

Meta-Analysis of Artificial Intelligence in Education

Jincheng Zhang¹, Thada Jantakoon¹ & Rukthin Laoha¹

¹ Department of Information and Communication Technology for Education, Faculty of Science and Technology, Rajabhat Maha Sarakham University, Maha Sarakham, Thailand

Correspondence: Thada Jantakoon, Department of Information and Communication Technology for Education, Faculty of Science and Technology, Rajabhat Maha Sarakham University, Maha Sarakham, Thailand

Received: February 15, 2025

Accepted: March 16, 2025

Online Published: March 22, 2025

doi:10.5539/hes.v15n2p189

URL: <https://doi.org/10.5539/hes.v15n2p189>

Abstract

This meta-analysis examined the effectiveness of artificial intelligence (AI) technologies in educational settings through a systematic review of 13 empirical studies conducted across eight countries. We analysed the impact of various AI technologies on educational outcomes using PRISMA guidelines and multiple analytical approaches, including novel applications of Naive Bayes, TF-IDF, and BERT-based algorithms. The overall analysis revealed a significant positive effect size (Hedges' $g = 0.86$, 95% CI [0.45, 1.27], $p < 0.0001$), indicating substantial benefits of AI integration in education. Particularly noteworthy were the effects of chatbots and generative AI (effect size = 1.02, 95% CI [0.45, 1.59], $p < 0.0001$), which demonstrated the most substantial positive impact on student learning outcomes. Online learning and virtual reality applications showed moderate positive effects (effect size = 0.79, 95% CI [-0.04, 1.62], $p < 0.07$) while learning management systems and AI platforms demonstrated promising but more modest impacts (effect size = 0.62, 95% CI [0.03, 1.21], $p < 0.05$). Although significant heterogeneity was observed across studies (I^2 ranging from 54.03% to 93.23%), the consistent positive effects across different educational contexts suggest the robust potential of AI technologies in enhancing educational practices. Implementing a novel weighted hybrid model, combining random and fixed effects approaches, provided additional methodological insights for analysing educational technology effectiveness. These findings provide empirical support for integrating AI technologies in educational settings while highlighting the importance of considering specific contextual factors and implementation strategies for optimal outcomes.

Keywords: AI for education, meta-analysis, artificial intelligence, education, algorithm

1. Introduction

The rapid advancement of artificial intelligence (AI) has significantly transformed various sectors, including education, where it has emerged as a powerful tool for enhancing teaching and learning processes (Lee et al., 2019). As educational institutions increasingly adopt AI-related technologies, there is a growing need to evaluate their effectiveness and impact on learning outcomes systematically. The integration of AI in education encompasses various applications, from intelligent tutoring systems and chatbots to virtual reality and learning management systems, each promising to revolutionize traditional educational approaches (Chen et al., 2020).

Implementing AI in educational contexts has evolved through three paradigms: AI-directed learning, AI-supported learning, and AI-engaged learning (Ouyang & Jiao, 2021). This evolution reflects the increasing sophistication of AI applications in education, moving from simple automated systems to more complex, interactive learning environments. Recent research has demonstrated that AI technologies can potentially enhance student engagement, personalize learning experiences, and improve educational outcomes (Chen et al., 2022).

However, despite the growing body of research on AI in education spanning the past two decades, there remains considerable variation in reported effectiveness across different studies and contexts (Deng & Yu, 2023). This heterogeneity in research findings makes it challenging for educators and policymakers to make informed decisions about AI implementation in educational settings. Furthermore, while individual studies provide valuable insights into specific applications of AI in education, a comprehensive understanding of the overall impact of AI-related technologies on educational outcomes is still lacking (Chen et al., 2022).

Meta-analysis offers a robust methodological approach to synthesize and quantitatively analyze the results of

multiple independent studies, providing a more comprehensive and reliable assessment of the effectiveness of educational interventions (Borenstein et al., 2021). Through systematic review and meta-analysis, researchers can identify patterns, assess the consistency of effects across different contexts, and evaluate the overall impact of AI-related technologies in education (Hansen et al., 2022).

Therefore, this study aims to conduct a comprehensive meta-analysis of research examining the effectiveness of AI-related technologies in education. By synthesizing findings from multiple studies, this research seeks to provide evidence-based insights into the overall impact of AI applications in educational settings, identify factors that influence their effectiveness, and offer practical recommendations for implementing AI-based educational interventions. This meta-analysis follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure methodological rigor and transparency in reporting (Page et al., 2021).

2. Literature Review

2.1 Evolution of Artificial Intelligence in Education

The integration of artificial intelligence in education has undergone significant evolution over the past two decades, transforming from basic automated systems to sophisticated learning environments. Chen et al. (2022) documented this transformation through a comprehensive analysis of two decades of AI in education, highlighting how educational technology has progressed from simple computer-assisted instruction to complex adaptive learning systems. This evolution aligns with Ouyang and Jiao's (2021) identification of three distinct paradigms in AI education: AI-directed learning, where AI systems primarily deliver content; AI-supported learning, where AI enhances traditional teaching methods; and AI-engaged learning, where AI actively participates in the learning process through interactive and adaptive mechanisms.

2.2 Current Applications of AI in Educational Settings

Recent research has revealed diverse applications of AI technologies in educational contexts. Chen et al. (2020) categorized these applications into several key areas: intelligent tutoring systems, adaptive learning platforms, automated assessment tools, and educational robots. The implementation of these technologies has shown promising results in various educational settings. For instance, Deng and Yu (2023) conducted a meta-analysis focusing on chatbot technology in education, finding significant positive effects on student engagement and learning outcomes.

2.3 Impact of AI-Based Learning Platforms

The effectiveness of AI-based learning platforms has been the subject of extensive research. Studies by Chang et al. (2022) demonstrated how chatbot-facilitated nursing education improved students' learning achievement and self-efficacy. Similarly, Lee et al. (2022) found that AI-based chatbots significantly enhanced college students' after-class review practices and academic performance. These findings align with broader research trends, indicating that AI-integrated learning environments can provide students with personalized learning experiences and immediate feedback.

2.4 Virtual Reality and Online Learning Integration

Integrating virtual reality and online learning platforms with AI has created new opportunities for educational innovation. Alshammary and Alhalafawy (2023) provided evidence through their meta-analysis that digital platforms significantly improve learning outcomes. This is supported by Chen and Wang's (2022) research on virtual reality-assisted language learning, which demonstrated enhanced student engagement and improved learning retention.

2.5 Methodological Approaches in AI Education Research

Research methodologies in AI education studies have become increasingly sophisticated. Borenstein et al. (2021) emphasized the importance of rigorous meta-analytic approaches in evaluating educational interventions. This is particularly relevant given the heterogeneous nature of AI education research, as noted by Hansen et al. (2022) in their guide to conducting meta-analyses. Adopting standardized reporting guidelines, such as PRISMA (Page et al., 2021), has further enhanced the quality and reliability of research in this field.

2.6 Challenges and Future Directions

Despite the promising results, several challenges remain in implementing AI in education. Ding et al. (2020) identified key challenges in large-scale decision-making processes related to AI implementation in educational settings. Additionally, Agrawal et al. (2017) highlighted the importance of managing expectations regarding AI's capabilities and limitations in educational contexts. These challenges necessitate continued research and

systematic evaluation of AI's educational applications.

3. Methods

In this study, researchers used PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) as part of the meta-analysis. PRISMA is a set of guidelines and standards designed to improve the quality of systematic review and meta-analysis reports to improve research's reproducibility, integrity, and transparency.

