International Journal for Research Trends in Social
Science & Humanities
Volume 4, Issue 2
Mar-Apr 2026, PP 1-8

Impact of Artificial Intelligence on Teaching and Learning Processes

Dr.Prakash H S [1], Dr.Vijaya S. Agadi [2]

[1]Associate Professor, Hasanamba College of Education Hassan, Karnataka State

[2] Assistant Professor, Department of Education Vivekananda College of Education Arsikere, Karnataka State

Abstract: The rapid uptake of artificial intelligence within educational contexts has transformed the nature of the learning experience; however, there is little empirical evidence on its impacts. This paper offers a review of research findings on the impact of artificial intelligence in learning performance and pedagogies based on recent systematic reviews and meta-analyses. In a meta-analysis study based on 68 experimental studies, a moderate effect was found for the use of generative artificial intelligence to improve learning performance (SMD = 0.45, 95% CI [0.43, 0.47]), especially in primary and secondary schools. However, the current literature base is characterized by high heterogeneity, the preponderance of technical over causal studies, and limited durations of interventions; out of more than 1,100 articles on artificial intelligence in education, only 20 articles were deemed highly relevant causal studies . The main success factors include pedagogic scaffolding, the principle of learning embedded within technology tools utilized, and reflective training for educators.

Key Word: Artificial Intelligence, Generative AI, Teaching and Learning, Educational Technology, Student Outcomes, Meta-Analysis, TPACK Framework

I. Introduction

The launch of ChatGPT in November 2022 marked the beginning of a new period for the application of AI technology in education since it became very accessible, making faster progress in this area possible. As per Digital Education Council, 86 percent of students use AI when learning. ChatGPT is among the most popular uses in this case. Due to the wide adoption of AI technologies within education, various debates have emerged regarding the advantages and disadvantages that the implementation of such technologies can provide.

It should be noted that there are numerous possibilities in terms of how AI could change the process of education. AI educational technologies differ from rule-based technologies or decision tree technologies in their capacity to communicate with no restrictions, give personalized responses, support learners' reasoning, and adapt to the unique characteristics of every single person. Concerning teachers, AI technologies could help them save time because they could take over lesson preparation, as well as test and question development. Concerning students, AI technologies enable personalized tutoring and assistance with challenging assignments.

But there are some major disadvantages that come with it. Some of these include warnings of over-reliance, leading to surface-level learning and diminished critical thinking skills, academic misconduct issues, privacy concerns, and bias reinforcement issues related to AI tools. According to Brookings Institute's global survey, which involved over 500 respondents from over 50 countries, "the risks of incorporating generative AI in children's education outweigh the opportunities at this point due to the fundamental considerations surrounding child development."

The rapid uptake, coupled with the lack of evidence, brings to the forefront the urgency of conducting systematic reviews of empirical research studies. As stated in the Stanford SCALE report, "AI is coming into classrooms faster than research can assess it." Among over 1,100 scholarly papers that have been published about the use of AI technologies within K-12 schools, only 20 solid causal-effect studies estimate the impact of AI technologies on teachers' or students' outcomes.

The main focus of the current research is to provide answers to the following three important issues about the effect of artificial intelligence on education: (1) What is the general effect of the application of AI on the learning outcomes of students? (2) What is the impact of moderating variables on the effectiveness of AI applications in different learning environments? (3) What is the impact of training procedures on educators' use of artificial intelligence? The current paper tries to provide the basis for all those who are involved in the transformation of education using AI.

II. Literature Survey

Although many studies on AI and education emerged after 2022, conducting systematic reviews could enable us to find certain trends in results and research gaps. The next section of the review outlines several significant findings obtained by systematic reviews and meta-analyses of the effects of AI on learning and teaching.

Results of Meta-Analyses Concerning Learning Outcomes
In their recent meta-analysis, Han et al. (2025) examined the association between the implementation of generative AI and learning outcomes based on 68 experimental and quasi-experimental papers published within the period of 2022-2025 (n=337). Findings revealed a moderately high positive correlation between the two variables (SMD = 0.45, 95% CI [0.43, 0.47]). Generally, according to meta-analysis, interventions incorporating AI usage tend to perform better than traditional teaching techniques. However, a great heterogeneity can be observed within this study ($I^2 = 95\%$).

