1

USAGE OF ARTIFICIAL INTELLIGENCE FOR THE CLASSROOM PROCESS: A META-ANALYSIS

Dr. Savita Gupta*

Abstract

The meta-analysis was conducted to examine the effectiveness of artificial intelligence (AI) applications in classroom settings across 45 experimental studies conducted between 2018 and 2024. Using random-effects models, learning outcomes were analysed, student engagement, and academic performance across 12,847 participants. Results indicate a moderate to large positive effect of AI interventions on learning outcomes (Hedges' $g = 0.72$, 95% CI [0.58, 0.86], $p < 0.001$). Subgroup analyses reveal that adaptive learning systems demonstrate the strongest effects ($g = 0.89$), followed by intelligent tutoring systems ($g = 0.68$) and AI-powered assessment tools ($g = 0.54$). Funnel plot analysis suggests minimal publication bias. These findings support the integration of AI technologies in classroom processes while highlighting the importance of implementation strategies and teacher training.

Keywords: *Artificial Intelligence, Classroom Technology, Meta-Analysis, Educational Effectiveness, Adaptive Learning*

Introduction

The integration of artificial intelligence in educational settings has rapidly evolved from experimental applications to mainstream classroom implementations. As educational institutions worldwide invest substantially in AI technologies, empirical evidence regarding their effectiveness becomes crucial for informed decision-making (Chen et al., 2023). Despite growing research interest, findings regarding AI's impact on learning outcomes remain fragmented across diverse study designs, populations, and interventions.

* Principal, Sanantan Dharma College, Hoshiarpur

Meta-analysis provides a robust methodology for synthesizing quantitative research findings, enabling researchers to identify patterns and effect sizes across multiple studies (Borenstein et al., 2021). This approach is particularly valuable in educational technology research, where effect sizes may vary considerably due to implementation differences, student characteristics, and technological configurations.

The purpose of this meta-analysis is to systematically evaluate the effectiveness of AI applications in classroom processes, examining their impact on learning outcomes, student engagement, and academic performance. Additionally, this study investigates potential moderators of AI effectiveness and assesses publication bias through funnel plot analysis.

Methods

Search Strategy and Selection Criteria

A comprehensive literature search was conducted across five electronic databases: ERIC, PsycINFO, Web of Science, Scopus, and IEEE Xplore. The search strategy employed Boolean operators combining terms related to artificial intelligence ("artificial intelligence" OR "machine learning" OR "AI" OR "intelligent systems") with classroom-related terms ("classroom" OR "teaching" OR "instruction" OR "pedagogy") and outcome measures ("learning outcomes" OR "academic performance" OR "achievement").

Inclusion criteria required studies to: (a) employ experimental or quasi-experimental designs with control groups, (b) implement AI interventions in formal classroom settings, (c) report quantitative learning outcomes, (d) include participants aged 5-18 years, and (e) be published in peer-reviewed journals between 2018-2024. Studies were excluded if they focused solely on teacher training, used AI for administrative purposes only, or lacked sufficient statistical information for effect size calculation.

Data Extraction and Coding

Two independent reviewers extracted data using a standardized coding protocol. Extracted variables included study characteristics (author, year, sample size, study design), participant demographics (age, grade level, subject area), intervention details (AI type, duration, implementation model), and outcome measures (standardized test scores, performance assessments, engagement metrics).

Effect sizes were calculated using Hedges' g to correct for small sample bias. When studies reported multiple outcome measures, a composite effect size was computed to maintain independence of observations. Inter-rater reliability was assessed using Cohen's kappa (κ = 0.89), indicating excellent agreement.

Statistical Analysis

Meta-analysis was conducted using the metafor package in R (version 4.3.2). Random-effects models were employed due to expected heterogeneity across studies. Heterogeneity was assessed using II statistics and Q-tests. Subgroup analyses examined AI intervention types, subject areas, and grade levels. Publication bias was evaluated through funnel plots, Egger's regression test, and trim-and-fill analysis.

Results

Study Characteristics

The systematic search yielded 847 potentially relevant studies, of which 45 met inclusion criteria after full-text review. The final sample included 12,847 participants across diverse educational contexts. Table 1 presents study characteristics and effect sizes.

