Effect of Artificial Intelligence tools on Academic Performance of students: A Meta Analysis

Zaheer Abbas [1*] & Hafiz Tahir Jameel[2]

Journal of Educational Psychology and Pedagogical Sciences (JEPPS)

ISSN:2791-0393 (Print) eISSN: 2791-0407

Vol.5, No. 2, (Jul-Dec, 2025): 46-66

[1*]PhD Scholar, Department of Educational Planning, Policy studies & Leadership, Allama Iqbal Open University, Islamabad. Pakistan. E-Mail: abasszaheer71@gmail.com. (https://orcid.org/0009-0003-3736-7797)

[2]Assistant Professor, Department of Special Education, Allama Iqbal Open University, Islamabad. Pakistan. E-Mail: tahir.jameel@aiou.edu.pk. (https://orcid.org/0000-0003-4948-5929)

Submitted 17 July 2025

Accepted 10 Dec 2025

Published 31 Dec 2025

DOI: https://doi.org/10.52587/jepps.v5i2.121

https://jepps.su.edu.pk/article/53

OPEN ACCESS

ABSTRACT

Artificial intelligence (AI) has gained rapid development in the field of education for improving classroom instruction, assessment and students' learning outcomes. This meta-analysis specifically focused on the overall impact of AI on the academic performance of students. Using "Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)" guidelines, the study synthesized 18 studies retrieved from three databases ("Directory of Open Access Journals - DOAJ, PubMed, and Google Scholar") published between 2015 to June 2025 and used the information extracted from the studies for the analysis. The studies comprised a sample of 3436 participants from various educational levels fulfilled the inclusion criteria. For the analysis, the data from the studies was coded and entered into "Jaffery's Amazing Statistical Program (JASP) version 0.18.3". The results showed that there was a moderate to large (0.62) pooled effect size under the random-effects model. Moderating effects of four variables, type of institution, sample size, area of the subject and AI types were examined. Results depicted that the sample size moderated the effectiveness of AI in predicting the academic performance of students significantly. However, other moderating variables, that is, type of institution, area of the subject and AI type did not show any significant effects. Based on the findings, the meta-analysis presented practical implications for both policy and practice.

Keywords: Artificial Intelligence, Academic Performance, Moderators, Forest Plot, Funnel Plot.

1. Introduction

Artificial intelligence has gained rapid momentum in all fields of life including education for the last few decades. The research has found that artificial intelligence has been impacting all areas of education starting from early childhood care and education to higher education. Within education, the major areas being impacted by artificial intelligence include students' learning outcomes, teachers' learning, classroom teaching, assessment, and tasks related to management and administration. The foremost outcome of using artificial intelligence can be

on the overall effectiveness of educational institutions.

There is increasing volume of studies carried out on how artificial intelligence influences the academic learning outcomes of students in educational institutions (Fu, et al., 2024). Literature reports that AI solutions provide learning platforms for students in enhancing their achievements by compensating for lack of qualified teachers, teaching and learning resources, and traditional teaching processes, and so on. However, there is a need to investigate how AI relates with the learning outcomes of students. This gap in literature warrants critical reviews of the existing research and a need to explore the tools of AI and how they are associated with the performance of students in schools by scrutinizing the recent research. Searching Google and Google Scholar using the strings of words like 'Meta-analysis', artificial intelligence', and 'academic achievements' resulted in only a few reviews (Arar et al., 2025; Bond et al., 2024) with no major focus only on AI and its relationship with the academic achievements of students suggesting an intensive meta-analysis on it. The focus of the reviews has remained around the role of AI in education, classroom instruction, early childhood education, higher education, teachers' professional development and so on (Arar et al., 2025; Bond et al., 2024; Salas-Pilco, et al., 2022a; Sharadgah & Sa'di, 2022). Therefore, a meta-analysis of the AI and students' academic outcomes was conducted.

In addition, the findings of the previous research on how AI influences the results of students is inconsistent. For example, a few studies highlighted the positive impact of AI on academic performance (Setiawan, 2025; Vieriu & Petrea, 2025; Shahzad et al., 2024; Zhang et al., 2024) while another group of studies showed negative or hindering results on the academic achievements of students (Klimova & Pikhart, 2025; Abbas et al., 2024; Basha, 2024; Weck, 2024). Similarly, a few studies also highlighted that there is not any significant relationship between the technologies and the learning outcomes of the students (Koć-Januchta, 2020). In addition, there are inconsistencies in the methodologies adopted by these studies making it difficult to draw conclusions from the studies. Therefore, it warranted a meta-analysis exploring how AI tools influence the academic scores of students.

2. Research Objectives

The objectives of the research were to:

1. explore the AI impact on academic performance of students.
2. investigate the variables moderating the AI impact on academic performance of students.

3. Research Hypotheses

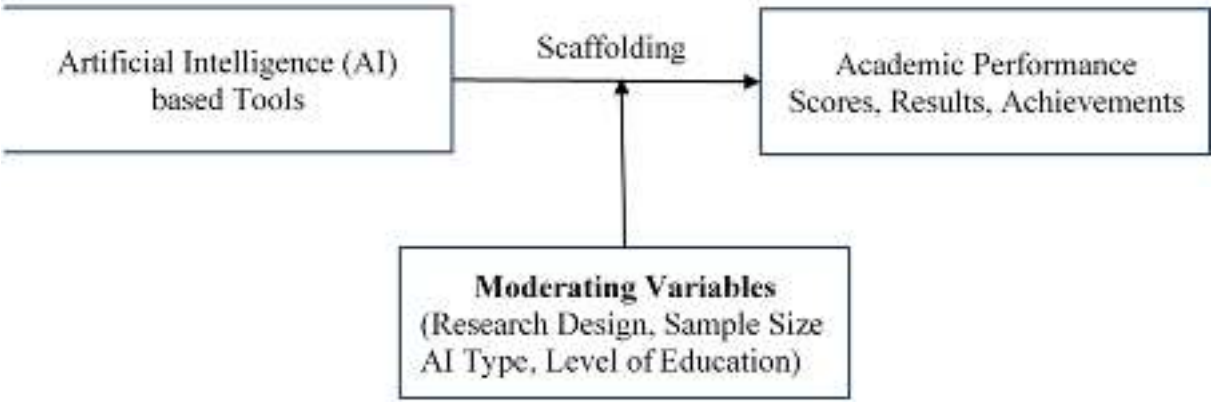
H$_o$1: There is no statistically significant AI impact on academic performance of students.

H$_o$2: There is no variable moderating the AI impact on academic performance of students.

4. Theoretical Framework

The interplay of artificial intelligence and the academic achievements of learners in based on Vygotsky's social sociocultural theory of cognitive development involving the notion of "zone of proximal development". Artificial intelligence supports academic performance through various scaffolding tools for developing learning of the students including adaptive tools and intelligence tutoring and so on (Roll & Wyllie, 2016). However, its poorly implemented usage may result in over reliance on it and negatively affect cognitive abilities of the students (Ng et. al., 2021). Figure 1 depicts the hypothesized relationship among the artificial intelligence, scaffolding process and the achievements of learners with the moderating variables for the study.