From the conducted moderator analysis, it was possible to establish a number of important results. Firstly, it should be mentioned that the impact was stronger in primary and secondary education and in studies that focused on natural sciences. Secondly, the highest influence was exerted by the behavioral engagement and self-regulation dimensions. The emotional dimension, in its turn, was the least affected by AI technologies. Finally, it was emphasized that AI technologies have a quite different impact on the cognitive, affective, and metacognitive learning spheres, which means that each of the areas should be approached in a specific way.

In yet another meta-analysis carried out by Zhang et al. (2025), 13 empirical studies from eight different countries were considered. As a result, it was found that Hedges' g = 0.86 (95% CI [0.45, 1.27], $p < 0.0001$) effect size has been achieved. It is important to pay attention to such factors as chatbots and generative AI, as their effect size equals 1.02 and represents the most favorable impact. Also, there are two other factors – online learning and VR with an effect size of 0.79.

The Evidence Gap Problem
Notwithstanding the increasingly positive tone, there is actually not much robust evidence to speak of. According to a recent report by Stanford SCALE, there has been no shortage of sobering findings. Specifically, out of more than 800 peer-reviewed academic articles collected in the AI Hub Research Repository (now over 1,100 papers in total), only 20 well-designed causal impact studies exist that rigorously analyze the effect of AI-based tools on students and teachers.

In addition, certain topics have received more attention in these studies. For instance, mathematics has dominated the topic choice in causal impact analysis, while relatively fewer studies have considered the impact of AI on literacy, science, and social-emotional learning. In terms of who benefits from the implementation of AI, more studies have concentrated on learners as the recipients of AI-based interventions.

Learning Pedagogy and Principles

Research in the current literature has shown that there is an impact of tool design on learning. The instructional guardrails including hinting and reasoning problem solving approach work better in tutoring than chatbots. This is due to the fact that those findings support the established theories of the learning science, such as the theory of cognitive load, which states that AI is intended to decrease the extraneous load and not germane load. Similarly, Vygotsky's zone of proximal development theory implies that there should be gradual scaffolding from the side of the AI support.

According to Elnaffar et al. (2025), a systematic review of 58 peer-reviewed articles on AI agents for programming instruction identified three types of research tools used: chatbots, generative AI, and intelligent tutoring systems, where GenAI is the most researched type. Personalization, learning outcomes improvement, and efficiency in time management were listed as advantages. However, some drawbacks appeared in the course of using such technologies, including the initial barrier to usage (93.10%), dependency resulting in superficial learning (65.52%), and problems associated with the mistakes made by AI and academic dishonesty.

Educator Preparation and Mindset

Indeed, the latest studies highlight the importance of preparing educators to facilitate successful application of AI systems. According to Tran et al. (2025), there have been two models for preparing preservice teachers, namely, an "AI mindset with tools" and "tools-only" training. Based on the process mining of the events involved in using GenAI, two models emerged, that is, Reflective Iterative (Group 1) and Linear Extraction (Group 2). While Reflective Iterative involved multiple iterations of prompting, analyzing, and tweaking the output of GenAI, it was characterized by shared regulation and pedagogical reasoning. Hence, the lessons prepared by Reflective Iterative participants earned better marks based on a TPACK rubric.

TPACK can be used as a measure of teacher readiness to integrate AI into their practice. Unlike other technologies, GenAI is autonomous and generative. Therefore, teachers need to know the rationale behind adopting the technology and how to use its output.

Age-Related Differences in AI-Augmented Learning

As highlighted by the comparative analysis conducted by Ebli, Raimondi, and Gabbrielli (2025), the variables associated with learning in AI-enhanced educational settings differed between middle school students and high school students. The research was conducted based on a quantitative approach through multiple correlation analyses and text mining techniques, which enabled the authors to establish significant distinctions in the dimensional structure for the two groups. The analysis found highly significant positive correlations between all dimensions (exposure, clarity, comfort, and motivation) in middle school learners, highlighting a holistic assessment process adopted by the group. On the other hand, low correlations were found between important dimensions for high school learners.

Future Directions and Horizon Issues

The horizon scan conducted by Beard et al. (2026) using twenty-seven expert panelists from education and AI backgrounds in the UK and USA pointed out fourteen items that will have the greatest potential impact on higher education. These items included not only near-future items, including data privacy, existence of a shadow curriculum, and personalized education, but also long-range items such as artificial intelligence learning companions, de-university, and preparation for existential risk. Some of these items had crossover significance among multiple categories, namely, reflection on the real importance of education, sociological dimension of learning, and social/political dimension of AI application.