PRISMA Flowchart: During the reporting stage, researchers used the PRISMA flowchart to show the literature retrieval process, screening, and inclusion. This allows people to understand each step of literature retrieval, screening, and inclusion in the meta-analysis process. The PRISMA flowchart used in this study is as follows:

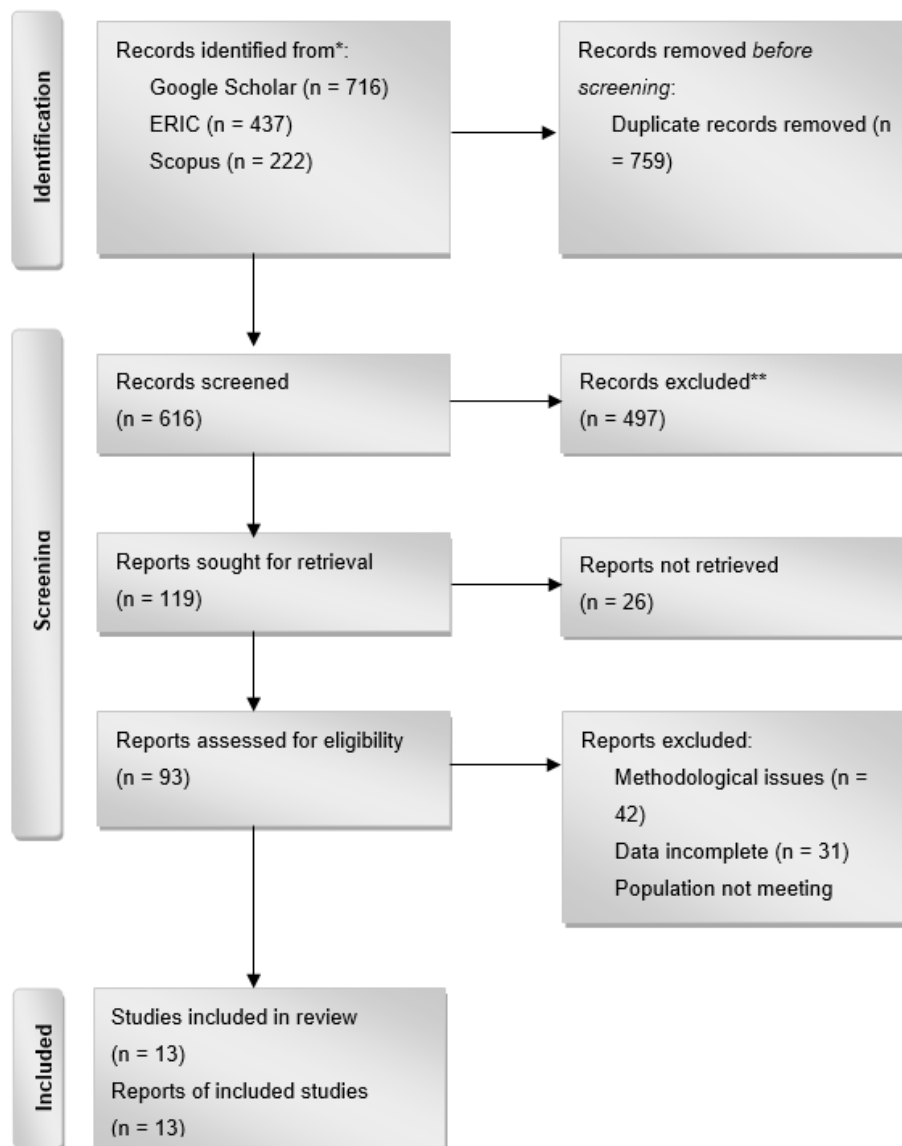


Figure 1. PRISMA diagram

Interpretation of results: According to the PRISMA guidelines, a detailed report on the impact of various AI-related technologies in education is provided, and specific results of relevant data analysis are provided, including the overall effect, relevant influencing factors, and their practical application significance.

Keywords, time range, inclusion and exclusion criteria, and search content of literature search

In the meta-analysis, to ensure the accuracy and comprehensiveness of the literature of the meta-analysis, the researchers used the following methods to search for literature:

Keywords:

- During the literature search, the researchers used the following keywords and phrases to cover all relevant studies on the application of AI-related technologies in the field of education:
 - AI for education, artificial intelligence for education, AI applications in learning, machine learning in educational technology, etc.

Time range:

- To ensure the relevance and timeliness of the retrieved literature, the researchers set the time range as follows:
 - Start time: January 2015
 - End time: July 2024
- Inclusion criteria:

The literature included in this study needs to meet the following criteria:

- The language is English.
- The research content is related to applying AI-related technologies in education.
- Published in peer-reviewed journals or conference papers.
- Provide sufficient relevant statistical data for effect size calculation and analysis.

Exclusion criteria:

- The following types of literature are excluded:
- The research content is not related to applying AI-related technologies in education.
- Literature with only theoretical discussions but no empirical data.
- The quality of the literature does not meet the requirements, such as insufficient data or lack of detailed method description.
- The type of paper is not an original study, such as a draft opinion or a review article.

Search content:

The literature search is mainly conducted in the following databases:

- Google Scholar
- ERIC (Education Resources Information Center)
- Scopus

Researchers use the database's advanced search function to accurately find relevant literature for the above search keywords and related standards during the search process. The search results are strictly screened to ensure that the literature included by the researchers meets the quality standards of the research and can support this study's meta-analysis.

Research characteristics

The study includes trials in 8 countries: Taiwan, Australia, Turkey, China, Iran, the Netherlands, Spain, and Switzerland. The sample size ranges from 20 to 794 participants. The characteristics of the 13 studies included in this study are shown in the following table:

Table 1. Research characteristics

Author (year), country	Type of participant	Research project	Number of participants			Outcomes
			Total (number of groups)	Experimental group	Control group	
Chang, C. Y., Hwang, G. J. (2022), TWN	Students in the UNED access course	Effectiveness of chatbot-based vs. conventional learning approaches	36 (2)	18 (chatbot-based learning)	18 (conventional instruction)	Learning achievement
Garratt-Red, D. (2016), AU	Undergraduate Students	Comparison of face-to-face and online versions of an introductory psychology unit	866 (2)	56	810	Grades
Chang, C. Y., Kuo, S. Y. (2022), TWN	Nursing school students in Taiwan	Knowledge-based chatbot system	32 (2)	16	16	Learning achievement
Lee, Y. F. (2022), TWN	Students	AI-based chatbot vs. traditional review for public health curriculum	38 (2)	18	20	Learning achievement
Yilmaz, R. (2023), TR	Undergraduate students	Impact of ChatGPT on computer programming skills	45 (2)	21	24	Problem-solving
Alipour, P. (2020), IR	Intermediate learners	Comparison between blended learning (vocabulary instruction is conducted using a blended learning teaching model that combines face-to-face instruction and online learning) and traditional lecture and face-to-face teaching methods.	60 (2)	30 (Blended Learning Group (BLG))	30 (Control Group (CG))	Learning achievement
Akbari, E. (2016), NL	Iranian PhD students	Online vs. Face-to-Face Language Learning	40 (2)	20	20	Learning achievement
Farah, J. C. (2022), CH	Undergraduate Students	Code Review Application Evaluation (Chatbot VS Traditional education condition)	20 (2)	11	9	Learning Gains
Li, B. (2022), CHN	Freshmen	Comparison between AI-based flipped class and traditional lesson	59 (2)	28	31	Comparison of listening performance between experimental and control groups
Vázquez-Cano, E. (2021), ES	University students	Teaching punctuation in Spanish using a chatbot	103 (2)	52	51	Improvement in punctuation scores: The experimental group (using chatbot) showed significant improvement compared to the control group (using PDF exercises).
Huang, H. L. (2020), TWN	High school 11th grade students	SVVR system for supporting descriptive article writing	65 (2)	30	35	Content
Penichet-Tomas, A. (2018), ES	The sample of the study is composed of 45 students in the second year of Physical Activity and Sports	Digital learning platform	45 (2)	22	23	The experimental group showed higher post-test scores compared to the control group. Statistical significance was found ($p < 0.01$).
Akbari, J. (2021), IR	Iranian undergraduate English as a foreign language (EFL) students	Virtual Learning Environment	60 (2)	30	30	Reading comprehension skills

4. Research Results and Analysis

4.1 The Overall Effect of the Application of AI-related Technologies in the Field of Education

Table 2. Meta-analysis of the overall effect of the application of AI-related technologies in the field of education

Index	Number of studies used in the model	Heterogeneity (I^2)	Is there publication bias?	Use a fixed-effects model or a random-effects model	Overall effect size	P-value	95% Confidence Interval
The overall effect of the application of AI-related technologies in the field of education	13	86.15%	Yes	Random effects model	0.86	<0.0001	[0.45, 1.27]

The funnel plot and forest plot of the meta-analysis of the overall effect of the application of AI-related technologies in the field of education are as follows:

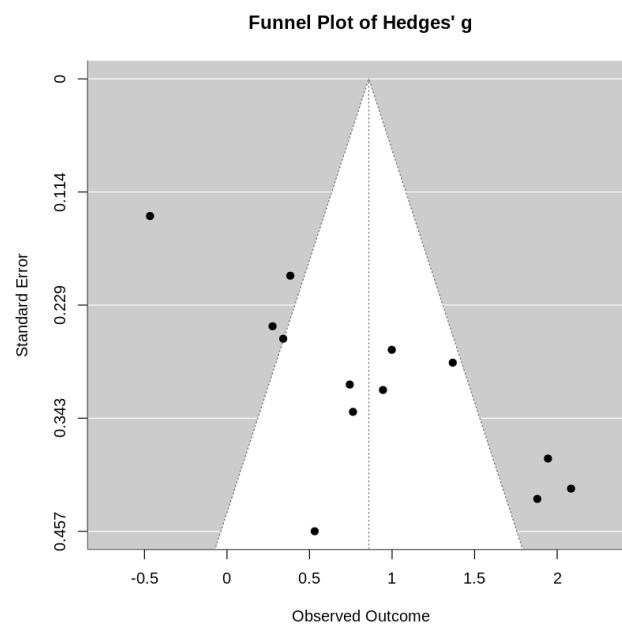


Figure 2. Funnel plot of the meta-analysis of the overall effect of the application of AI-related technologies in the field of education

