large language models (LLMs) can improve formative feedback, scaffold inquiry, and personalize learning pathways, particularly in STEM domains. However, effects vary by context, teacher expertise, data quality, and assessment alignment. Ethical concerns (bias, transparency, data privacy) and teacher preparation remain primary barriers. We propose a socio-technical framework for integrating AI into science classrooms and outline mixed-methods designs to evaluate pedagogical impact, equity outcomes, and teacher change. Key recommendations emphasize teacher agency, curriculum alignment to scientific practices, transparency, and iterative evaluation.**

***Index Terms*—AI in Education, Intelligent Tutoring Systems, Large Language Models, Formative Assessment, Equity, Teacher Professional Development**

I. INTRODUCTION

Recent advances in machine learning, natural language processing, and educational data mining have produced tools with the potential to transform science education. AI can provide individualized feedback, model complex simulations, scaffold inquiry-based labs, and support formative assessment at scale. At the same time, these technologies raise concerns about fairness, transparency, and the nature of scientific learning when part of the reasoning process is externalized to algorithms. This paper reviews the empirical and theoretical literature to synthesize what is known about AI's influence on science teaching and learning and to propose principled directions for research and practice (1).

II. LITERATURE REVIEW

(I) HISTORICAL ROOTS: ITS AND ADAPTIVE SYSTEMS

Intelligent Tutoring Systems (ITS) and adaptive instructional systems have decades of evidence showing moderate improvements in STEM learning outcomes, especially when providing stepwise feedback and individualized pacing. Meta-analyses indicate ITS have positive effects compared with traditional classroom instruction, though often smaller than one-on-one human tutoring; effect sizes vary by subject, fidelity of domain model, and deployment conditions. ([ida.org][2])

(II) EDUCATIONAL ROBOTICS AND HANDS ON SCIENCE LEARNING

Educational robots and sensor-enabled labs bring AI to embodied science learning-helping students test hypotheses, collect real-time data, and engage in iterative experimentation. Reviews indicate robotics can support engagement and conceptual understanding in K-12 science, but teacher scaffolding and curricular integration are essential.

(III) LARGE LANGUAGE MODELS (LLMS) AND GENERATIVE AI

The rapid spread of LLMs (e.g., GPT-family, others) has created new affordances for drafting explanations, generating formative questions, simulating "what-if" scenarios, and offering conversational tutoring.

Surveys and recent syntheses show LLMs can enhance accessibility and creativity but raise risks: hallucinations (incorrect but plausible outputs), overreliance by learners, and propagation of biases present in training data. Early studies suggest LLMs can be powerful assistants when paired with scaffolds and verification practices.

(IV) ASSESSMENT, FEEDBACK, AND LEARNING ANALYTICS

AI enables new forms of formative assessment (automated scoring, instant hints, and model-based feedback) and learning analytics that reveal students' problem-solving paths. The promise is rapid, personalized feedback aligned with scientific practices (modeling, explanation, evidence-based reasoning). However, validity and alignment with curricular goals are ongoing concerns (4).

(V) EQUITY, ETHICS AND TEACHER PROFESSIONAL LEARNING

Multiple reviews emphasize equity risks (biased outputs, uneven access to infrastructure), the need for teacher professional development, and policy for privacy and accountability. Successful implementations foreground teacher agency-AI as augmentation, not replacement-and explicit instruction in AI literacy within science curricula.

III. CONCEPTUAL FRAMEWORK: A SOCIO-TECHNICAL MODEL FOR AI IN SCIENCE EDUCATION

I propose the S-TAR model (Socio-technical Alignment for Responsible AI), which centers four interacting components:

(i) Student cognition and epistemic practices-how AI supports hypothesis formation, modeling, argumentation from evidence, and met cognition.
(ii) Teacher mediation and pedagogy-teacher roles in scaffolding, interpreting AI feedback, and maintaining epistemic norms.
(iii) Technology affordances and limitations-accuracy, explain ability, latency, domain coverage (e.g., physics vs. ecology), and propensity to hallucinate.

IV. CONTEXTUAL ETHICS AND GOVERNANCE data privacy, equity of access, algorithmic bias, assessment validity, and institutional policy.

The framework guides design (align tool affordances with target scientific practices), implementation (teacher PD, classroom workflows), and evaluation (learning outcomes + equity metrics + teacher outcomes).

V. METHODOLOGICAL RECOMMENDATIONS FOR EMPIRICAL STUDY

To robustly evaluate AI's influence on science education, combine these methods:

Cluster randomized controlled trials (RCTs) where feasible (e.g., comparing curriculum + AI tutor vs. curriculum alone) for causal estimates of learning gains. Include pre/post content tests and transfer tasks.

MIXED-METHODS STUDIES: classroom observations, teacher interviews, and analysis of interaction logs to surface mechanisms (how feedback was used, when students verified AI outputs).

LEARNING ANALYTICS AND SEQUENCE MINING: analyze fine-grained traces (click streams, solution steps) to model how AI feedback changes problem-solving strategies.