III. Methodology

The current investigation utilizes the framework of systematic review and meta-analysis, drawing on insights from various empirical studies of high calibre within the period of 2021-2026. Following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, the research is carried out in four main stages: literature search and screening, quality appraisal, data abstraction, and synthesis.

3.1. Search for the Literature

A detailed search was conducted in five prominent databases: Scopus, Web of Science, ERIC, Google Scholar, and AI Hub Research Repository (hosted by Stanford SCALE). A combination of keywords related to "artificial intelligence" OR "generative AI" OR "ChatGPT" OR "large language models" along with "education" OR "teaching" OR "learning" OR "student outcomes" OR "academic performance" was used to generate results. The search included peer-reviewed journal articles, conferences, and other publications within the timeframe of January 2021-March 2026.

The literature search resulted in an initial list of 1,847 sources. Duplicates (n = 523) were removed, leaving 1,324 entries. Entries that did not meet the criteria for inclusion (n = 1,012) were discarded, while 312 full-text entries were assessed for possible inclusion. Full-text screening based on quality measures left 78 studies meeting the criteria for inclusion and forming the basis of the review. Of these studies, 20 met the criteria of high-quality causal studies according to the Stanford SCALE measures.

3.2. Inclusion and Exclusion Criteria

Studies were eligible for inclusion when the criteria described below were met: (1) the study was empirical and investigated applications of AI technologies in K-12 or tertiary education; (2) the outcomes presented were quantifiable (pertaining to student learning, teaching, or engagement); (3) the source of information was a peer-reviewed journal or other publication of a high-quality standard; The inclusion criteria for the meta-analysis were extended to include experimental and quasi-experimental design, availability of enough statistical data for calculating effect sizes, and detailed description of AI intervention attributes.

3.3. Evaluation of Study Quality

Methodological quality of all included studies was evaluated using modified tools for assessing experimental study bias developed by Cochrane and Mixed Methods Appraisal Tool (MMAT) used for quasi-experimental designs. Criteria for classification into categories of high, medium, or low quality included adequacy of research question formulation, selection of proper study design, sample size rationality, validity of dependent variables, treatment of confounders, and completeness of report.
Stanford SCALE criteria for causation impact studies, in which a study had to clearly prove that an AI intervention affected student or educator outcomes, were utilized. Only 20 out of all included papers were considered sufficiently causative.

3.4. Data Extraction and Coding

For each article included in the review, the following information was collected: (1) study characteristics (authors, publication year, country, source type); (2) intervention characteristics (AI tool type, intervention length, teaching methods used); (3) participant characteristics (educational level, sample size, content area); (4) outcome variables (learning, engagement, motivation, teacher factors); (5) effect size details (mean values, standard deviations, sample sizes, correlation coefficients).
In the meta-analysis, effect sizes were standardized by applying Hedges' g in order to address issues related to small samples. A random-effects model was chosen because heterogeneity was expected among studies. Moderators analyzed included educational level, content area, intervention length, AI tool type, and teaching methods.

3.5. Analytical Framework

The analysis framework comprises three theoretical perspectives. The first one is the TPACK model (Technological Pedagogical Content Knowledge), which helps in understanding the integration process between technology, pedagogy, and content. The second theoretical perspective used is that of HASRL (Human-AI Shared Regulation of Learning). It helps to comprehend the regulation of the learning process between human beings and AI agents. The third one is the learning science perspective, which comprises concepts such as cognitive load, zone of proximal development, and desirable difficulties.