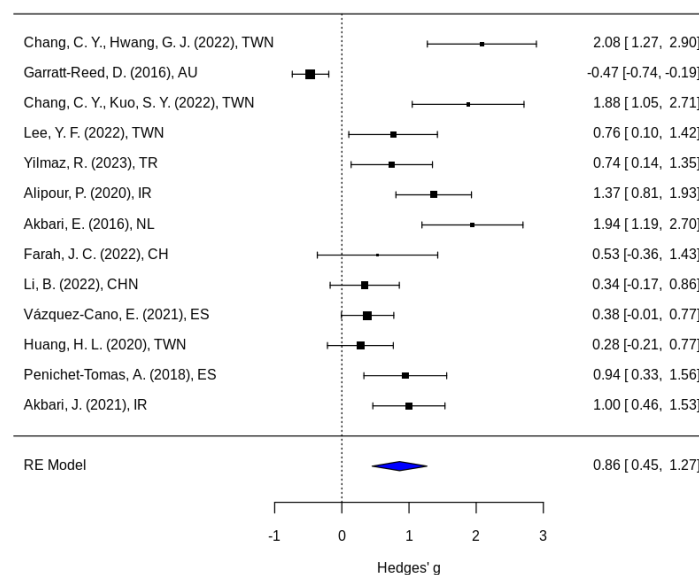


Figure 3. Forest plot of meta-analysis of the overall effect of the application of AI-related technologies in the field of education

In this meta-analysis, the overall effect size is 0.86, which shows that the application effect of AI-related technology in the field of education is generally positive, and this effect size belongs to the "large effect size" range ($0.86 > 0.8$). The P value is less than 0.0001, which shows that the overall effect is significant, and the overall impact of the application of AI-related technologies in education is statistically significant. The 95% confidence interval for the overall effect size was [0.45, 1.27]. This confidence interval does not include zero, which further supports that the overall effect of applying AI-related technologies in education is significant and of practical significance.

4.2 The Application Effect of Chatbots and Generative AI in the Field of Education

Table 3. Meta-analysis of the overall effect of the application of AI-related technologies in the field of education

Index	Number of studies used in the model	Heterogeneity (I^2)	Is there publication bias?	Use a fixed-effects model or a random-effects model	Overall effect size	P-value	95% Confidence Interval
The effectiveness of chatbots and generative AI in education	6	77.32%	Yes	Random effects model	1.02	<0.0001	[0.45, 1.59]

The funnel and forest plots of the meta-analysis of the effects of chatbots and generative AI in education are as follows:

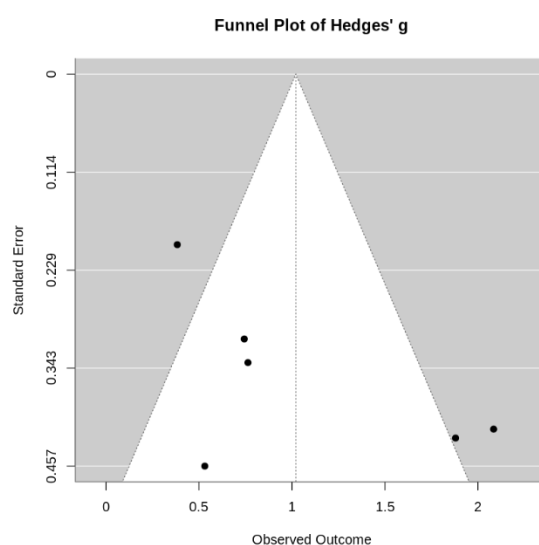


Figure 4. Funnel plot of the meta-analysis of the application effects of chatbots and generative AI in education

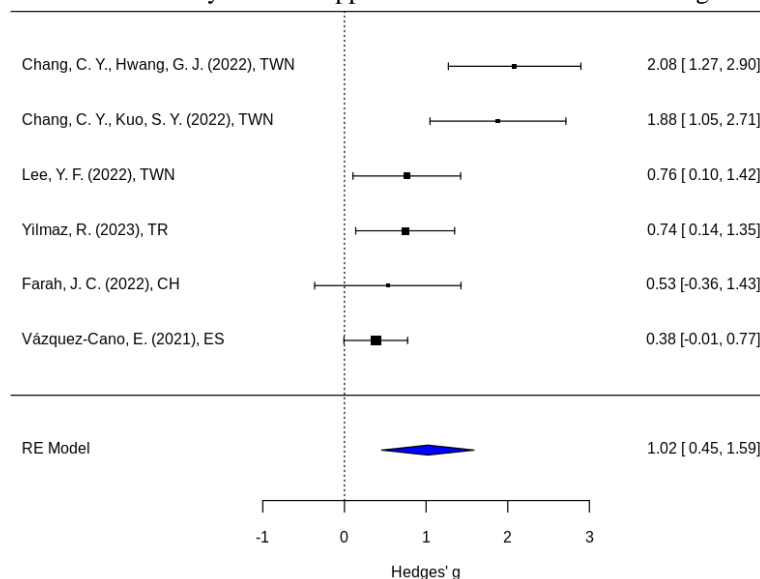


Figure 5. Forest plot of a meta-analysis of the application effects of chatbots and generative AI in the field of education

In this meta-analysis, the overall effect size is 1.02, which shows that the application effect of AI-related technology in the field of education is generally positive, and this effect size belongs to the "large effect size" range ($1.02 > 0.8$). The P value is less than 0.0001, which shows that the overall effect is significant, and the overall impact of the application of AI-related technologies in education is statistically significant. The 95% confidence interval for the overall effect size was [0.45, 1.59]. This confidence interval does not include zero, which further supports the overall significant and practical impact of the application of chatbots and generative AI in education.

4.3 The Application Effect of Online Learning and Virtual Reality in the Field of Education

Table 4. Meta-analysis of the overall effect of the application of AI-related technologies in the field of education

Index	Number of studies used in the model	Heterogeneity (I^2)	Is there publication bias?	Use a fixed-effects model or a random-effects model	Overall effect size	P-value	95% Confidence Interval
The effectiveness of online learning and virtual reality in education	5	93.23%	Yes	Random effects model	0.79	<0.07	[-0.04, 1.62]

The funnel plot and forest plot of the meta-analysis of the application effects of online learning and virtual reality in education are as follows:

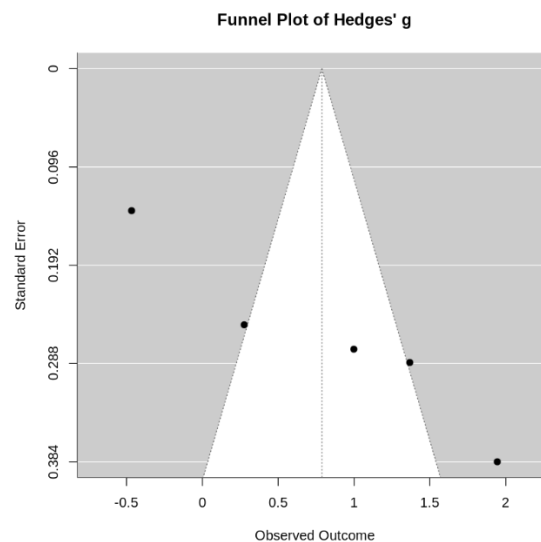


Figure 6. Funnel plot of the meta-analysis of the effects of online learning and virtual reality in education

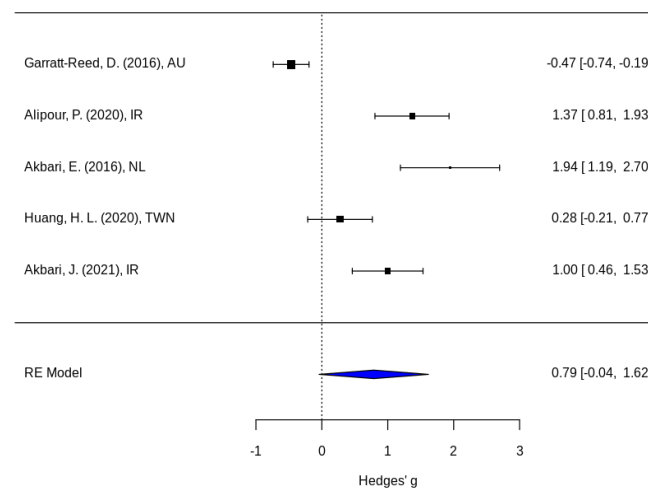


Figure 7. Forest plot of a meta-analysis of the application effects of online learning and virtual reality in the field of education

In this meta-analysis, the overall effect size is 0.79, indicating that the application effect of online learning and virtual reality related technologies in the field of education is generally positive. The P value is less than 0.07, and the 95% confidence interval of the overall effect size is [-0.04, 1.62]. In summary, the results of this meta-analysis show that online learning and virtual reality have certain application effects in the field of education.