EQUITY-FOCUSED ANALYSES: disaggregate outcomes by socio-economic status, language background, and prior achievement to detect differential effects.

VALIDITY STUDIES FOR AUTOMATED ASSESSMENT: align automated scoring rubrics with expert human scoring and analyze reliability across diverse student responses.

EXAMPLE MEASUREMENT INSTRUMENTS: concept inventories in target science domains, rubrics for scientific argumentation, surveys for epistemic beliefs and AI trust.

VI. SYNTHESIS OF EMPIRICAL FINDINGS (WHAT THE LITERATURE SHOWS)

(I). LEARNING GAINS: ITS and adaptive platforms produce small-to-moderate average gains in STEM subjects; LLM-augmented tutors show promise but evidence is still emerging. Gains are largest when AI provides timely, targeted feedback and is tightly aligned with curriculum objectives (2).

(II). ENGAGEMENT AND MOTIVATION: Educational robotics and interactive simulations increase engagement and task persistence-especially in younger learners-when combined with teacher-guided inquiry.

(III). TEACHER PRACTICE: Teachers who receive PD on AI tools and maintain control over assessment choices tend to use AI more effectively; otherwise, tools can be marginalized or used superficially (4).

(IV). EQUITY RISKS AND ACCESS: Without policy and infrastructure investment, AI can widen gaps-unequal device access, biases in language/knowledge representation, and differential teacher capacity.

(V). ASSESSMENT VALIDITY AND HALLUCINATION RISK: Generative models can produce convincing but incorrect science explanations (hallucinations). Systems that combine LLMs with verification modules or retrieval-augmented approaches show more promise for reliable educational use (3).

(vi). Practical guidelines for teachers, designers, and policymakers

FOR TEACHERS

- Treat AI as an assistant-use outputs as starting points, and teach students verification strategies (cross-checking, experimental validation).
- Integrate AI-supported tasks that require scientific practices (e.g., designing experiments, constructing explanations) rather than only fact retrieval.
- Seek PD that includes hands-on use of ed-AI tools and pedagogical scenarios.

FOR DESIGNERS

- Prioritize explain ability: provide traceable reasoning or sources for claims, and expose uncertainty.
- Align feedback with curricular rubrics (not only correctness).
- Build retrieval-augmented systems to reduce hallucinations and to tie outputs to authoritative resources.

FOR POLICYMAKERS

- Invest in equitable access (devices, internet) and in teacher preparation.
- Require transparency and data-privacy standards for ed-AI vendors.
- Support independent evaluation of AI tools in real classrooms before large-scale adoption (4).

VII. LIMITATIONS OF EXISTING RESEARCH AND OPEN QUESTIONS

(I) HETEROGENEITY OF EFFECTS: Many studies report varied effect sizes; more replication across contexts (countries, grade levels, subjects) is needed.

(II) SHORT-TERM FOCUS: Much evidence addresses immediate learning gains; long-term retention, transfer, and changes to students' scientific reasoning need study.

(III) LLM RELIABILITY: Rapidly evolving models mean prior findings may age quickly; ongoing evaluation is necessary.

(IV) ETHICAL MEASUREMENT: How to measure and mitigate algorithmic bias in formative feedback remains underdeveloped.

VIII. CONCLUSION

AI offers substantial opportunities to enhance science education—improving personalized feedback, enabling richer simulations, and scaling formative assessment—when designed and implemented with attention to pedagogy, transparency, and equity. The most promising approach positions AI as a socio-technical augmentation: powerful tools under teacher mediation, aligned to scientific practices, and governed by clear ethical and policy safeguards. Future research should prioritize rigorous causal studies, equity analyses, and the development of AIaware curricula that teach students to use and critique AI as part of scientific inquiry.

REFERENCES

[1] The impact of artificial intelligence on scientific practices: an emergent area of research for science education—Oxford/Education report (review of AI in science education). ([education.ox.ac.uk][1])
[2] The impact of large language models on higher education — Frontiers in Education (2024): overview of LLM opportunities and risks. ([Frontiers][6])
[3] Large Language Models for Education: A Survey and Outlook — arXiv (2024): technical/educational survey of LLMs. ([arXiv][3])
[4] A systematic review of AI-driven intelligent tutoring systems in K-12 — PMC/NCBI (2025): meta-review of ITS impacts and classroom designs. ([pmc.ncbi.nlm.nih.gov][5])
[5] Artificial Intelligence and the Future of Teaching and Learning— US Dept. of Education report (2023): policy and formative assessment guidance. ([U.S. Department of Education][4]
[6] https://arxiv.org/pdf/2403.18105?utm_source=chatgpt.com "[PDF] Large Language Models for Education: A Survey and Outlook - arXiv"
[7] https://pmc.ncbi.nlm.nih.gov/articles/PMC12078640/?utm_source=chatgpt.com"A systematic review of AI-driven intelligent tutoring systems (ITS) in."
[8] 8.https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2024.1392091/full?utm_source=chatgpt.com "The impact of large language models on higher education - Frontiers"