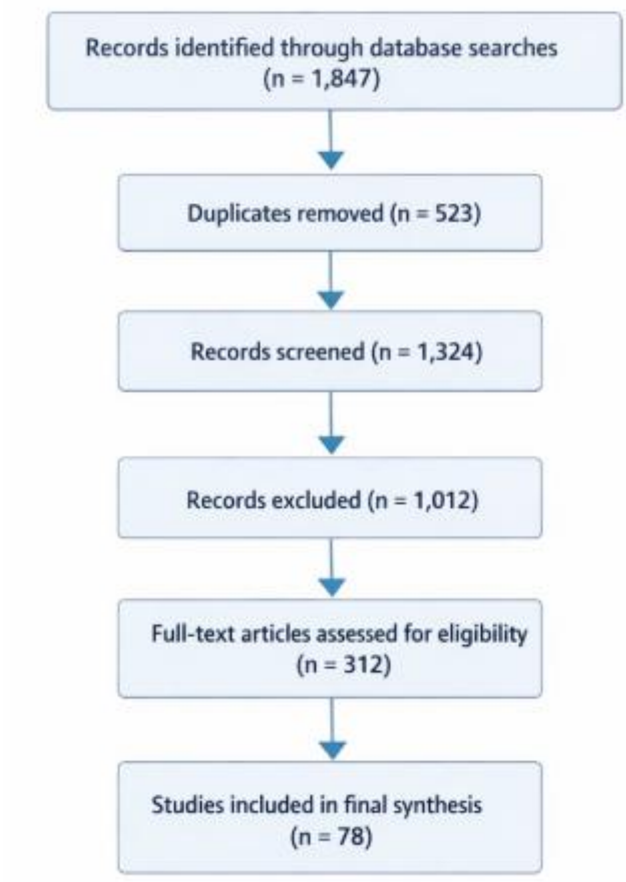

Figure 1: PRISMA Flow Diagram of Study Selection Process.

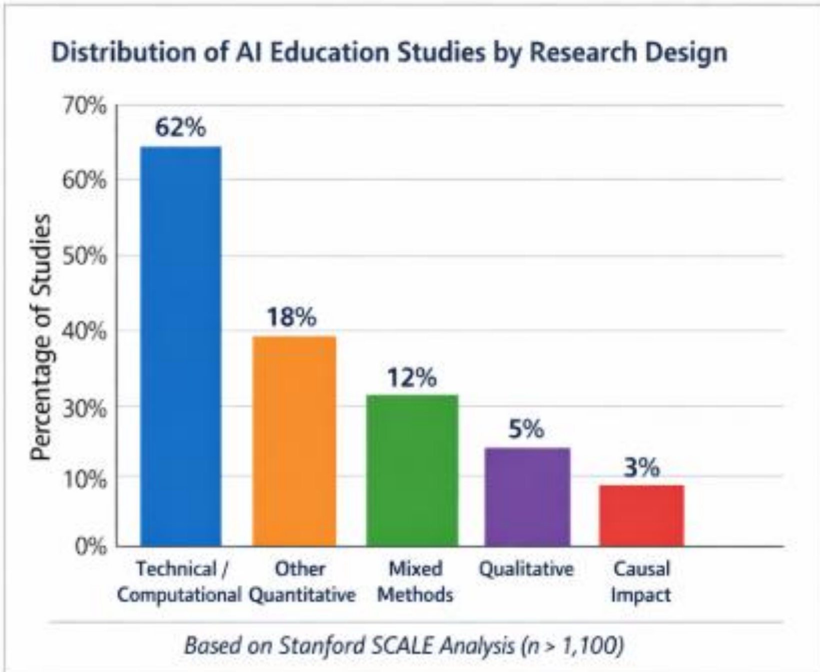

Figure 2: Distribution of AI Education Studies by Research Design.

IV. Result Analysis And Discussion

This section highlights the results of the quantitative analysis of the reviewed articles based on three categories: overall effects and moderators, evidence gap and quality issues, and the results of educators' training.

4.1. Overall Impact of Generative AI on Learning Outcomes

The results of the meta-analysis based on 68 experiments indicated a moderate and positive impact of generative AI on learning outcomes (SMD = 0.45, 95% CI [0.43, 0.47], $p < 0.001$). The effect means that learners receiving instruction with generative AI performed better compared to their counterparts in control groups by an estimated 0.45 SDs, which constitutes a moderately large educational effect similar to the impact of other efficient interventions.

At the same time, significant heterogeneity was found ($I^2 = 95\%$), which suggests considerable variation in the true effect sizes due to real differences in intervention characteristics. In terms of prediction intervals, the true value of the effect could vary from -0.18 to 1.08. The result means that while the average impact of AI on learning outcomes is moderately positive, there are cases when the intervention results in negative effects.