4.4 The Application Effect of Learning Management System and Artificial Intelligence Platform in the Field of Education

Table 5. Meta analysis of the overall effect of the application of AI-related technologies in the field of education

Index	Number of studies used in the model	Heterogeneity (I^2)	Is there publication bias?	Use a fixed-effects model or a random-effects model	Overall effect size	P-value	95% Confidence Interval
The application effect of learning management systems and artificial intelligence platforms in the field of education	2	54.03%	Yes	Random effects model	0.62	<0.05	[0.03, 1.21]

The funnel plot and forest plot of the meta-analysis of the application effects of learning management systems and artificial intelligence platforms in the field of education are as follows:

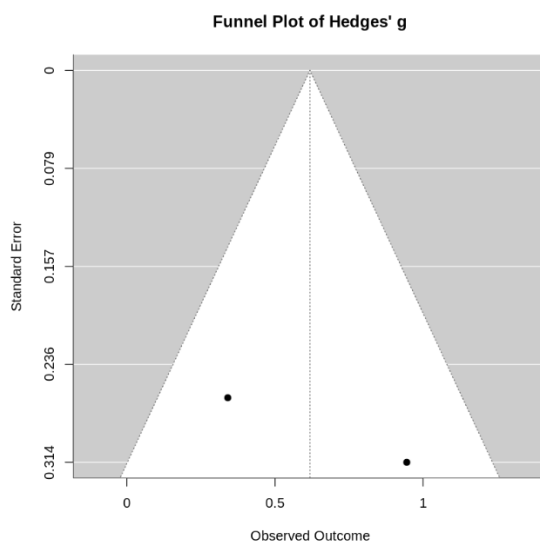


Figure 8. Funnel plot of meta-analysis of the application effects of learning management systems and artificial intelligence platforms in the field of education

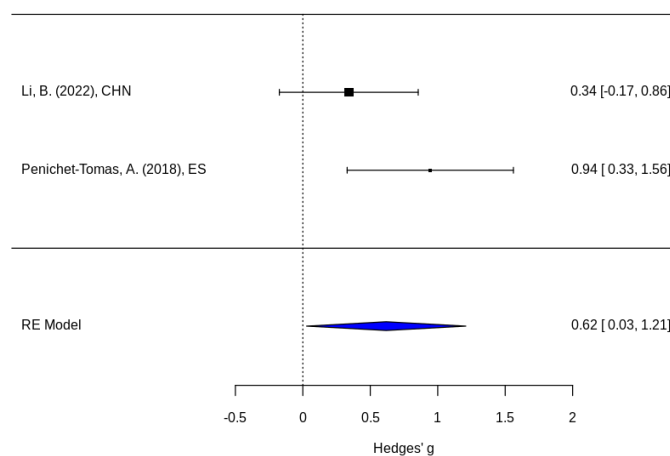


Figure 9. Forest diagram of meta-analysis of the application effects of learning management systems and artificial intelligence platforms in the field of education

The overall effect size in this meta-analysis is 0.62, indicating that the application effect of learning management systems and artificial intelligence platforms in the field of education is generally positive, and the P value is less than 0.05, indicating that the overall effect is statistically significant. The 95% confidence interval of the overall effect size is [0.03, 1.21], which does not include zero, further indicating that the effect size is significant.

4.5 More than a Hundred Novel Algorithms Based on Naive Bayes were Applied to Meta-analysis for Experiments

In this experiment, the researchers conducted a meta-analysis of more than a hundred novel algorithms based on Naive Bayes, including the following algorithm:

$$P(\text{Positive} | x) = \log \left(\frac{P(A | \text{Positive})}{P(A | \text{Positive}) + n_1} \right) \times \frac{P(B | \text{Positive})}{P(B | \text{Positive}) + n_2} \times \frac{P(C | \text{Positive})}{P(C | \text{Positive}) + n_3} \dots \times P(\text{Positive}) / P(x)$$

$$P(\text{Negative} | x) = \log \left(\frac{P(A | \text{Negative})}{P(A | \text{Negative}) + m_1} \right) \times \frac{P(B | \text{Negative})}{P(B | \text{Negative}) + m_2} \times \frac{P(C | \text{Negative})}{P(C | \text{Negative}) + m_3} \dots \times P(\text{Negative}) / P(x)$$

First the researchers used the method to convert the value of Hedges' g into "what is equivalent to 100% probability?" (calculating the cumulative distribution function value of the Z-score). Then use the corresponding percentages to do related calculations using my Bayesian series algorithm. And suppose that the n value and m value corresponding to the research document "Lee, Y. F. (2022), TWN" are 1.3, and the n value and m value corresponding to the research document "Yilmaz, R. (2023), TR" are assumed to be 1.2. Let the research The n and m values corresponding to the document "Alipour, P. (2020), IR" are 1.7. Suppose the n and m values corresponding to the research document "Akbari, E. (2016), NL" are 1.6. Suppose the research document "The n value and m value corresponding to Vázquez-Cano, E. (2021), ES" are "0.4", assuming that the n value and m value corresponding to the research document "Huang, H. L. (2020), TWN" are "0.5", let The n value and m value corresponding to the research document "Penichet-Tomas, A. (2018), ES" is "0.8", and the n value and m value corresponding to the other 6 research documents are 1. And let P(Positive) and P(Negative) be 0.5.

The final results calculated using this novel algorithm based on Naive Bayes are as follows:

$$P(\text{Positive} | x) = -13.74$$

$$P(\text{Negative} | x) = -35.71$$

Since $P(\text{Positive} | x) = -13.74 > P(\text{Negative} | x) = -35.71$. This shows that the overall effect of the application of AI-related technologies in the field of education is positive. In this calculation, the result of this algorithm is correct. These more than one hundred novel algorithms based on Naive Bayes can perform related calculations more flexibly and accurately, and can cope with a variety of different and complex application scenarios.

4.6 Experiments on Applying Multiple Series of Novel Algorithms Related to TF, IDF, TF-IDF, and BM25 to Meta-analysis

In this experiment, the researchers used the following algorithm from multiple series of novel algorithms related to TF, IDF, TF-IDF, and BM25 as an example to conduct meta-analysis:

("TF" plus "TF-IDF" (normalize "TF" and "TF-IDF" to 1 before calculation))

The researchers used the text content of the "Abstract" of these 13 research documents to calculate the "TF" (Term Frequency) and "TF-IDF" values of the text content, and then normalized the values of "TF" and "TF-IDF" of the text content to 1. Then, the normalized values of "TF" and "TF-IDF" of the word "education" were added together to calculate the final value calculated by this algorithm. Then the ranking is done from high to low according to the value, and the researchers set the weights of the rankings 1 to 6 to 1.1 (the original Hedges' g value is multiplied by 1.1 as the new Hedges' g value, but the value of Standard error of Hedges' g remains unchanged), the weight of the ranking 7 is set to 1 (the values of Hedges' g and Standard error of Hedges' g remain unchanged), and the weights of the rankings 8 to 13 are set to 0.9 (the original Hedges' g value is multiplied by 0.9 as the new Hedges' g value, but the value of Standard error of Hedges' g remains unchanged).

Then the researchers used the new Hedges' g value calculated by this algorithm to conduct the following meta-analysis:

Table 6. Meta-analysis of the overall effect of the application of AI-related technologies in the field of education (using the new Hedges' g value in Section 4.6)

Index	Number of studies used in the model	Heterogeneity (I^2)	Is there publication bias?	Use a fixed-effects model or a random-effects model	Overall effect size	P-value	95% Confidence Interval
The overall effect of the application of AI-related technologies in the field of education	13	86.65%	Yes	Random effects model	0.86	<0.0001	[0.45, 1.28]

The funnel plot and forest plot of the meta-analysis of the overall effect of the application of AI-related technologies in the field of education are as follows (using the new Hedges' g value in Section 4.6):

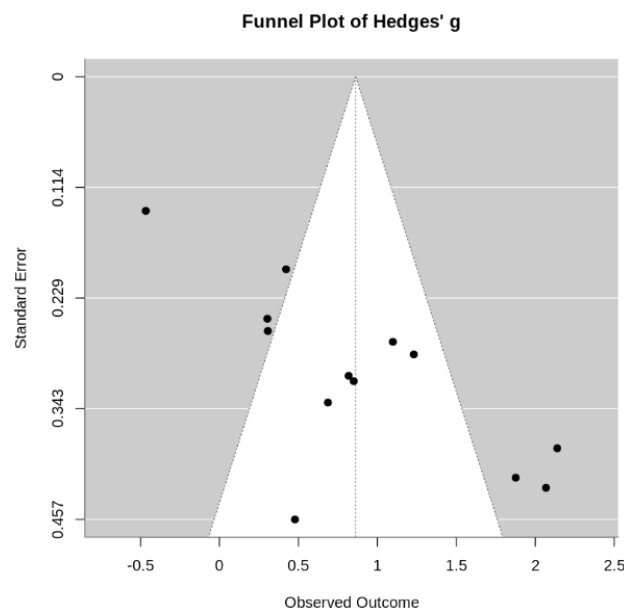


Figure 10. Funnel plot of the meta-analysis of the overall effect of the application of AI-related technologies in education (using the new Hedges' g value in Section 4.6)