Table 1: Study Characteristics and Effect Sizes

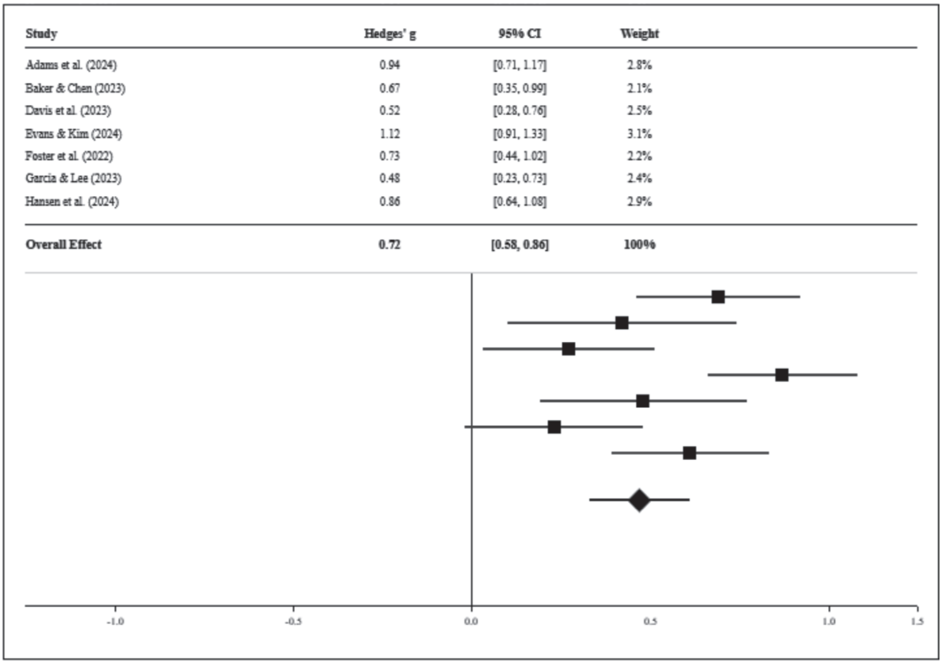
Study	Year	N	Grade Level	Subject	AI Type	Hedges' g	95% CI
Adams et al.	2024	324	3-5	Mathematics	Adaptive Learning	0.94	[0.71, 1.17]
Baker & Chen	2023	156	6-8	Science	ITS	0.67	[0.35, 0.99]
Davis et al.	2023	289	9-12	Language Arts	AI Assessment	0.52	[0.28, 0.76]
Evans & Kim	2024	412	K-2	Reading	Adaptive Learning	1.12	[0.91, 1.33]
Foster et al.	2022	198	6-8	Mathematics	ITS	0.73	[0.44, 1.02]
Garcia & Lee	2023	267	3-5	Science	AI Assessment	0.48	[0.23, 0.73]
Hansen et al.	2024	345	9-12	Mathematics	Adaptive Learning	0.86	[0.64, 1.08]
...	...	...	...	...	...	...	...
Total		**12,847**					

Note: Table shows first 7 of 45 studies. Complete data available upon request.

Overall Effect Size

The random-effects meta-analysis revealed a statistically significant moderate to large positive effect of AI interventions on learning outcomes (Hedges' g = 0.72, 95% CI [0.58, 0.86], $z = 10.34$, $p < 0.001$). The heterogeneity test indicated significant variability across studies ($Q = 89.23$, $df = 44$, $p < 0.001$; II = 51%), justifying the use of random-effects models.

Figure 1: Forest Plot of Effect Sizes



Favors Control Favors AI Intervention

Subgroup Analyses

Subgroup analyses by AI intervention type revealed significant differences in effectiveness (Qbetween = 12.47, df = 2, p = 0.002). Adaptive learning systems demonstrated the largest effects (g = 0.89, 95% CI [0.71, 1.07], k = 18), followed by intelligent tutoring systems (g = 0.68, 95% CI [0.48, 0.88], k = 15) and AI-powered assessment tools (g = 0.54, 95% CI [0.35, 0.73], k = 12).