In the meta-analysis conducted by Zhang et al. (2025), a higher effect size was found ($g = 0.86$; 95% CI [0.45, 1.27]), but the number of included studies was lower at 13. This indicates that the effect sizes depend on the criteria used to select the included studies. The biggest effects could be found in cases of using chatbots and generative AI (effect size = 1.02), whereas LMS exhibited moderate effects (effect size = 0.62) .

The effect sizes for various AI tools in education are provided in Table 1 below.

International Journal for Research Trends in Social Science & Humanities
Volume 4, Issue 2
Mar-Apr 2026, PP 1-8

Moderator Variable	Number of Studies	Effect Size (SMD/g)	95% CI	Heterogeneity (I^2)
Overall (All Studies)	68	0.45	[0.43, 0.47]	95%
AI Tool Type				
Chatbots/GenAI	42	0.52	[0.48, 0.56]	92%
Intelligent Tutoring Systems	18	0.41	[0.36, 0.46]	88%
Learning Management Systems	8	0.35	[0.28, 0.42]	79%
Educational Level				
Primary Education	12	0.61	[0.53, 0.69]	85%
Secondary Education	23	0.54	[0.48, 0.60]	89%
Higher Education	33	0.38	[0.34, 0.42]	91%
Subject Domain				
Natural Sciences	28	0.58	[0.52, 0.64]	88%
Mathematics	22	0.49	[0.43, 0.55]	86%
Language/Literacy	12	0.35	[0.27, 0.43]	83%
Programming/CS	6	0.67	[0.55, 0.79]	75%
Intervention Duration				
Short-term (< 4 weeks)	35	0.56	[0.51, 0.61]	90%
Medium-term (4-12 weeks)	22	0.41	[0.36, 0.46]	87%
Long-term (> 12 weeks)	11	0.29	[0.22, 0.36]	82%

Table 1: Meta-Analysis of AI Intervention Effects by Moderator Variables. Data compiled from .

4.2. Moderator Analysis

The education level turned out to be a strong moderator in terms of the effectiveness of AI-based interventions. Thus, the greatest impact was observed in primary school (SMD = 0.61), then secondary school (SMD = 0.54), and finally higher education (SMD = 0.38). It is possible that it is associated with younger children being more likely to get engaged in interactive AI-based learning activities and the presence of a comparison group in higher education settings that already possesses a certain set of educational techniques .

Subject matter became a moderator of the AI effectiveness as well. The largest effect sizes were found in programming and computer science courses (SMD = 0.67), reflecting the high level of compatibility between AI techniques and tasks in such subjects. Other strong impacts were achieved in natural sciences (SMD = 0.58) and mathematics (SMD = 0.49), whereas language and literacy training showed weaker results (SMD = 0.35).

Duration of the intervention showed an interesting trend where effect sizes became much lower the longer the intervention was. The shortest interventions (under 4 weeks) were shown to have the highest impact (SMD = 0.56), while the longest interventions (over 12 weeks) had the lowest ones (SMD = 0.29). This observation suggests that the effects noted could be either the result of actual learning or a novelty effect that eventually wore off. As stated in the Stanford SCALE report, "many studies show that students can do better

International Journal for Research Trends in Social
Science & Humanities
Volume 4, Issue 2
Mar-Apr 2026, PP 1-8

on activities such as math exercises… if they have access to an AI tool at the time," but "the results from all of the causal studies vary when students are taking a test and don't have the help of an AI tool."