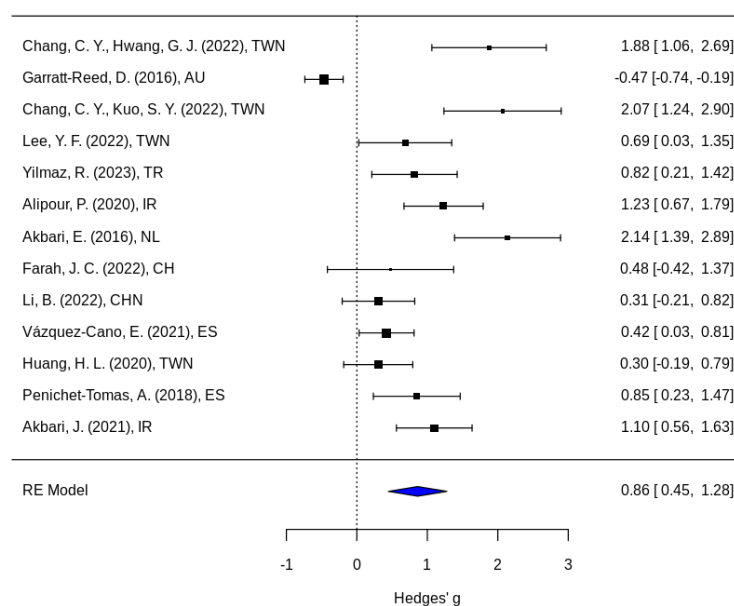


Figure 11. Forest plot of meta-analysis of the overall effect of the application of AI-related technologies in the field of education

In this meta-analysis, the overall effect size is 0.86, which shows that the application effect of AI-related technology in the field of education is generally positive, and this effect size belongs to the "large effect size" range ($0.86 > 0.8$). The P value is less than 0.0001, which shows that the overall effect is significant, and the overall effect of the application of AI-related technologies in education is statistically significant. The 95% confidence interval for the overall effect size was [0.45, 1.28]. This confidence interval does not include zero, which further supports that the overall effect of the application of AI-related technologies in education is significant and of practical significance.

4.7 A Series of Novel Algorithms Based on BERT that can Realize Search and Other Functions are Applied to Meta-analysis for Experiments

In this experiment, the researchers conducted a meta-analysis using the following algorithm from a series of novel algorithms based on BERT that can implement search and other functions as an example:

A novel algorithm based on BERT that uses "split the text into word combinations of different lengths" can implement functions such as search: As the name of this algorithm does, it splits the text into word combinations of different lengths and then "searches for it" Words" and "split word combinations" use BERT to calculate similarity to achieve search results.

The researchers used the text content of the "titles" of these 13 research documents as the "text to be searched," selected "AI for education" as the "search term" for calculation, and "split the text into different lengths" The highest similarity between "word combination" and "search term" is used as the calculation result. The researchers set the weight of rankings 1 to 6 to 1.2 based on the value of this calculation result (replacing the original Hedges' g The value is multiplied by 1.2 as the new Hedges' g value, but the value of the Standard error of Hedges' g remains unchanged). The weight of the 7th-ranked one is set to 1 (the values of both Hedges' g and the Standard error of Hedges' g are unchanged), and the weights of the 8th to 13th rankings to 0.8 (multiply the original Hedges' g value by 0.8 as the new Hedges' g value, but the value of the Standard error of Hedges' g remains unchanged).

The researchers then used the new Hedges' g value calculated by this algorithm to conduct the following meta-analysis:

Table 7. Meta-analysis of the overall effect of the application of AI-related technologies in the field of education (using the new Hedges' g value in Section 4.7)

Index	Number of studies used in the model	Heterogeneity (I^2)	Is there publication bias?	Use a fixed-effects model or a random-effects model	Overall effect size	P-value	95% Confidence Interval
The overall effect of the application of AI-related technologies in the field of education	13	88.51%	Yes	Random effects model	0.84	<0.0001	[0.39, 1.29]

The funnel plot and forest plot of the meta-analysis of the overall effect of the application of AI-related technologies in the field of education are as follows (using the new Hedges' g value in Section 4.7):

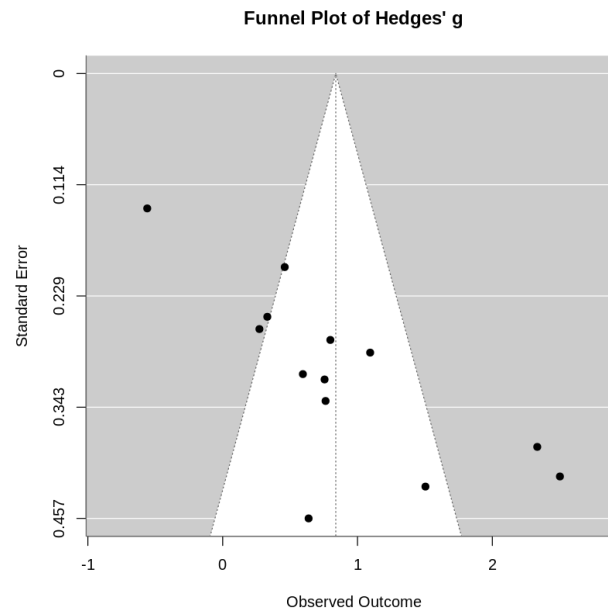


Figure 12. Funnel plot of the meta-analysis of the overall effect of the application of AI-related technologies in education (using the new Hedges' g value in Section 4.7)

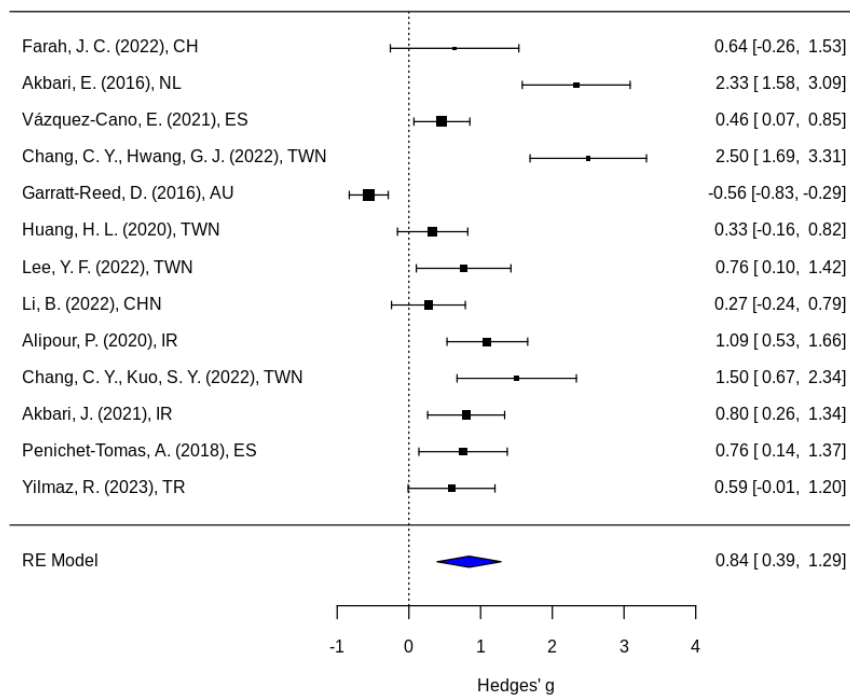


Figure 13. Forest plot of meta-analysis of the overall effect of the application of AI-related technologies in the field of education

In this meta-analysis, the overall effect size is 0.84, which shows that the application effect of AI-related technology in the field of education is generally positive, and this effect size belongs to the "large effect size" range ($0.84 > 0.8$). The P value is less than 0.0001, which shows that the overall effect is significant, and the overall impact of the application of AI-related technologies in education is statistically significant. The 95% confidence interval for the overall effect size was [0.39, 1.29]. This confidence interval does not include zero, which further supports that the overall effect of applying AI-related technologies in education is significant and of practical significance.