Figure 1.



The correlation between artificial intelligence tools and quantitative learning outcomes of students is not linear rather it is affected and moderated by several other variables. For the meta-analysis, those variables were included for which relevant data was available in the extracted and selected studies. The study focused on academic performance of the students instead of broader learning outcomes as the outcomes may not be quantifiable due to their broader qualitative nature. In addition, the qualitative aspects of the learning outcomes were possible to analyze using the software package. Therefore, the academic performance in terms of marks, scores, cumulative grade point averages were taken.

5. Literature Review

How AI relates Education

Technology has increasingly been developing and evolving in education since its inception in the early 1900s (Almasri, 2024). However, with the introduction of artificial intelligence (AI) as a form of education has progressively been used in almost all aspects of education. It has improved access to education by addressing physical barriers through several online and distance learning modes. It has also enhanced the learning possibilities for the people for different physical and other disabilities through various digital assistance aids and other measures (Setiawan, 2025). In addition to access, it has also been supporting the educational managers in their administrative tasks while streamlining their activities including students' admissions and enrollments, managing human and financial resources, scheduling classes and managing transition, tracking students and other HR attendance and performance and so on. Technology integrated tools have also been in use for fostering communication among students, teachers, parents and other stakeholders for managing inquiries and developing collaborations. Lastly, the analysis of accurate and efficient data helps the policymakers in taking timely decisions for policy formula and updating it related to all affairs of the schools (Núñez-Canal, et al., 2026).

Anticipated Benefits of Artificial Intelligence for Classroom Instruction

Artificial intelligence is considered instrumental in transforming the learning process into educational institutions. On one hand, it assists the teachers in devising student-centered learning classroom instruction making it more engaging and inclusive for the learners while on the other hand, AI provides opportunities for students to enhance their learning outcomes through personalized learning experiences, ensuring their engagements in classrooms and improving academic scores (Zawacki-Richter et al., 2019).

AI transforms classroom instructions by facilitating teachers in the classroom instruction

process resulting in improved learning outcomes of the students. For instance, AI provides teachers with resources and activities for developing engaging and effective lessons for their classrooms.

Such lessons may help teachers in achieving the lesson goals in more efficient ways while saving preparation time of teachers. The focus of such lessons will also be on developing individualized learning plans and activities for catering individual needs and differences of the learners. In addition, AI-powered tools can also assist teachers in classroom management through attendance tracking and progress monitoring solutions. Such solutions also contain automated grading systems for optimizing effective assessment and performance monitoring. It will free up teachers for arranging other hands-on and minds-on engagements and activities for students in classrooms. By fostering student-centered activities through AI-powered tools, the process of teaching and learning can be improved for achieving the improved learning outcomes of students.

AI also helps in developing personalized learning experiences for the students. For example, AI can adapt the learning contents for the students based on their needs and interests. It can also adjust the difficulty level and pace of learning for the students based on their own personalized needs. Students can also explore relevant resources and activities specifically tailored to their goals and learning styles. Such resources and activities can help them in tracking their learning and progress in predefined learning routes considering their self-paced progress. Similarly, AI powered intelligent tutoring solutions can provide virtual platforms for the students supporting one-on-one support for learning and offering real-time feedback and guidance on the tasks completed by the students. Through its chatbots and virtual assistants, it may also facilitate the students through interactive learning and real time answer the questions during the one-on-one interactions. Such personalized learning experiences may support students in developing better understanding of the concepts, enhancing skills, and acquiring dispositions and resultantly enhancing their academic outcomes.

How do Artificial Intelligence Tools Influence Students' Achievements?

Review of literature revealed that there is a growing number of studies on the how AI mediates students' scores. Most of the studies have reported positive impacts of AI on their academic learning improving their performance, self-efficacy, engagements, and attitudes towards education. For instance, Vieriu and Petrea (2025) conducted their study on 85 second year students who had direct exposure to technologies in their classrooms. The findings revealed that AI tools significantly predict the achievements of students, such as on their performance in examination, projects, and examination grades improving their results. A similar study conducted by Shahzad et al., (2024) on the effects of social media particularly AI school performance and mental well-being of around 400 university students reported positive impacts on their learning and well-being. Zhang et al., (2024) also conducted an exploratory study on 20 university students in Portugal on how AI generated videos impact their learning outcomes and acceptance using a mixed model. The outcomes of pre-and-post tests revealed that the AI generated videos and material significantly improved the learning outcomes of students suggesting AI generating videos and materials as important pedagogical tools. The previous studies discussed above have explored how AI relates with the school results of students, however, these studies have fallen short of presenting a comprehensive synthesis of how AI correlates with the school-based achievements of students. For this reason, this research systematically analyzed the impact of artificial intelligence on the cumulative results of students.

6. Methodology

The study employed "Preferred Reporting Items for Systematic Reviews and Meta-Analysis" (PRISMA) principles for the analysis of metadata and synthesize the results of quantitative studies on relationship between AI tools and their impact on the school-based achievements of students. These guidelines provide a checklist of items for detailed information about the included studies of the meta-analysis, search strategy, "inclusion and exclusion criteria" and so on. The detailed report of each step is given. For ethical considerations, each step of the search strategy, "inclusion and exclusion criteria" of the studies, number of search results and analysis steps are clearly mentioned for understanding and reporting. The authors followed ethical principles and guidelines in each step to ensure transparency and responsibility.

Sources Search Strategy

For searching the sources, a comprehensive search strategy was developed. For this purpose, three open access databases, namely, "Google Scholar", "PubMed", and "Directory of Open Access Journals" "(DOAJ)". were identified. These databases were searched using several keywords. Table 1 presents the lists of keywords for both artificial intelligence and academic performance.

Table 1. Terms related to AI, Academic Performance and Educational Instituions

Terms related to AI	Terms Related to Academic Performance	Terms Related to Educational Institutions
Artificial Intelligence	Academic performance	University
Chatbots	Academic achievements	College
ChatGPT	Learning outcomes	School
Language processing	Results	Elementary
Machine learning	Scores	Secondary
ICT	Marks	Primary
		Profession
		Medical
		Engineering

The terms mentioned in Table 1 were used by combining one term from column 1 and one from 2 and using any keyword from column 3. For instance, the key terms were combined as "Artificial Intelligence and Academic Performance", "Artificial Intelligence and Academic Achievements", "AI and Academic Performance". Synonyms for academic performance were also used like achievements, scores, and results.

In advanced search options, different filters were applied, like in category, only journal articles', in type, "peer reviewed" articles, and in timeline, articles from "2015 to present" were selected. Boolean operators like 'AND' and 'OR' were also used for the precision in search results. Search results in each database were counted. Then individual articles were open and read to explore the relevance of each article with the research questions.