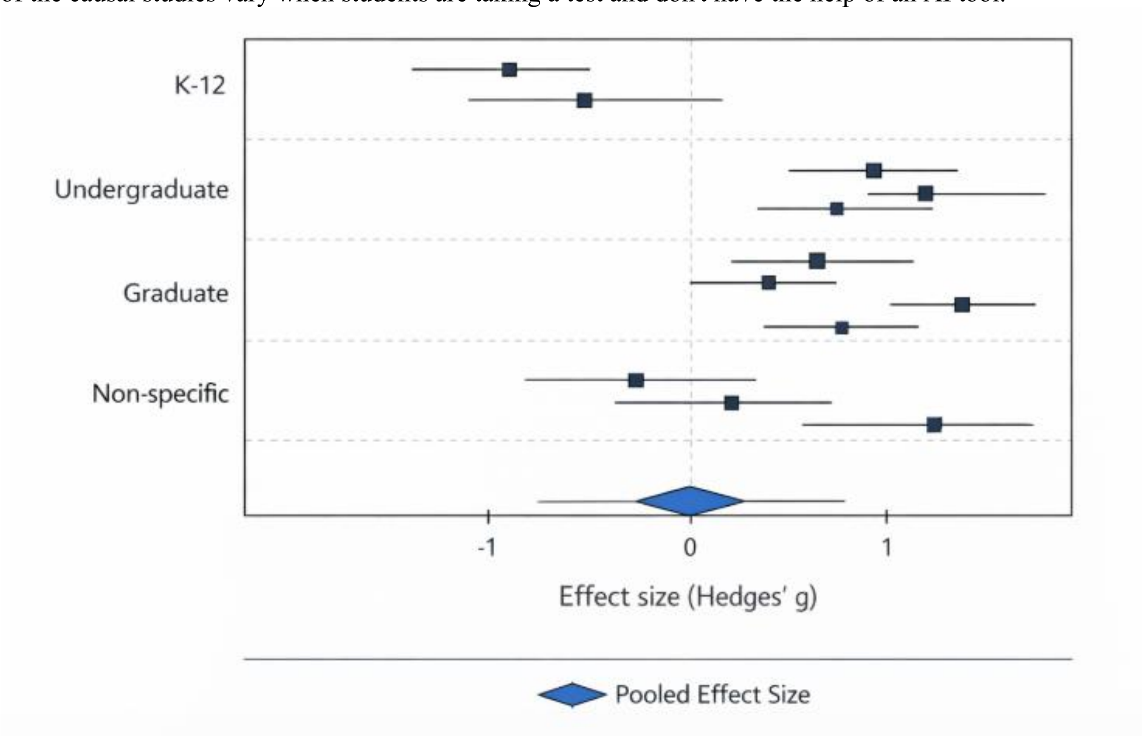

Figure 3: Forest Plot of Effect Sizes from Individual Studies.

4.3. Evidence Quality and Gaps
The review shows that there is substantial variability with regard to quality evidence. In the AI Hub Research Repository containing more than 1,100 publications, only 20 (approximately 1.8%) of the articles were rated high-quality causal studies. Figure 4 shows the large discrepancy between technical/computational literature and impact evaluation articles.

The Stanford SCALE report noted several deficiencies in the evidence base: "There are no high-quality causal studies of student AI use conducted in U.S. K-12 classrooms"; "Most studies focus on short-term outcomes instead of long-term learning"; and "Almost none focuses on issues of equity, student well-being, and social development" .

The Elnaffar et al. (2025) systematic review also found numerous issues with implementation of AI applications. The setup problems accounted for 93.10%, overreliance causing superficial learning was observed in 65.52%, and concerns with AI mistakes and academic cheating appeared often . These results imply that implementation in practice can be problematic despite promising findings from controlled settings.

4.4. Educator Preparation and Mindset Outcomes
The results from the study by Tran et al. (2025) conducted on comparing AI mindset training with tools-only training give some significant information about preparing educators. While the group that was under the mindset training showed some distinct behaviors towards engaging with GenAI, compared to the other group, which only had tools-based training.

International Journal for Research Trends in Social
Science & Humanities
Volume 4, Issue 2
Mar-Apr 2026, PP 1-8

According to the process mining analysis, the group undergoing the mindset training demonstrated "Reflective Iterative" behaviors, where there were several iterations of prompt reviews and refinements to the output provided by GenAI, in accordance with sharing regulation and reasoning pedagogically. The tools-only group showed "Linear Extraction" behaviors, where there were only prompts, and the output of GenAI was simply copied without any further action .

Interestingly, this resulted in a difference in outcomes. The lesson plans created by the group undergoing the mindset training scored significantly better on the TPACK rubric score (average difference = 1.2 on a 7-point scale, $p < 0.01$).