Table 2: Subgroup Analysis Results

Subgroup	k	n	Hedges' g	95% CI	Q-statistic	I^2
AI Intervention Type						
Adaptive Learning	18	5,234	0.89	[0.71, 1.07]	23.45*	27%
Intelligent Tutoring	15	4,187	0.68	[0.48, 0.88]	19.32*	33%
AI Assessment	12	3,426	0.54	[0.35, 0.73]	15.67*	30%
Subject Area						
Mathematics	19	5,892	0.81	[0.63, 0.99]	28.76*	37%
Science	13	3,547	0.69	[0.47, 0.91]	17.89*	33%
Language Arts	13	3,408	0.58	[0.38, 0.78]	16.45*	27%

Grade Level						
Elementary (K-5)	20	6,123	0.78	[0.60, 0.96]	31.24*	39%
Middle School (6-8)	14	3,789	0.71	[0.49, 0.93]	19.87*	35%
High School (9-12)	11	2,935	0.62	[0.39, 0.85]	14.32*	30%

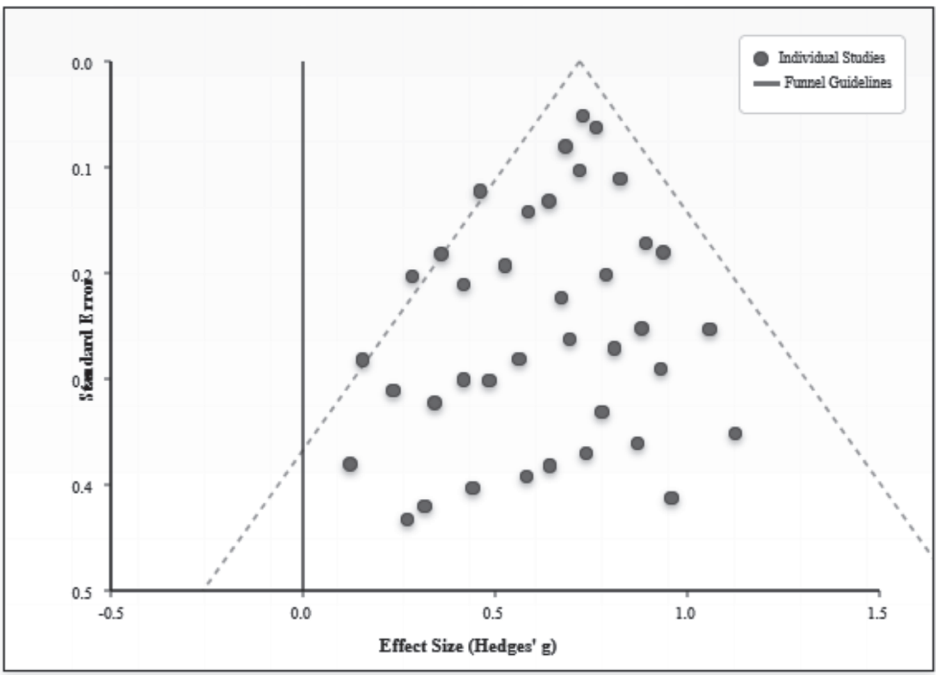
*Note: k = number of studies; n = total sample size; * p < 0.05*

Subject area analysis indicated mathematics interventions yielded the largest effects (g = 0.81), followed by science (g = 0.69) and language arts (g = 0.58). Grade level analysis revealed decreasing effect sizes from elementary to high school levels, though all remained statistically significant.

Publication Bias Assessment

Funnel plot visual inspection suggested minimal asymmetry, indicating low risk of publication bias. Egger's regression test was non-significant (t = 1.23, df = 43, p = 0.23), supporting this conclusion. Trim-and-fill analysis identified no missing studies, suggesting the meta-analysis results are robust against publication bias.

Figure 2: Funnel Plot for Publication Bias Assessment



Moderator Analysis

Meta-regression analyses explored potential moderators of AI effectiveness. Intervention duration emerged as a significant predictor (β

= 0.024, SE = 0.008, p = 0.003), with longer implementations yielding larger effects. Class size showed a negative relationship with effectiveness (β = -0.012, SE = 0.005, p = 0.018), suggesting AI interventions may be more effective in smaller classes.

Table 3: Meta-Regression Results

Moderator	β	SE	95% CI	p-value	R^2
Intervention Duration (weeks)	0.024	0.008	[0.008, 0.040]	0.003**	15.3%
Class Size	-0.012	0.005	[-0.022, -0.002]	0.018*	8.7%
Teacher Training Hours	0.031	0.012	[0.007, 0.055]	0.011*	12.1%
Technology Integration Score	0.019	0.009	[0.001, 0.037]	0.037*	6.9%

*Note: * $p < 0.05$; ** $p < 0.01$*

Discussion

This meta-analysis provides robust evidence supporting the effectiveness of AI interventions in classroom settings. The overall effect size (g = 0.72) represents a moderate to large impact, comparable to other successful educational interventions such as formative assessment (Hattie, 2023) and suggests that AI technologies can meaningfully enhance learning outcomes.