Inclusion and Exclusion Criteria

For narrowing down the search and selection process of studies, an explicit "inclusion and exclusion criteria" were developed. Table 2 presents the criteria.

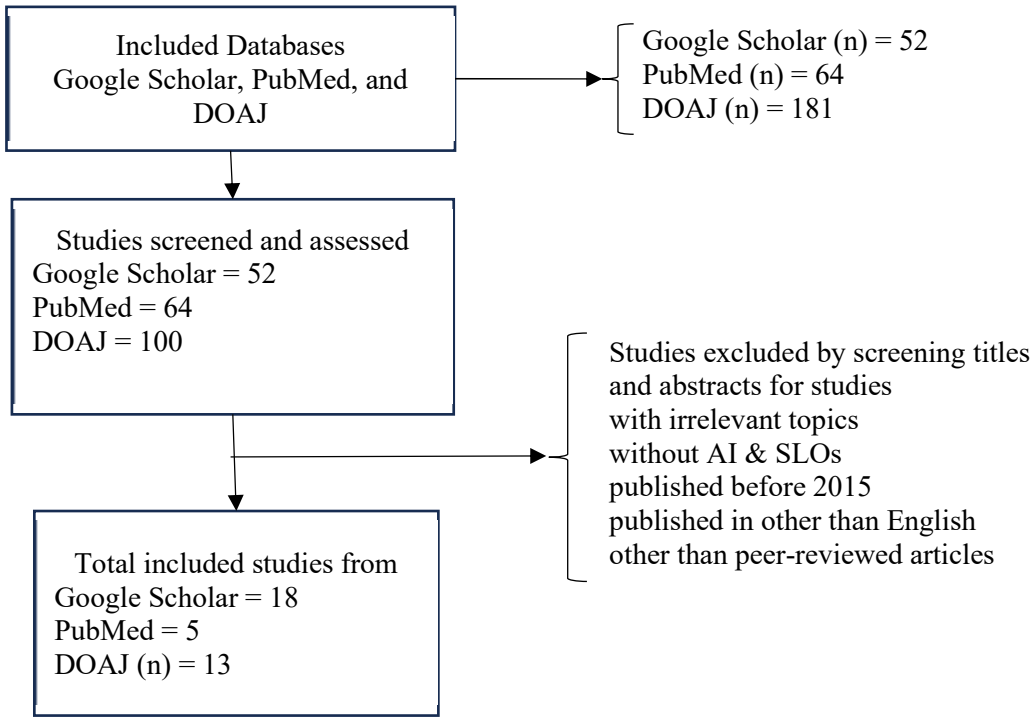
Table 1: Inclusion and Exclusion Criteria used in the study

Categories	Inclusion Criteria	Exclusion Criteria
Areas of Studies	Studies should contain AI and students' learning outcomes in any form.	Studies without AI and academic performance or with only general areas of education were excluded.
Publication date	The studies should have been published between 2015 to 2025.	Studies published before 2015 were excluded. The period 2015–2025 was chosen due to its relevance to the evolution of AI technologies such as ChatGPT and adaptive learning systems.
Publication type	Studies included must be from the peer-reviewed journal articles.	Studies including reports, proceedings, unpublished theses, dissertations, or chapters of the books were excluded.
Research type	Studies with different quantitative research designs were included, like quasi-experimental or true-experimental studies, or correlational surveys etc. Studies should also report sufficient data for the extraction of effect size	Studies other than quantitative research designs including mixed-method studies were excluded.
Language	The studies using English language were included.	Studies written in other languages were excluded.
Outcome Measure	Studies must report quantitative data on academic performance like marks, percentages, grades etc.	Studies that report only perceptions, attitudes or any other qualitative data were excluded.

Search Selection

The databases selected for the search were explored separately using the keywords and Boolean operators. The search and selection process is presented in Figure 2.

Figure 2: Search and Selection Process



In DOAJ database, when search process resulted in 181 indexed articles were found in 19 pages. The titles of first 100 entries (10 pages of search result) were browsed to explore the relevance of the studies with the research topic. 35 titles of the articles were found relevant and then the abstracts of the articles were read to explore further relevance of the studies. Based on the review of the abstracts and inclusion criteria, 13 relevant research articles were selected for analysis.

The PubMed database was also searched using the keywords. The applied filters were "Free full text", "Adaptive Clinical Trial", "Classical Article", "Clinical Trial", "Comparative Study", "Controlled Clinical Trial", "Equivalence Trial", "Evaluation Study", "Introductory Journal Article", "Periodical Index", "Randomized Controlled Trial", "Clinical Trial", "Veterinary", and from "2015 – 2025". The search process reported 64 results. All titles of all the 64 articles were read and 15 titles were selected based on their relevancy. The abstracts of 15 articles were read based on the "inclusion and exclusion criteria". After review, 5 articles were selected due to their relevancy based on the criteria.

"Google Scholar", the third database, was also searched using the keywords. The search resulted in 52 index studies. The titles of the articles were read, and 29 articles were selected for the review of abstracts. After reviewing the abstracts, 18 articles were identified as the most relevant for the meta-analysis based on the "inclusion and exclusion criteria".

After finalizing 36 articles from all three databases, the selected articles were reviewed to explore any possible duplicate articles. Table 3 presents the details of duplicate articles in databases.

Table 3. Duplicate articles in database

1. **Databases**	**Total Selected Articles**	**Duplicates**	**Finalized Removing Duplicates**
DOAJ & PubMed	18	05	13
DOAJ & Google Scholar	31	13	18
PubMed and Google Scholar	23	05	18
DOAJ, PubMed & Google Scholar	**36**	**18**	**18**

From Table 3, it is evident that there were duplicate articles in any two selected databases as well as in all the three selected databases. After removing the duplicate articles, 18 research articles were finalized for the meta-analysis.

Data Extraction from Individual Studies

For the extraction of data from the selected studies, a template was developed for data entry. The major data entries included, "title of the study", author(s), "year of publication", country/region, study type, sample size, type of AI, academic performance, and name of journal". After extracting data from all the selected studies, coding was carried out. The codes were given to all entries required for analysis through the software, including research design, sample size, subject area, type of AI used, and education level. The codes and dimensions are presented in Table 4.