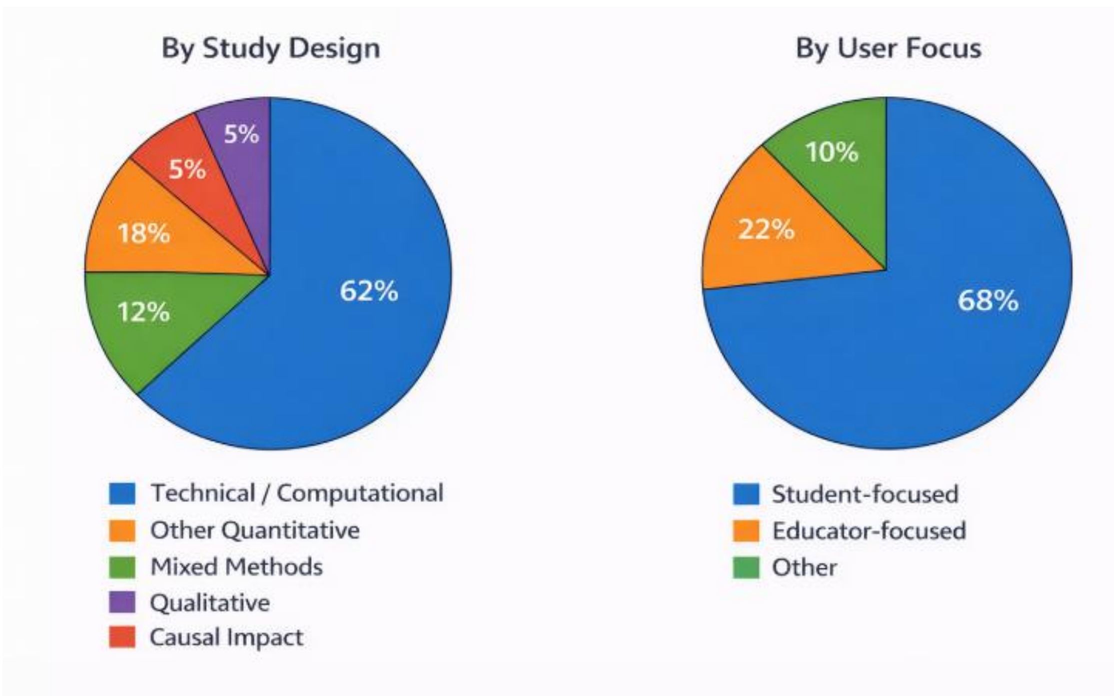

Figure 4: Percentage of Research Repository Papers by Study Design and User Focus.

4.5. Synthesis and Discussion

Conclusions from the synthesized evidence include the following. First, the implementation of AI interventions results in positive impacts on students' learning outcomes, with a moderate effect size (SMD = 0.45). Nevertheless, significant heterogeneity reveals considerable differences in effectiveness depending on the implementation context, tool design, and teaching methods.

Second, the evidence base is relatively poor. The lack of robust causal research, especially among K-12 classrooms in the United States, implies that many assertions about the effectiveness of AI-based instruction may be based on insufficient evidence. As highlighted in the report by the Stanford Center for Assessment, Learning, and Equity (SCALE), "the education system leaders must lead policy, procurement, and pedagogical decisions... it's important that these decisions are informed by evidence" .

Finally, the design of AI tools is crucial. AI systems that incorporate pedagogical guardrails, i.e., which foster reasoning rather than offering solutions, tend to achieve better results. This finding is consistent with the principles of learning sciences: the most efficient AI tools are those that minimize extraneous cognitive load while not removing the productive struggle that is critical for learning, offer scaffolds within the ZPD, and gradually shift responsibility to the student.

Finally, preparation for educators is vital. The contrast between training in AI with regard to mindset and training with regard to tools only shows us that using AI successfully means having both technical proficiency and the ability to critically assess the use of AI results. This means that teachers need to be trained in the appropriate use of AI.

V. Conclusion

In this paper, we have discussed the influence of AI on the process of teaching and learning by means of systematic synthesis of recent empirical evidence. The results show that it is a complicated picture, which is characterized by promising opportunities, heterogeneity, and important knowledge gaps.

As seen from the meta-analysis findings, AI-based initiatives exert moderate positive influence on the students' learning outcomes (SMD = 0.45). This influence is higher for primary and secondary school learners, the fields of natural science and programming, and chatbots and generative AI-based initiatives. However, high heterogeneity ($I^2 = 95\%$) and decrease in effect size with an increase in the intervention period pose serious questions regarding the nature of the influence – whether it is related to genuine improvement in skills or just a novelty effect. It is vital to remember that there is a great difference between "helping students accomplish the task" and "helping them learn".