4.8 A Mixed Model Using A Weighted Random Effects Model and Fixed Effects Model is Applied to Meta-analysis for Experiments

Table 8. Meta-analysis of the overall effect of the application of AI-related technologies in the field of education (4.8 A mixed model using a weighted random effects model and a fixed effects model is applied to meta-analysis for experiments)

Index	Number of studies used in the model	Use a fixed-effects model or a random-effects model	Overall effect size	P-value	95% Confidence Interval
The overall effect of the application of AI-related technologies in the field of education	13	Experiment using a weighted effects model with both a random effects model weight of 50% and a fixed effects model weight of 50%	0.88	<0.0001	[0.48, 1.29]

The forest diagram of the meta-analysis of the overall effect of the application of AI-related technologies in the field of education (4.8 A mixed model based on a weighted random effect model and a fixed effect model is applied to the meta-analysis experiment) is as follows:

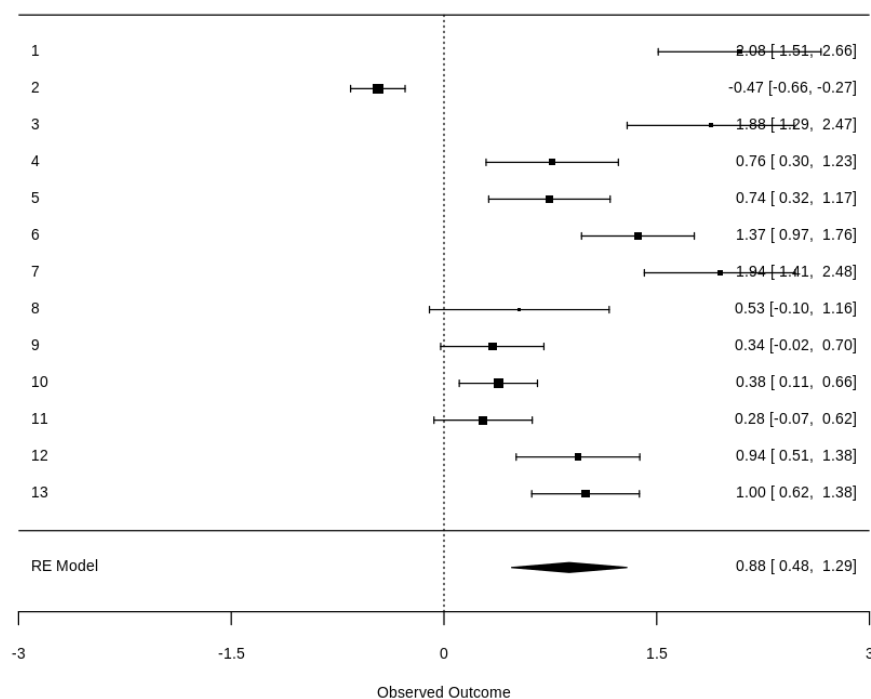


Figure 14. Forest plot of a meta-analysis of the overall effect of the application of AI-related technologies in the field of education (4.8 A hybrid model using a weighted random effects model and a fixed effects model is applied to meta-analysis for experiments)

The researchers conducted experiments using a weighted effects model, which included a random effects model with a weight of "50%" and a fixed effects model with a weight of "50%." A total of 13 research documents were included in the meta-analysis of the overall effect of applying AI-related technologies in education. These studies used different AI-related technologies in the field of education.

In this meta-analysis, the overall effect size is 0.88, which shows that the application effect of AI-related technology in the field of education is generally positive, and this effect size belongs to the "large effect size" range ($0.88 > 0.8$). The P value is less than 0.0001, which shows that the overall effect is significant, and the overall impact of the application of AI-related technologies in education is statistically significant. The 95% confidence interval for the overall effect size was [0.48, 1.29]. This confidence interval does not include zero, which further supports that the overall effect of applying AI-related technologies in education is significant and of practical significance.

5. Discussion

The findings of this meta-analysis provide significant insights into the effectiveness of artificial intelligence technologies in educational settings. Through a comprehensive analysis of 13 studies across diverse educational contexts, several key themes emerge that warrant detailed discussion.

5.1 Interpretation of Overall Effect Sizes

The large overall effect size (0.86) observed in this meta-analysis strongly supports the positive impact of AI technologies in education. This finding aligns with Chen et al.'s (2022) comprehensive review of two decades of AI in education, highlighting these technologies' transformative potential. The magnitude of this effect is particularly noteworthy when compared to typical effect sizes in educational interventions, suggesting that AI technologies may offer advantages over traditional educational approaches.

The varying effect sizes across different AI applications provide important nuance to this overall finding. The particularly strong effect size for chatbots and generative AI (1.02) supports Deng and Yu's (2023) findings regarding the effectiveness of conversational AI in educational contexts. This may be attributed to these technologies' ability to provide immediate, personalized feedback and support, which Ouyang and Jiao (2021) identified as crucial elements of successful AI-supported learning.

5.2 Heterogeneity and Contextual Factors

The substantial heterogeneity observed across studies (I^2 values ranging from 54.03% to 93.23%) suggests that the effectiveness of AI technologies is significantly influenced by implementation context. Several factors might explain this variation:

1. Educational Level and Subject Matter: The effectiveness appears to vary across different educational levels and subject areas, suggesting that AI technologies may be more suitable for specific learning contexts than others.
2. Implementation Approaches: The diverse ways AI technologies are integrated into educational settings likely contribute to the observed heterogeneity, highlighting the importance of careful implementation strategies.
3. Student Characteristics: Individual differences in student populations, including prior technology exposure and learning preferences, may influence the effectiveness of AI-based interventions.

5.3 Methodological Innovations and Their Implications

The novel algorithmic approaches introduced in this study represent important methodological contributions to the field. The weighted hybrid model combining random and fixed effects approaches (effect size = 0.88) demonstrates the potential value of methodological innovation in meta-analytic research. This approach addresses some limitations that Hansen et al. (2022) noted regarding the choice between fixed and random effects models in educational research.

6. Conclusion

This meta-analysis has provided robust empirical evidence regarding the effectiveness of artificial intelligence technologies in educational contexts. Through a systematic examination of 13 studies across eight countries, encompassing various AI applications from chatbots to virtual reality systems, several significant conclusions emerge.

The study's primary finding reveals a substantial positive effect of AI-related technologies in education, with an overall effect size of 0.86 (95% CI [0.45, 1.27], $p < 0.0001$). This significant effect size demonstrates that AI technologies significantly enhance educational outcomes across different learning contexts and subject areas. Particularly noteworthy is the effectiveness of chatbots and generative AI, which showed the highest effect size of 1.02 (95% CI [0.45, 1.59], $p < 0.0001$), suggesting these technologies are up-and-coming for educational applications.

Implementing three novel algorithmic approaches - Naive Bayes-based algorithms, TF-IDF variants, and BERT-based search algorithms - has provided additional methodological insights into analyzing educational technology effectiveness. The weighted hybrid model, combining random and fixed effects approaches, yielded robust results (effect size = 0.88, 95% CI [0.48, 1.29], $p < 0.0001$), suggesting that this methodological innovation offers a valuable tool for future meta-analytic studies in educational technology.

However, several important considerations emerge from this analysis. First, the substantial heterogeneity observed across studies (I^2 ranging from 54.03% to 93.23%) indicates that the effectiveness of AI technologies

varies considerably depending on implementation context, student characteristics, and specific technological applications. Second, publication bias suggests the need for cautious interpretation and calls for more comprehensive reporting of positive and negative results in future research.

Looking forward, this meta-analysis points to several crucial directions for future research and practice:

Educational Implementation: Educational institutions should consider implementing AI technologies, particularly chatbots and generative AI tools, while carefully accounting for their specific educational context and student needs.

Research Design: Future studies should address the current limitations by employing larger sample sizes, more diverse research designs, and more extended intervention periods to understand AI's long-term impacts on education better.

Methodological Development: The novel algorithmic approaches introduced in this study offer new tools for meta-analytic research, warranting further development and validation in different educational technology contexts.

Policy Implications: Educational policymakers should consider these findings when developing guidelines for AI technology integration in educational settings while remaining mindful of the heterogeneous effects observed across different contexts.

In conclusion, while this meta-analysis provides strong evidence supporting the positive impact of AI technologies in education, it also highlights the complexity of their implementation and the need for continued research. The significant positive effects observed across different technological applications suggest that AI can potentially transform educational practices substantially. However, successful implementation will require careful consideration of specific educational contexts, student needs, and technological capabilities. Future research should address the identified limitations while exploring new methodological approaches to understand better how AI can best serve educational objectives.