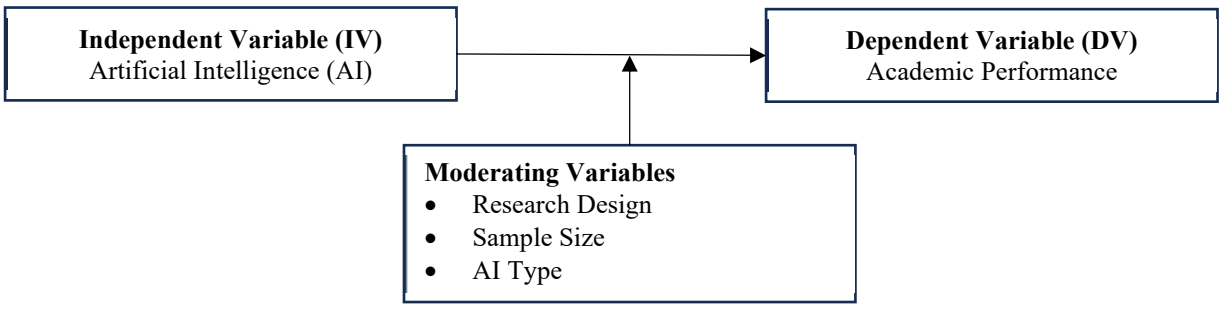
Table 4. Codes and Dimensions

Dimensions	Variables	Categories
Research related Characteristics	Research Design	Experimental, Non-Experimental
	Sample Size	1-50, 51-100, 101-300, >300
Opportunity to Learn	Grade	University, College, School
	AI Type	AI Interventions, Chatbots, Gamified Ai, Others

7. Data Analysis

For the analysis of the data, a modified version of an "opportunity to learn (OTL)" model was used to analyze the effect of AI on the academic performance of students. The analytical framework developed for the study comprised of the moderating variables, such as research design, sample size, AI type, and level of education presented in Figure 3.

Figure 3. Analytical Framework of the study



i. Descriptive Results

The purpose of the meta-analysis was to investigate whether AI predicts the school-based academic achievements based on the 18 selected studies published between the year 2015 and 2025. The details of the selected research studies are given in Table 5.

Table 5. Research studies selected for the study

Authors	Article Title	Journal	Peer-Reviewed	Indexed In
"Azamatova et al. (2023)"	"The Effect of Using Artificial Intelligence and Digital Learning Tools based on Project-Based Learning Approach in Foreign Language Teaching on Students' Success and Motivation"	"Int. J. of Education in Mathematics, Science, and Technology	Yes	"Google Scholar"
"Han et al. (2022)"	"Analysis of the effect of an artificial intelligence chatbot educational program on non-face-to-face classes": "A quasi-experimental study"	BMC Medical Education"	Yes	"PubMed, Google Scholar DOAJ"
"Effiong et al. (2024)"	"Effect of EduGPT artificial intelligence on academic performance of educational technology students": "A quasi -experimental study"	"Prestige Journal of Education"	Yes	"Google Scholar"
"Tong et al. (2022)"	"The effectiveness of blended learning on students' academic achievement, self-study skills and learning attitudes": "A quasi-experiment study in teaching the conventions for coordinates in the plane"	"Heliyon"	Yes	"PubMed, DOAJ Google Scholar"
"Yang et al. (2025)"	"The effectiveness of ChatGPT in assisting high school students in programming learning": "Evidence from quasi-experimental research"	"Interactive Learning Environments"	Yes	"Google Scholar"
"Elsayed et al. (2024)"	"Teacher support in AI-assisted exams: an experimental study to inspect the effects on demotivation, anxiety management in exams, L2 learning experience, and academic success"	"Language Testing in Asia"	Yes	"DOAJ, Google Scholar"

Authors	Article Title	Journal	Peer-Reviewed	Indexed In
"Eltahir & Babiker (2024)"	"The influence of artificial intelligence tools on student performance in e-learning environments - Case study"	"Electronic Journal of e-Learning"	Yes	"Google Scholar"
"Song & Song (2023)"	"Enhancing academic writing skills and motivation": "Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students"	"Frontiers in Psychology"	Yes	"PubMed, DOAJ Google Scholar"
"Gürefe et al. (2024)"	"Investigating the impact of the AI supported 5E (AI-s5E) instructional model on spatial ability"	MDPI (various journals)	Yes	"PubMed, DOAJ Google Scholar"
"Ali et al. (2025)"	"Impact of artificial intelligence on the academic performance and test anxiety of pharmacy students in objective structured clinical examination": "A randomized controlled trial"	"Int. J. of Clinical Pharmacy"	Yes	"Google Scholar PubMed"
"Liao et al. (2023)"	"Design and implementation of an AI-enabled visual report tool as formative assessment to promote learning achievement and self-regulated learning": "An experimental study"	"British Journal of Educational Technology"	Yes	"DOAJ, Google Scholar"
"Zhang et al. (2025)"	"Does ChatGPT-based reading platform impact on foreign language paper reading? Evidence from a quasi-experimental study on Chinese undergraduate students"	"Education and Information Technologies"	Yes	"DOAJ, Google Scholar"
"Dai et al. (2023)"	"An embodied, analogical and disruptive approach of AI pedagogy in upper elementary education": "An experimental study"	"British Journal of Educational Technology"	Yes	"DOAJ, Google Scholar"
"Menekse et al. (2025)"	"Enhancing student reflections with natural language processing-based scaffolding": "A quasi-experimental study in a large lecture course"	"Computers and Education: Artificial Intelligence"	Yes	"DOAJ, Google Scholar"

Authors	Article Title	Journal	Peer-Reviewed	Indexed In
"Klarin et al. (2024)"	"Adolescents use and perceived usefulness of generative AI for schoolwork": "Exploring their relationships with executive functioning and academic achievement"	"Frontiers in Artificial Intelligence"	Yes	"PubMed, DOAJ Google Scholar"
"Almogren et al. (2024)"	"Integrated technological approaches to academic success": "Mobile learning, social media, and AI in visual art education"	"Digital Object Identifier (check actual journal name)"	Yes	"DOAJ, Google Scholar"
"Essel et al. (2022)"	"The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education"	"Int. J. of Educational Tech in Higher Education"	Yes	"DOAJ, Google Scholar"
"Baltezarević & Baltezarević (2025)"	"Digital game-based learning (DGBL) effect on students' academic performance"	"Int. J. of Cognitive Research in Science, Engineering and Education"	Yes	"Google Scholar"
"Arkün-Kocadere & Çağlar-Özhan (2024)"	"Video lectures with AI-generated instructors: low video engagement, same performance as human instructors"	"Int. Review of Research in Open and Distributed Learning"	Yes	"DOAJ, Google Scholar"

All the studies mentioned in Table 4 used AI based interventions in the experimental or treatment group while conventional or routine-based teaching in control group. In total, there were 3436 participants. Of the 18 selected studies, there were 6 (33.3%) experimental, 12 (67%) were non-experimental including quasi-experimental, correlational and mixed-method research. Regarding the sample size, most of the studies 8 (44.4%) consisted of 51-100 participants, followed by 5 (27.7%) studies with 151 or greater number of participants, 3 studies (16.6%) with 1-50 participants, and 2 studies (11.11%) with 101-150 participants. Similarly, 10 (55.5%) studies were conducted in university settings, 6 (33.3%) studies in school and 2 (11.11%) studies in professional college settings. The focus of most of the studies remained on math, science or other similar disciplines 7 (38.8%), followed by technology 6 (33.3%) and language 5 (27.7%). Most of the studies 9 (50 %) either used gamification, mobile application and other forms of artificial intelligence, followed by ChatGPT 5 (27.7%), and Chatbots 4 (22.2%).