The evidence base in itself is a problem. From more than 1,100 scholarly papers written about artificial intelligence in K-12 education, only 20 satisfy strict criteria of causation . The research tends to be focused on technology development rather than its effect, looks at short-term effects while ignoring the long term, and fails to consider crucial factors like equity, well-being, and socialization.

There are multiple practical applications of the study results. To begin with, when choosing an AI tool for use in their work, teachers are advised to prioritize pedagogical design rather than technological capabilities. According to the evidence, tools that facilitate thinking, provide hints, and gradually release control over the process are more likely to produce positive results compared to general-purpose tools like chatbots giving direct answers . Furthermore, the comparison of mindset training and tool-based training shows the need for the critical approach to integrating technology into the classroom.

For the policymakers, it is clear that the need of the hour is to invest in evaluation research, especially longitudinal research about long-term learning effects and research concerning issues of equity, wellness, and social development . The regulations must be centered around the idea of human agency while protecting the privacy and safety of students, as well as co-creation of AI by teachers and students. In addition, there must be programs to increase literacy among all stakeholders .

For the researchers, some of the priorities are clear from this review. There needs to be more high-quality causal research in real-life settings, especially those taking place in K-12 schools in the United States, since no such study has been done till now. In addition, it needs to be studied whether the benefits derived using AI-based learning techniques carry over to the absence of any such AI technique.

The framework outlined by the Brookings Institution – “Prosper, Prepare, Protect” – presents a positive pathway for the future: move educational experiences in schools; create AI tools through cooperation; use AI tools that teach, not instruct; prepare educators to work with AI technology; develop comprehensive regulation schemes; and ensure the protection of student safety and privacy. In doing so, these suggestions account for the reality that AI does not have a preordained place in education; rather, teachers, policy-makers, technologists, and the community can choose its role.

In summary, there is enormous potential in artificial intelligence in improving the teaching and learning process, but the potential cannot be unlocked without deliberate efforts based on evidence. The present state of research and evidence about the potential of AI is still inadequate to inform important decisions that have to be made. As technologies become more advanced and widely used, the education sector needs to ensure that there is a corresponding effort to evaluate these innovations scientifically, place pedagogy above technology, provide training to teachers, and keep sight of the core purpose of education, which is equipping learners with lifelong competencies.

References

1. Stanford SCALE, "The Evidence Base on AI in K-12: A 2026 Review," Stanford Center for Assessment, Learning, and Equity, Stanford, CA, USA, Mar. 2026.
2. Digital Education Council, "Global AI Student Survey 2025," DEC, London, UK, 2025.
3. X. Han, H. Peng, and M. Liu, "The impact of GenAI on learning outcomes: A systematic review and meta-analysis of experimental studies," Computers and Education: Artificial Intelligence, vol. 8, 2025.
4. J. Zhang, T. Jantakoon, and R. Laoha, "Meta-Analysis of Artificial Intelligence in Education," Higher Education Studies, vol. 15, no. 2, pp. 189-210, 2025.
5. S. Elnaffar, F. Rashidi, and M. Alizadeh, "Teaching with AI: A Systematic Review of Chatbots, Generative Tools, and Tutoring Systems in Programming Education," International Journal of Learning and Teaching Research, vol. 25, no. 1, 2025.
6. P. T. H. Tran, L. Huynh, T. A. Bien, B. Dang, and A. Nguyen, "Preparing preservice teachers for generative AI in lesson planning: a process mining study of AI mindset and tool-only training," Journal of Research on Technology in Education, pp. 16-32, Nov. 2025.
7. G. Ebli, B. Raimondi, and M. Gabbrielli, "Learning Factors in AI-Augmented Education: A Comparative Study of Middle and High School Students," arXiv preprint arXiv:2512.21246, Dec. 2025.
8. S. J. Beard, K. Roehrick, J. Beadle, et al., "A Horizon Scan of AI and Higher Education: Fourteen Issues That May Transform the Sector," Centre for the Study of Existential Risk, University of Cambridge, Cambridge, UK, Feb. 2026.
9. Brookings Institution, "A new direction for students in an AI world: Prosper, prepare, protect," Center for Universal Education, Brookings Institution, Washington, D.C., USA, Jan. 2026.
10. P. Mishra and M. J. Koehler, "Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge," Teachers College Record, vol. 108, no. 6, pp. 1017-1054, 2006.