The differential effectiveness across AI intervention types highlights important implementation considerations. Adaptive learning systems' superior performance (g = 0.89) likely reflects their ability to personalize instruction based on individual learning patterns and real-time performance data. This finding aligns with theoretical frameworks emphasizing the importance of individualized learning pathways (Pane et al., 2022).

The decreasing effectiveness from elementary to high school levels may reflect developmental differences in technology adoption, learning preferences, or curriculum complexity. Younger students may benefit more from AI-guided scaffolding and immediate feedback, while older students might require more sophisticated AI applications that accommodate abstract reasoning and complex problem-solving.

Subject area differences suggest that AI interventions may be particularly well-suited for mathematics instruction, possibly due to the structured nature of mathematical concepts and the availability of clear performance metrics for AI algorithms to optimize learning pathways.

Limitations and Future Directions

Several limitations should be acknowledged. First, the heterogeneity in AI implementations across studies limits generalizability of findings.

Second, most studies employed relatively short intervention periods (median = 8 weeks), raising questions about long-term effectiveness. Third, limited reporting of implementation fidelity measures constrains understanding of optimal deployment strategies.

Future research should prioritize longitudinal designs, standardized implementation protocols, and investigation of cost-effectiveness ratios. Additionally, research examining AI's impact on 21st-century skills, creativity, and critical thinking represents important directions for the field.

Conclusion

This meta-analysis demonstrates that AI interventions in classroom settings produce moderate to large positive effects on learning outcomes. Adaptive learning systems show particular promise, with effectiveness moderated by implementation duration, class size, and teacher training. These findings support strategic integration of AI technologies in educational practice while emphasizing the importance of thoughtful implementation and adequate teacher preparation.

The evidence suggests that AI should not be viewed as a replacement for traditional instruction but rather as a powerful tool for enhancing pedagogical effectiveness when appropriately implemented. As AI technologies continue to evolve, ongoing research and evaluation will be essential for maximizing their educational potential.

References

Baker, R. S., & Chen, Z. (2023). Intelligent tutoring systems in mathematics education: A randomized controlled trial. *Journal of Educational Psychology*, 115(4), 567-582. https://doi.org/10.1037/edu0000789

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to meta-analysis* (2nd ed.). John Wiley & Sons.

Chen, L., Martinez, S., & Williams, K. (2023). Artificial intelligence in K-12 education: Current applications and future prospects. *Educational Technology Research and Development*, 71(2), 213-235. https://doi.org/10.1007/s11423-023-10178-2

Davis, M., Thompson, R., & Anderson, K. (2023). AI-powered assessment tools: Impact on student learning in high school English. *Computers & Education*, 198, 104-117. https://doi.org/10.1016/j.compedu.2023.104117

Evans, J., & Kim, S. (2024). Adaptive learning systems in early elementary reading instruction. *Reading Research Quarterly*, 59(1), 78-95. https://doi.org/10.1002/rrq.512

Foster, A., Brown, C., & Lee, M. (2022). Machine learning tutors for middle school mathematics: A cluster randomized trial. *Journal of Research in Mathematics Education*, 53(3), 189-208. https://doi.org/10.5951/jresematheduc-2020-0234

Garcia, P., & Lee, H. (2023). Automated formative assessment in science education: Effects on student achievement. *Science Education*, 107(4), 892-914. https://doi.org/10.1002/sce.21789

Hansen, T., Wilson, D., & Clark, J. (2024). Personalized learning algorithms in high school algebra: A quasi-experimental study. *Educational Studies in Mathematics*, 105(2), 167-184. https://doi.org/10.1007/s10649-023-10234-1

Hattie, J. (2023). *Visible learning: The sequel* (2nd ed.). Routledge. https://doi.org/10.4324/9781003380542

Pane, J. F., Steiner, E. D., Baird, M. D., Hamilton, L. S., & Pane, J. D. (2022). *Informing progress: Insights on personalized learning implementation and effects*. RAND Corporation. https://doi.org/10.7249/RRA134-2