For analyzing the data, Jaffery's Amazing Software Program (JASP) version 13 was used. The analysis of the data mainly included calculating pooled effect size from the individual effect sizes from the studies to explore the overall relationship between AI and academic performance of students in academic settings. "Hedge's g" was calculated to reduce sampling bias which is mostly evident in small samples, and it is evident in Cohen's d. Hedge' g statistically reduces this bias. For exploring the pooled effect size, random effects model was

used as the studies selected for the research vary in terms of research design, population and sample size, and other methodological variations. Q-statistics are used to show level of heterogeneity in effect sizes while heterogeneity explains the degree of differences or variability in effect sizes. were used to justify the random effects model as the Q-statistics were significant, suggesting the heterogeneity of the true effect sizes among the studies. Forest plot was used for visual summary of the pooled effect size as well as the effect sizes and confident intervals of each study for their relative weight in the meta-analysis. For the analysis of the moderators, a direct module in meta-analysis was not available in the software. However, there was an option of using R console in the program for using script to compute the effects of moderators. For this purpose, scripts were prepared and used for the analysis of moderators in R console. Results retrieved by running the script were then formatted in tables for presentation and interpretation (Table 4 to 6).

Overall Effect Size

For analyzing the variability across 18 effect sizes, heterogeneity analysis was carried out. Results revealed that ($Q\ [18] = 199.75$, $p < .017$, $I^2 = 95\ \%$) there is substantial variation among the effect sizes suggesting that this variable may not be because of the sampling error rather than between-studies differences. The results also showed that the variance the variability among effect sizes was partly affected by other variables suggesting a need for the moderators' analysis. Based on the results of variability and the various characteristics of the study including different designs, AI types, level of institutions etc., random-effects model was used for the analysis of the data.

For exploring the patterns of overall effect size, a test of the pooled effect size was employed. Results are presented in Table 5.

Table 6: Pooled Effect Size

Estimate	Standard Error	t	df	p	95% CI		95% PI	
0.62	0.23	2.70	16.000	0.017	0.13	1.11	-1.27	2.5

Results indicated that there is a significant pooled effect size d = 0.62, Standard Error = 0.23, t(18) = 2.70, p = 0.017, [95% CI [0.13, 1.27]. These results indicate a moderate to large overall effect size across all the studies. The 95% Prediction Interval (PI) values ranged from -1.27 to 2.50 indicating an average positive effect with large variation in true effect size within the individual studies. These variations are beyond what could be expected by chance suggesting a random-effect model for analysis. The effects of individual effect sizes of the studies on the overall effect size and the pooled effect size are presented in Figure 4.

Figure 4. Forest Plot Depicting Effect Sizes

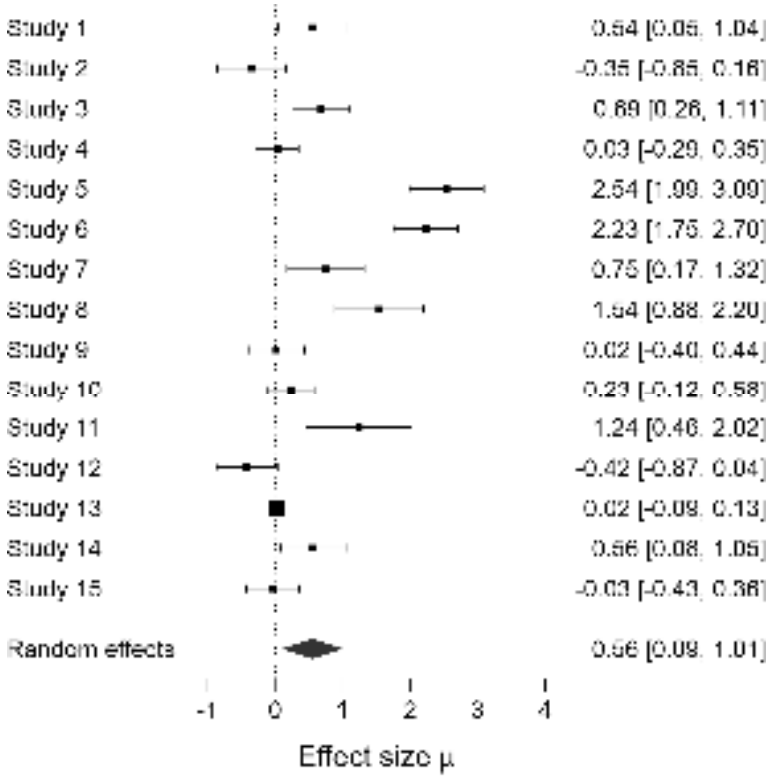


Figure 4 depicts the effect sizes of 18 studies and the heterogeneity with a confidence interval of 95%. The lines horizontally shown in the plot depict intervals of confidence for each study. The wider lines reflect higher uncertainty in the estimates of their effect sizes. The sizes of the squares within each horizontal line showed the assigned weight for each study during analysis of the data from each study. The larger size of the square shows the studies with larger precision particularly due to their bigger sample sizes. The given diamond at the bottom of the larger horizontal line (x-axis) represents the weighted average of the effect sizes with confidence interval of 95% calculated through random effect model which explains the between-studies variability.

During the analysis, it was already found that there were large variations in true effect size within the individual studies suggesting a random-effect model for analysis; therefore, random effect model was used. The forest plot highpoints the high-level variability in the effect sizes across all 16 studies as evident in the overlapping broad and confidence intervals. Additionally, the plot also shows that most of the effect sizes and their confidence intervals are on the same side of the line suggesting moderate heterogeneity.

In the forest plot, the diamond is on the left side of zero which is also evident from the pooled effect size (0.62) highlighting a positive effect size in favor of the treatment variable, that is, the impact of artificial intelligence on the academic performance of students. Similarly, the diamond and the confidence intervals do not cross the vertical line at zero showing that the effect size is statistically significant.

To explore the effects of moderators on the effectiveness of AI on students' academic performance, four moderating variables, that is, research design, sample size, AI type and level of education were taken. Analysis was carried out to explore the weighted effect sizes of the moderators and results are presented in Table 7.