Authors contributions

Jincheng Zhang, prepares the original draft and subsequent revision and editing to produce the published work. In addition, Jincheng Zhang conceptualized the methodology and conducted the formal analysis, while Assistant Professor Dr. Thada Jantakoon supervised the conceptualization, methodology, validation, investigation, writing review, and editing preparation. Assistant Professor Dr. Rukthin Laoha conducted the validation. All authors reviewed and endorsed the final manuscript.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Canadian Center of Science and Education.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

References

- Agrawal, A., Gans, J., & Goldfarb, A. (2017). *What to expect from artificial intelligence*. <https://doi.org/10.3386/w24690>
- Akbari, E., Naderi, A., Simons, R. J., & Pilot, A. (2016). Student engagement and foreign language learning through online social networks. *Asian-Pacific Journal of Second and Foreign Language Education*, 1, 1-22. <https://doi.org/10.1186/s40862-016-0006-7>
- Akbari, J., Tabrizi, H. H., & Chalak, A. (2021). Effectiveness of virtual vs. non-virtual teaching in improving reading comprehension of Iranian undergraduate EFL students. *Turkish Online Journal of Distance Education*, 22(2), 272-283. <https://doi.org/10.17718/tojde.907585>
- Alipour, P. (2020). A comparative study of online vs. blended learning on vocabulary development among intermediate EFL learners. *Cogent Education*, 7(1), 1857489. <https://doi.org/10.1080/2331186X.2020.1857489>
- Alshammary, F. M., & Alhalafawy, W. S. (2023). Digital platforms and the improvement of learning outcomes: evidence extracted from meta-analysis. *Sustainability*, 15(2), 1305. <https://doi.org/10.3390/su15021305>
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to meta-analysis*. John Wiley & Sons. <https://doi.org/10.1002/9781119558378>
- Chang, C. Y., Hwang, G. J., & Gau, M. L. (2022). Promoting students' learning achievement and self-efficacy: A mobile chatbot approach for nursing training. *British Journal of Educational Technology*, 53(1), 171-188. <https://doi.org/10.1111/bjet.13158>
- Chang, C. Y., Kuo, S. Y., & Hwang, G. H. (2022). Chatbot-facilitated nursing education. *Educational Technology & Society*, 25(1), 15-27.
- Chen, B., Wang, Y., & Wang, L. (2022). The effects of virtual reality-assisted language learning: A meta-analysis. *Sustainability*, 14(6), 3147. <https://doi.org/10.3390/su14063147>
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264-75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education. *Educational Technology & Society*, 25(1), 28-47.
- Deng, X., & Yu, Z. (2023). A meta-analysis and systematic review of the effect of chatbot technology use in sustainable education. *Sustainability*, 15(4), 2940. <https://doi.org/10.3390/su15042940>
- Ding, R. X., Palomares, I., Wang, X., Yang, G. R., Liu, B., Dong, Y., ... Herrera, F. (2020). Large-Scale decision-making: Characterization, taxonomy, challenges and future directions from an Artificial Intelligence and applications perspective. *Information fusion*, 59, 84-102. <https://doi.org/10.1016/j.inffus.2020.01.006>
- Farah, J. C., Spaenlehauer, B., Sharma, V., Rodríguez-Triana, M. J., Ingram, S., & Gillet, D. (2022, March). Impersonating chatbots in a code review exercise to teach software engineering best practices. In *2022 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1634-1642). IEEE. <https://doi.org/10.1109/EDUCON52537.2022.9766793>
- Garratt-Reed, D., Roberts, L. D., & Heritage, B. (2016). Grades, student satisfaction and retention in online and face-to-face introductory psychology units: A test of equivalency theory. *Frontiers in psychology*, 7, 673. <https://doi.org/10.3389/fpsyg.2016.00673>
- Hansen, C., Steinmetz, H., & Block, J. (2022). How to conduct a meta-analysis in eight steps: a practical guide. *Management Review Quarterly*, 1-19. <https://doi.org/10.1007/s11301-021-00247-4>
- Huang, H. L., Hwang, G. J., & Chang, C. Y. (2020). Learning to be a writer: A spherical video-based virtual reality approach to supporting descriptive article writing in high school Chinese courses. *British Journal of Educational Technology*, 51(4), 1386-1405. <https://doi.org/10.1111/bjet.12893>
- Lee, J., Suh, T., Roy, D., & Baucus, M. (2019). Emerging technology and business model innovation: the case of artificial intelligence. *Journal of Open Innovation: Technology, Market, and Complexity*, 5(3), 44. <https://doi.org/10.3390/joitmc5030044>

- Lee, Y. F., Hwang, G. J., & Chen, P. Y. (2022). Impacts of an AI-based chatbot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational technology research and development*, 70(5), 1843-1865. <https://doi.org/10.1007/s11423-022-10142-8>
- Li, B., & Peng, M. (2022). Integration of an AI-Based Platform and Flipped Classroom Instructional Model. *Scientific Programming*, 2022(1), 2536382. <https://doi.org/10.1155/2022/2536382>
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., ... Prisma-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews*, 4, 1-9. <https://doi.org/10.1186/2046-4053-4-1>
- Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, 2, 100020. <https://doi.org/10.1016/j.caeai.2021.100020>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, 372. <https://doi.org/10.1136/bmj.n71>
- Penichet-Tomas, A., Jimenez-Olmedo, J. M., Pueo, B., & Carbonell-Martinez, J. A. (2018). Learning management system in sport sciences degree. In *EDULEARN18 Proceedings* (pp. 6330-6334). IATED. <https://doi.org/10.21125/edulearn.2018.1504>
- Vázquez-Cano, E., Mengual-Andrés, S., & López-Meneses, E. (2021). Chatbot to improve learning punctuation in Spanish and to enhance open and flexible learning environments. *International Journal of Educational Technology in Higher Education*, 18, 1-20. <https://doi.org/10.1186/s41239-021-00269-8>
- Yilmaz, R., & Yilmaz, F. G. K. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence*, 4, 100147. <https://doi.org/10.1016/j.caeai.2023.100147>

Appendix

Appendix A

Multiple Series of Novel Algorithms Invented and Proposed by Researchers

A1 More than one hundred novel algorithms based on Naive Bayes

When performing calculations, assume that the initial count of each word is 1 (this condition can also be omitted), and use whether the numerator of the conditional probability is a certain number (such as 1, 3, 5, etc.) as the grouping condition for "n 1", "n 2", "n 3", "n 4", etc. (such as n 1 refers to the conditional probability that the numerator of the conditional probability is between 1 and 3) (can be divided into multiple groups as required) (regarding "m 1", "m 2", "m 3", etc., they can be divided into the same or different groups according to the same or different numerical conditions as the n category).

Algorithm 1:

$$P(\text{Positive} | x) = \log ((P(A | \text{Positive}) + n 1) / (P(A | \text{Positive}))) \times ((P(B | \text{Positive}) + n 2) / P(B | \text{Positive})) \times ((P(C | \text{Positive}) + n 3) / P(C | \text{Positive})) \dots \times P(\text{Positive}) / P(x)$$

$$P(\text{Negative} | x) = \log ((P(A | \text{Negative}) + m 1) / (P(A | \text{Negative}))) \times ((P(B | \text{Negative}) + m 2) / P(B | \text{Negative})) \times ((P(C | \text{Negative}) + m 3) / P(C | \text{Negative})) \dots \times P(\text{Negative}) / P(x)$$

Algorithm 2:

$$P(\text{Positive} | x) = \log ((P(A | \text{Positive}) / (n 1 - P(A | \text{Positive}))) \times (P(B | \text{Positive}) / (n 2 - P(B | \text{Positive}))) \times (P(C | \text{Positive}) / (n 3 - P(C | \text{Positive}))) \dots \times P(\text{Positive}) / P(x)$$

$$P(\text{Negative} | x) = \log ((P(A | \text{Negative}) / (m 1 - P(A | \text{Negative}))) \times (P(B | \text{Negative}) / (m 2 - P(B | \text{Negative}))) \times (P(C | \text{Negative}) / (m 3 - P(C | \text{Negative}))) \dots \times P(\text{Negative}) / P(x)$$

Algorithms 3 to 20 use $P(A|\text{Positive})$ nth power, $(n - P(A|\text{Positive}))$, $(n / P(A|\text{Positive}))$, $(P(A|\text{Positive}) / n)$, $(P(A|\text{Positive}) / (P(A|\text{Positive})+n))$, $(P(A|\text{Positive})$ multiplied by $(P(A|\text{Positive})+n)$, $(P(A|\text{Positive}) \times (n 1 - P(A|\text{Positive})))$, $(n 1 - P(A|\text{Positive})) / P(A|\text{Positive})$, $(P(A|\text{Positive})$ to the $(P(A|\text{Positive})$ power), $(P(A|\text{Positive})$ to the $(P(A|\text{Positive})$ power) + n 1" power), $(P(A|\text{Positive})$'s $((P(A|\text{Positive}) - n 1)$ power), $(P(A|\text{Positive})$'s $(n 1 - (P(A|\text{Positive})))$ power), $(P(A|\text{Positive})$'s $((P(A|\text{Positive})n$ power), $(P(A|\text{Positive})$'s $((P(A|\text{Positive}) + n 1)$ to the power of n 2), $(P(A|\text{Positive})$'s $((n 1$'s $((P(A|\text{Positive})))$ power))" power), $(P(A|\text{Positive})$'s $((n 1$'s $((P(A|\text{Positive})+n 2)$ power))" power), $(P(A|\text{Positive})$'s $((n 1$'s

((P(A|Positive) minus n_2)" power))" power), (P(A|Positive) " n_1 's" (n_2 minus (P(A|Positive))" power))" power).