Table 7: Moderators' Analysis

Variable	Category	*K*	*Hedge's g*		*Qw*	*df*	*p*	*Qb*	*p*
			Mean	*SE*					
Type of Institution								1.87	0.392
	University	10	1.027	0.351	61.76	9	<0.001		
	College	2	0.942	1.288	18.90	1	<0.001		
	School	6	0.318	0.257	94.42	5	<0.001		
Subject Area of Intervention								3.41	0.182
	Language	5	0.82	0.29	60.91	4	<0.001		
	Technology	6	0.65	0.41	33.92	5	<0.001		
	Math/Science	7	1.04	0.28	29.31	6	<0.001		
Size of Sample								26.84	<0.001
	1–50	3	1.43	0.33	25.13	2	<0.001		
	51–100	8	0.94	0.22	132.39	7	<0.001		
	101–150	2	0.44	0.18	11.58	1	0.010		
	>150	5	0.04	0.11	29.5	4	0.01		
AI Intervention								1.09	0.780
	ChatGPT	5	0.72	0.32	30.67	4	<0.001		
	Chatbot	4	1.05	0.25	172.47	3	<0.001		
	Gamified and other Forms AI	9	0.92	0.44	16.41	8	<0.001		

According to the results presented in Table 6, the type of institution moderated the impact of AI interventions on the academic performance of students non- significantly [Qb = 1.87, df = 3, p = 0.392]. Breaking down the results by level of institutions provide significant insights. For example, the level of institutions for university context yielded a moderate to high effect size (g = 1.03) with a substantial variability (Q = 61.76, df = 9, p < .001). Similarly, college level education resulted in a moderate to high effect size (0.94) with a considerable low variability (Q = 18.90, df = 1, p < .001). Lastly, the school level education showed a low effect size (0.318) with a considerable high variability (Q = 94.42, df = 17, p < .001). These results suggest a more robust investigation into how the impact of moderators can be further explored for nuanced insights.

The analysis of second moderator in Table 5 revealed that the effect of the subject area as a moderator on the impact of AI on the academic performance of students was also non-significant [Qb = 3.41, df = 3, p = 0.182]. Further analysis depicted that the effect size for language area was moderate to high (0.82) with a moderate variability (Q = 60.91, df = 4, p < 0.001). The effect size for AI intervention on technology related fields was also moderate to high (0.65) with a comparatively low variability (Q = 33.92, df = 5, p <001). Lastly, the effect size for the intervention of AI on area of match, science and other disciplines showed a high effect size (1.04) with a considerable low variability (Q = 29.31, df = 6, p < 0.001).

These insignificant results suggest that there would be other factors which need to be considered rather than the role of the subject area of AI intervention.

The third moderator of the study was the sample size. The overall effect size for this moderator was significant showing that the sample size significantly influences the academic performance of the study [Qb = 26.84, df = 4, $p < 0.001$]. The breakdown of results by groups sample sizes reported that the sample with a size of 1-50 participants had a high effect size (1.43) with a low variability (Q = 25.13, df = 2, $p < 0.001$). The effect size for the second sample group with size 51-100 participants was also moderate to high (0.94) with a very high variability (Q = 132.39, df = 7, $p < 0.001$). The third group of studies with sample size of 101-150 participants had a low effect size (0.44) with a very low variability (Q = 0.18, df = 1, $p < 0.001$). Lastly, the group of studies with sample size higher than 150 participants also had a very low effect size (0.04) with again a very low variability (Q = 0.11, df = 4, $p < 0.001$). These results propose that the sample size could be an important factor which may need to be considered while exploring its impact on the overall influence of AI on school-based academic results.

The last moderator of the study was the type of AI used in the studies. According to results presented in Table 5, the type of AI had no significant effect on the academic performance of the students [Qb = 1.09, df = 4, p = 0.78]. Further breakdown of the type of AI showed that ChatGPT has a moderate to high effect size (0.72) with a low variability (Q = 0.32, df = 4, $p < 0.001$). The effect size for Chatbots was high (1.05) with a low variability (Q = 0.25, df = 3, $p < 0.001$) while the effect size for Gamified and other forms of AI was also high (0.92) with a low variability (Q = 0.44, df = 8, $p < 0.001$). The low variability results recommend that other methodological and contextual factors need to be included for a more robust analysis instead of the type of AI used in the study.

Publication Bias

The publication bias was explored using funnel plot presented in Figure 5.

Figure 5. the Funnel plot of publications

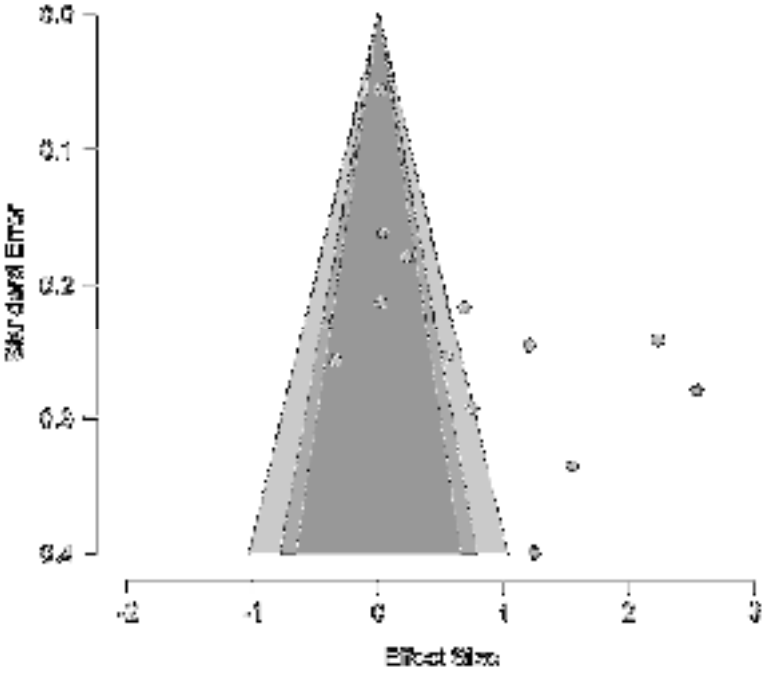


Figure 5 shows that there is slight asymmetry depicting a small degree of potential publication bias. It is evident from the plot that there was a comparatively higher representation of the studies with large effect sizes given on the rightward of the plot while no study on the left side. As there is a small asymmetry, therefore, it does not have any significant effect on the pool effect size of the study. A few studies with moderate standard error fell outside the expected range evident in the plot, however, their influence appeared limited in the meta-analysis. As most of the studies are clustered around the pooled values, therefore, it shows that the results are reliable. Hence, the findings of the meta-analysis are trustworthy and provide significant insights into the impact of artificial intelligence on academic performance of the students.

8. Discussion

This meta-analysis examined the impact of artificial intelligence on the academic performance of students. It used data from 18 research articles from different research journals. The analysis revealed that the impact of AI on the academic performance of students was positive as well as statistically significant. There was a moderate to large overall effect size under the random effect model. This positive effect of the meta-analysis was aligned with the results of the previous meta-analysis of Sun and Zhau (2024). This result suggests that when students are engaged in their learning process through the integration of artificial intelligence tools, they show better performance as compared to their counterparts. This difference can be as broader as 0.62 standard deviations from traditional learning settings. The major AI tools which may help in enhancing students' academic performance include gamified AI tools, machine learning, Chatbots and ChatGPT. These tools provide personalized teaching ensuring differentiated instruction and timely feedback to students. Such tools help students identify their progress and predict their performance through different adaptive learning systems.

From the moderators' analysis, it was evident that sample size significantly impacted the academic performance of students. Therefore, it was evident that the sample size influences the effectiveness of AI technologies on the academic performance of students and this finding corroborates with the findings of Zheng et al., (2021). The further breakdown of the results into sample groups revealed that the effect size for the sample group with 1-50 participants had the highest effect size as compared to other sample size groups. This finding proposes that studies with comparatively smaller sample sizes may depict larger effect sizes as compared to those with larger samples sizes. Dong et al. (2025) proposes that this result may be due to small sample size studies carried out in more environments which are more controlled in which interventions of AI might be designed addressing the specific needs and characteristics of the participants. The finding also suggests that it might be challenging in studies with larger sample size to create a personalized learning environment and achieve higher academic performance due to greater variations in the learning context as well as in personal characteristics of students (Nelson et al., 2015; Salvin & Smith, 2009).

However, the study did not find any significant effects of types of institutions, subject areas and types of AI tools as moderators on the effectiveness of AI on academic performance of students. This finding of the study is supported by the findings of the study of Steenbergen Hu and Cooper (2013) as they also found similar results. From these similar results, it can be hypothesized that the insignificant effects of types of institutions, subject areas and types of AI tools as moderators on the effectiveness of AI might be common in such similar studies. However, there is a need to conduct around these effects of moderators on the role of AI in academic performance of students to validate the assumption.

9. Limitations of the Study

The study has a few limitations. Firstly, the study used 18 research studies from which 3 studies did not use any experimental and control groups. Therefore, the interpretations of the findings need to be made in cautious as the study did not compare both types of groups. Secondly, the meta-analysis used studies from different levels of education, therefore, using these findings for specific settings other than the ones used in the meta-analysis should be considered carefully. Thirdly, the meta-analysis used only four moderators for exploring whether they influence the effectiveness of AI on academic performance of students, there could be other factors like students' socioeconomic backgrounds, teacher and school related factors and classroom environment which may also affect the effectiveness of AI and these variables were not controlled. Lastly, the meta-analysis used only three databases, and the meta-analysis used studies from the three databases. There would be more relevant studies in other databases which this meta-analysis might have missed. Inclusion of those studies from other databases may report different results.

10. Conclusion and Implications

This study examined the overall effect size of AI on academic performance of students using meta-analysis and reported a positive effect of AI. The meta-analysis synthesized different previously conducted studies and contributed to understanding of the current status of AI and its effectiveness on academic performance of students. The study also examined the moderating variables and their effects on AI and its effectiveness, including type of institutions, sample size, subject area, and type of AI tools and found sample size as a significant moderator.

Based on the findings, teachers should be encouraged to plan and implement AI interventions for improving academic achievements of students at all levels. For this purpose, educational managers should encourage teachers to use AI tools for improving the effectiveness of their classroom instruction. Educational institutions may provide such tools for the teachers and arrange professional development activities for integrating these tools in their teaching. Specifically, the highest effect size at university level education shows that there is a need to improve teaching and learning process at university classes through the integration of AI tools. As the effect size for science, mathematics and relevant subject areas was highest, it might be due to the reason that AI might have created such opportunities and experiences for students through which the abstract and complex concepts of science and mathematics are made easy for students to understand. Therefore, the structured and continuous use of AI tools in these and other subjects to be made in teaching and learning process (Zheng, et al., 2021). Finally, in this meta-analysis, there were three research studies which didn't provide complete information about their studies due to which effect size for the studies could not be computed. Therefore, it is suggested to researchers provide complete details and statistics of their studies because incomplete statistics may limit their usability in future studies (Hwang, 2022).

DECLARATION STATEMENTS

This paper is extracted from the first author's thesis

Conflict of Interest

None to declare

Ethical Permissions

This study received ethical clearance from the AIOU, Islamabad.

Funding
None to declare

Author Contribution
ZA conceptualized the study, collected & analyzed the data and drafted. JT analyzed data and reviewed final version of this manuscript.

11. References

Abbas, M., Jam, F. A., & Khan, T. I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. International Journal of Educational Technology in Higher Education Abbas *et al. Int J Educ Technol High Educ, 21*(10).

Ali, S. M., Rehman, S. & Cheema, E. (2025). Impact of artificial intelligence on the academic performance and test anxiety of pharmacy students in objective structured clinical examination: A randomized controlled trial. *International Journal of Clinical Pharmacy.* https://doi.org/10.1007/s11096-025-01876-5

Almasri, F. (2024). Exploring the impact of artificial intelligence in teaching and learning of science: A systematic review of empirical research. *Res Sci Educ* 54, 977–997.

Almogren, A. S., Al-Rahmi, W. M. & Dahri, N. A. (2024). Integrated technological approaches to academic success: Mobile learning, social media, and AI in visual art education. *IEEE Access, 12*, 175391-175413. https://ieeexplore.ieee.org/document/10752986

Arar, K., Tlili, A., Schunka, L., Salha, S. & Saiti, A. (2025). Reimagining Educational Leadership and Management Through Artificial Intelligence: An Integrative Systematic Review. *Leadership and Policy in Schools, 25*(1), 4-26.

Arkün-Kocadere, S., & Çağlar-Özhan, S. (2024). Video lectures with AI-generated instructors: Low video engagement, same performance as human instructors. *International Review of Research in Open and Distributed Learning, 25*(3), 350-369. https://doi.org/10.19173/irrodl.v25i3.7815

Azamatova, A., Bekeyeva, N., Zhaxylikova, K., Sarbassova, A., & Ilyassova, N. (2023). The effect of using artificial intelligence and digital learning tools based on project-based learning approach in foreign language teaching on students' success and motivation. *International Journal of Education in Mathematics, Science and Technology, 11*(6), 1458–1475. https://doi.org/10.46328/ijemst.3712

Baltezarević R. & Baltezarević I., (2025). Digital Game-Based Learning's (DGBL): Effect on students' academic performance. *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE), 13*(1), 127-140. https://doi.org/10.23947/2334-8496-2025-13-1-127-140

Basha Y. J. (2024). The negative impacts of AI tools on students in academic and real-life performance. *Int.J.Soci.Sci. 1*(3):1-16. Retrieved from https://doi.org/10.51470/IJSSC.2024.01.03.01

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*. Retrieved from https:// www.diva- portal. org/ smash/ get/ diva2: 18205 93/ FULLT EXT01. pdf

Dai, Y., Lin, Z., Liu, A. & Wang, W. (2023). An embodied, analogical and disruptive approach of AI pedagogy in upper elementary education: An experimental study. *British Journal of Educational Technology, 55*(1), 417-434. https://doi.org/10.1111/bjet.13371

Dong, L., Tang, X., & Wang, X. (2025). Examining the effect of artificial intelligence in relation to students' academic achievement: A meta-analysis. *Computers and Education: Artificial Intelligence, 8*, 100400. https://doi.org/10.1016/j.caeai.2025.100400