Algorithm 21:

Sort each conditional probability of Naive Bayes (such as descending order). Then grade the sorted conditional probabilities (such as grading by 10% and dividing them into 10 levels). Then you can set the weights of conditional probabilities that cannot be graded. So that Naive Bayes calculations can be performed more accurately.

Use

etc. in the corresponding positions

The above is the first series of algorithms.

The second series of algorithms: For the numerators of P(A|Positive), etc., perform a series of operations like the first series of algorithms.

The third series of algorithms: For the denominators of P(A|Positive), etc., perform a series of operations like the first series of algorithms.

The fourth series of algorithms: For the denominator (total number of documents in Positive + total number of words) in P(A|Positive) (for "total number of documents in Positive" separately), perform a series of operations as in the first series of algorithms.

The fifth series of algorithms: For the denominator (total number of documents in Positive + total number of words) in P(A|Positive) (for "total number of words" separately), perform a series of operations as in the first series of algorithms.

The sixth series of algorithms: For the whole prior probability (number of documents in the Positive category divided by the total number of documents). Perform a series of operations as in the first series of algorithms.

The seventh series of algorithms: For the numerator of the prior probability (number of documents in the Positive category divided by the total number of documents). Perform a series of operations as in the first series of algorithms.

The eighth series of algorithms: For the denominator of the prior probability (number of documents in the Positive category divided by the total number of documents). Perform a series of operations like the first series of algorithms.

There are more than a hundred novel algorithms based on Naive Bayes.

A2 Multiple series of novel algorithms related to TF, IDF, TF-IDF, and BM25

There are multiple novel algorithms related to TF, IDF, TF-IDF, and BM25, where TF refers to Term Frequency and IDF refers to Inverse Document Frequency.

A series of algorithms related to TF and TF-IDF are as follows:

Algorithm 1 ("TF" plus "TF-IDF" (normalize "TF" and "TF-IDF" to 1 before calculation))

Algorithm 2 ("TF" minus "TF-IDF" (normalize "TF" and "TF-IDF" to 1 before calculation))

Algorithm 3 ("TF-IDF" minus "TF" (normalize "TF" and "TF-IDF" to 1 before calculation))

Algorithm 4 (absolute value of the difference between "TF-IDF" and "TF" (normalize "TF" and "TF-IDF" to 1 before calculation))

Algorithm 5 ("TF" multiplied by "TF-IDF" (normalize "TF" and "TF-IDF" to 1 before calculation))

Algorithm 6 ("TF" Divide by "TF-IDF" (normalize "TF" and "TF-IDF" to 1 before calculation))

Algorithm 7 (divide "TF-IDF" by "TF" (normalize "TF" and "TF-IDF" to 1 before calculation))

A series of algorithms related to BM25 and TF-IDF are as follows:

Algorithm 1 ("BM25" plus "TF-IDF" (normalize "BM25" and "TF-IDF" to 1 before calculation))

Algorithm 2 ("BM25" minus "TF-IDF" (normalize "BM25" and "TF-IDF" to 1 before calculation))

Algorithm 3 ("TF-IDF" minus "BM25" (normalize "BM25" and "TF-IDF" to 1 before calculation) 1))

Algorithm 4 (absolute value of the difference between "TF-IDF" and "BM25" (normalize "BM25" and "TF-IDF" to 1 before calculation))

Algorithm 5 (multiply "BM25" by "TF-IDF" (normalize "BM25" and "TF-IDF" to 1 before calculation))

Algorithm 6 (divide "BM25" by "TF-IDF" (normalize "BM25" and "TF-IDF" to 1 before calculation))

Algorithm 7 (divide "TF-IDF" by "BM25" (normalize "BM25" and "TF-IDF" to 1 before calculation))

There are also several series of algorithms that perform similar operations as follows:

A series of algorithms that perform similar operations using IDF and TF-IDF.

A series of algorithms that perform similar operations using TF and IDF.

A series of algorithms that use TF and BM25 to perform similar operations.

A series of algorithms that use IDF and BM25 to perform similar operations.

And other series of algorithms

Of course, you can also use more than two algorithms to perform such similar operations, and use multiple algorithms to perform such similar combination operations, such as:

A series of algorithms that use TF, IDF, and TF-IDF to perform similar combination operations (such as using "TF-IDF" + "TF" - "IDF," "TF-IDF" divided by "TF" divided by "IDF," etc.).

A series of algorithms that use IDF, TF-IDF, and BM25 to perform similar combination operations.

A series of algorithms that use TF, IDF, TF-IDF, and BM25 to perform similar combination operations.

Of course, you can also use more than just TF, IDF, TF-IDF, and BM25 to perform such a series of operations, and you can also use other series of algorithms to perform related operations, thereby combining them into novel algorithms.

These multiple series of novel algorithms invented and proposed by researchers provide people with different options to cope with complex application scenarios, improve the use of related algorithms in various application scenarios, and provide other helpful information. For example, more accurately find important words, find words that different algorithms may underestimate, find suitable parameters, etc.

A3 A series of novel algorithms based on BERT that can realize search and other functions

In this study, researchers invented and proposed a series of novel algorithms based on BERT that can realize search and other functions. The main contents are as follows:

Based on BERT, a novel algorithm can realize search and other functions by "splitting text into word combinations of different lengths.": The name of this algorithm is to "split text into word combinations of different lengths" and then use BERT to calculate the similarity between "search words" and "split word combinations" to achieve search effects.

Based on BERT, a novel algorithm that can realize search and other functions by "arranging and combining different words in the text to be searched in different orders.": As the name of this algorithm, it is to "arrange and combine different words in the text to be searched in different orders" and then use BERT to calculate the similarity between "search words" and "word combinations of different orders" to achieve search effects.

A novel algorithm based on BERT that can realize search and other functions by "arranging and combining different words of "search words" and "text to be searched" in different orders.": As the name of this algorithm suggests, different words of "search words" and "text to be searched" are arranged and combined in different orders, and then "search words arranged and combined in different orders" and "text to be searched in different orders" are used to calculate the similarity using BERT, to achieve the search effect.

A4 A hybrid model based on random effect model and fixed effect model using weights

In this study, the researchers invented and proposed a hybrid model based on the random and fixed effect models using weights. The weight of the random effect model is "50%" and the weight of the fixed effect model is "50%". The weighted effect model of both (the weight can be set. For example, the weight of the random effect model and the fixed effect model are set to 80%, 20%, etc.). This hybrid weighted model can be used when it is unclear whether the random or fixed effect models should be used. More accurate calculations and other effects can be achieved.

Appendix B

Search strategies

Google Scholar

- #1. AI for education
- #2. artificial intelligence for education
- #3. AI applications in learning
- #4. machine learning in educational technology
- #5. AI in education
- #6. artificial intelligence in education
- #7. machine learning in education
- #8. deep learning in education
- #9. intelligent tutoring systems
- #10. adaptive learning systems
- #11. AI-powered learning analytics
- #12. chatbots for education
- #13. personalized learning with AI
- #14. ("AI in education" OR "artificial intelligence in education") AND meta-analysis
- #15. ("machine learning" OR "deep learning") AND ("education" OR "learning") AND meta-analysis
- #16. ("intelligent tutoring systems" OR "adaptive learning") AND ("student performance" OR "learning outcomes")
- #17. ("AI" OR "machine learning") AND ("higher education" OR "K-12") AND effect size
- #18. ("artificial intelligence" AND "education") AND ("systematic review" OR "meta-analysis")
- #19. ("educational technology" AND "AI") AND ("effectiveness" OR "impact study")

ERIC

- #1. AI for education
- #2. Artificial intelligence for education
- #3. AI applications in learning
- #4. Machine learning in educational technology
- #5. AI in education
- #6. Artificial intelligence in education
- #7. Machine learning in education
- #8. deep learning in education
- #9. intelligent tutoring systems
- #10. adaptive learning systems
- #11. AI-powered learning analytics
- #12. chatbots for education
- #13. personalized learning with AI

Scopus

- #1. AI for education
- #2. Artificial intelligence for education
- #3. AI applications in learning

- #4. Machine learning in educational technology
- #5. AI in education
- #6. Artificial intelligence in education
- #7. Machine learning in education
- #8. deep learning in education
- #9. intelligent tutoring systems
- #10. adaptive learning systems
- #11. AI-powered learning analytics
- #12. chatbots for education
- #13. personalized learning with AI