Effiong, A. A., Jacob, K. J., Udoh, V. I., & Joe, E. F. (2024). Effect of EduGPT artificial intelligence on academic performance of educational technology students: A quasi-experimental study. *Prestige Journal of Education, 7*(2), 107-119. https://openaccessglobal.com/wp-content/uploads/2025/01/edu-gpt.pdf

Elsayed, A. M., & Hussein, R. A. (2024). Teacher support in AI-assisted exams: An experimental study to inspect the effects on demotivation, anxiety management in exams, L2 learning experience, and academic success. *Lang Test Asia 14*(53). *https://doi.org/10.1186/s40468-024-00328-7*

Eltahir, M. E., & Babiker, F. M. (2024). The influence of artificial intelligence tools on student performance in e-learning environments: Case study. *Electronic Journal of e-Learning, 9*(22). https://doi.org/10.34190/ejel.22.9.3639

Essel, H.B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E., & Baah. P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *Int J Educ Technol High Educ* 19, 57. https://doi.org/10.1186/s41239-022-00362-6

Fu, Y., Weng, Z. & Wang, J. (2024). Examining AI Use in Educational Contexts: A Scoping Meta-Review and Bibliometric Analysis. *Int J Artif Intell Educ*. https://doi.org/10.1007/s40593-024-00442-w.

Gürefe, N., Akta, G. S. & Oksuz. H. (2024). Investigating the impact of the AI-supported 5E (AI-s5E) instructional model on spatial ability. *Behavioral Sciences, 14*(8). 10.3390/bs14080682

Han, J.-W., Park, J., & Lee, H. (2022). Analysis of the effect of an artificial intelligence chatbot educational program on non-face-to-face classes: A quasi-experimental study. *BMC Medical Education, 22*(1), 830. https://doi.org/10.1186/s12909-022-03898-3

Hwang, S. (2022). Examining the effects of artificial intelligence on elementary students' mathematics achievement: A meta-analysis. *Sustainability 14*, 13185.

Klarin, J., Hoff, E., T., Larsson, A. & Daukantaitė, D. (2024). Adolescents' use and perceived usefulness of generative AI for schoolwork: Exploring their relationships with executive functioning and academic achievement. *Frontiers in Artificial Intelligence, 7.* https://doi.org/10.3389/frai.2024.1415782

Klimova B and Pikhart M (2025) Exploring the effects of artificial intelligence on student and academic well-being in higher education: a mini-review. *Front. Psychol.* 16:1498132. Retrieved from doi: 10.3389/fpsyg.2025.1498132

Koć-Januchta, M. M., Schönborn, K. J., Tibell, L. A. E., Chaudhri, V. K., & Heller, H. C. (2020). Engaging with biology by asking questions: Investigating students' interaction and learning with an artificial intelligence-enriched textbook. *Journal of Educational Computing Research*, *58*(6), 1190-1224.

Liao, X., Zhang, X., Wang, Z., & Luo, H. (2024). Design and implementation of an AI-enabled visual report tool as formative assessment to promote learning achievement and self-regulated learning: An experimental study. *British Journal of Educational Technology*, *55*, 1253–1276. https://doi.org/10.1111/bjet.13424

Menekse, M., Putra, A. S., Kim, J., Butt, A. A., McDaniel, M. A., Davidesco, I., Cadieux, M., Kim, J. & Litman, D. (2025). Enhancing student reflections with natural language processing-based scaffolding: A quasi-experimental study in a large lecture course. *Computers and Education: Artificial Intelligence, 8*, 100397. https://doi.org/10.1016/j.caeai.2025.100397

Nelson, M. S., Wooditch, Al, & Dario, L. M. (2015). Sample size, effect size, and statistical power: A replication study of Weisburd's paradox. *Journal of Experimental Criminology, 11*, 141-163.

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041.

Núñez-Canal, M., De Obesso Arias, M. D., Pérez-Rivero, C. A., & Álvarez-de-Mon, I. (2026). Stakeholder engagement and strategic innovation in higher education through AI competency. *Journal of Innovation & Knowledge*, *13*, 100920.

Salas-Pilco, S. Z., Xiao, K., & Hu, X. (2022a). Artificial intelligence and learning analytics in teacher education: A systematic review. *Education Sciences*, *12*(8), 569.

Setiawan, R. Farisiyah, U. & Abidin, M. Z. (2025). Harnessing AI-based learning media in education: A meta-analysis of its effects on student achievement. *Participatory Educational Research (PER) 12*(1), 222-242.

Shahzad, M. F., Xu, S., Lim, W. M., Yang, X. & Khan, Q. R. (2024). Artificial intelligence and social media on academic performance and mental well-being: Student perceptions of positive impact in the age of smart learning. *Heliyon 10*(8), 29523.

Sharadgah, T. A., & Sa' di, R. A. (2022). A systematic review of research on the use of artificial intelligence in English language teaching and learning (2015‑2021): What are the current effects? *Journal of Information Technology Education: Research*, *21*.

Slavin, R. & Smith, D. (2009). The relationship between sample sizes and effect sizes in systematic reviews in education. *Educational Evaluation and Policy Analysis, 31*, 500-506.

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology 14*. https://doi.org/10.3389/fpsyg.2023.1260843

Sun, L. & Zhou, L. (2024). Does generative artificial intelligence improve academic achievement of college students? A meta-analysis. *Journal of Educational Computing Research 62*(7), 1896-1933.

Tong, D. H. T., Bui, P. U., & Lu, K. N. (2022). The effectiveness of blended learning on students' academic achievement, self-study skills and learning attitudes: A quasi-experiment study in teaching the conventions for coordinates in the plane. *Heliyon, 8*(12), e12657. https://doi.org/10.1016/j.heliyon.2022.e12657

Vieriu, A. M. & Petrea, G. (2025). The Impact of Artificial Intelligence (AI) on students' academic development. *Educ. Sci., 15*(3), 343.

Yang, T. C., Hsu, Y. C., & Wu, J. Y. (2025). The effectiveness of ChatGPT in assisting high school students in programming learning: evidence from a quasi-experimental research. *Interactive Learning Environments*, 1–18. https://doi.org/10.1080/10494820.2025.2450659

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39.

Zhang, Y. Lucas, M. Bem-haja, P. Pedro, L. (2024). The effect of student acceptance on learning outcomes: AI-generated short videos versus paper materials. *Computers and Education: Artificial Intelligence 7*, 100286.

Zhang, Y., Lai, X., Yi, S. & Lu, Y. (2025). Does ChatGPT-based reading platform impact foreign language paper reading? Evidence from a quasi-experimental study on Chinese undergraduate students. *Education and Information Technologies,* 30: 9737–9754 https://doi.org/10.1007/s10639-024-13190-0

Zheng, L., Niu, J., Zhong, L., Gyasi, J. F. (2021). The effectiveness of artificial intelligence on learning achievement and learning perception: A meta-analysis. *Interact. Learn. Environ*, 1–15.