Can artificial Intelligence Technology Promote the Improvement of Student Learning Outcomes?——Meta Analysis Based on 50 Experimental and Quasi Experimental Studies

Lijuan Wang[1,a], Miaomiao Zhao[2,b]

[a]327260692@qq.com; [b]35612724@163.com

[1]College of Management Bohai University ,Jinzhou, China
[2]College of education science Dalian University ,Dalian, China

Abstract. Now, the artificial intelligence (AI) technology is becoming a powerful driving tool for the development of new intelligent education. However, the academic community has not yet reached a consensus on the key issue of the impact of AI technology on student learning outcomes. After conducting a meta-analysis of 50 experimental studies related to AI teaching technology, it was found that it can significantly improve students' learning outcomes. The AI technology has a high improvement effect on both cognitive and non-cognitive learning outcomes for students. The application of different types of AI technology has a positive promoting effect on learning outcomes, but the differences between them are not significant. There are also significant differences in the learning effectiveness of AI technology for students of different age groups, with the most significant improvement being in the university. The AI technology has a significant positive impact on the learning outcomes of natural sciences and interdisciplinary fields, while the impact on the learning outcomes of cultural and artistic disciplines is not significant. Based on the above results, future research and practice on AI technology in education should formulate various new policies to further enhance the promoting effect of AI technology on student learning and actively explore effective models for the practice of AI technology based on human-machine collaboration.

Keywords: Artificial intelligence technology; Learning effectiveness; Meta analysis

1 Introduction

At present, there are two main views in the academic education field on whether artificial intelligence technology will affect the learning outcomes of students: one is a positive view, which believes that artificial intelligence technology can effectively improve or promote the learning outcomes; Another approach is to hold a reverse view, which suggests that the artificial intelligence teaching technology cannot significantly improve the learning outcomes of students. Many scholars have demonstrated through experiments that artificial intelligence technology has a significant impact on the learning outcomes. Zafar etc. designed and developed an intelligent tutoring and teaching system, and conducted SQL teaching experiments on computer science students, which showed that the intelligent tutoring system can improve the learning effectiveness; [1]Korkmaz etc. validated the positive effect of

EIMT 2024, March 29-31, Wuhan, People's Republic of China

DOI 10.4108/eai.29-3-2024.2347685

educational robots on enhancing students' learning motivation through quasi experimental research methods; [2] Juli à etc. analyzed and studied the effect of educational robots on cultivating spatial abilities of 12-year-old students. Through comparative experiments, they found that students who participated in educational robot projects had significantly greater positive changes in spatial abilities than those who did not participate in the course; [3] Abbasi etc. developed the experimental results showed that the chatbot system has a significant promoting effect on student memory retention and academic performance. However, there are also some experimental research results indicating that artificial intelligence technology cannot effectively improve the learning effectiveness of students. [4] Nugent etc. pointed out that educational robots do not have a positive impact on student learning outcomes; [5] Sun Lihui etc. used a quasi-experimental design to compare the effects of virtual classes. The results showed that there was no significant difference in student learning outcomes between the two classes; [6] Chang etc. also pointed out that educational robots have a certain negative effect on students' STEM learning attitudes; [7]Calvo Ferrer etc. showed that intelligent educational games had no effect on students' learning outcomes. [8]Overall, there is currently no consensus in both academia and education regarding the impact of artificial intelligence technology on learning outcomes, and the geometry of the impact of artificial intelligence technology on learning outcomes is still unclear.

2 Research design

The sample was mainly cross searched for papers published between 2010 and 2023 in international authoritative literature databases such as Web of Science, Springer Link, ScienceDirect, ERIC, IEEE, etc. When searching, relevant keywords related to artificial intelligence technology were used, including: Artificial Intelligence, Education, Intelligent Learning Systems, Robots, Wearable Technology, Smart Educational Games, etc. The learning outcomes keywords include: Learning Achievement, Learning Performance, Learning Gains, Learning Outcomes, Learning Effects, etc.

To ensure the rigor and effectiveness of the research, the following guidelines have been established for the inclusion of literature in this study: firstly, the research topic must be highly relevant research on the impact of artificial intelligence technology on student learning outcomes; Secondly, research must be empirical and exclude non empirical research. The study must include an experimental group and a control group. Students in the experimental group use artificial intelligence technology, while students in the control group use traditional learning methods. Alternatively, the experiment may adopt pre-tests and post-tests, with pre-tests before and after artificial intelligence technology intervention and post-tests after intervention. The study must include various statistical raw data that can calculate the effect value. After screening, there were a total of 50 articles that fully met the meta-analysis criteria, including a total of 70 effect values. The specific process is shown in Figure 1.

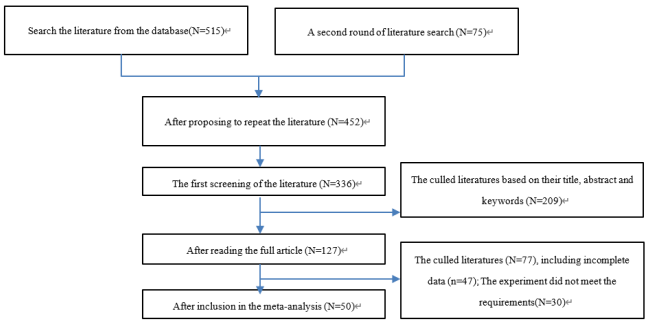

Figure 1. Prisma process for sample literature screening

The meta-analysis is a scientific research method that uses specific measurement and statistical analysis techniques to reanalyze multiple independent quantitative studies of the same topic in order to obtain a more universal conclusion. [9] The mainstream tools currently available for meta-analysis include R language, Stata, SAS, Review Manager, and Comprehensive Meta Analysis (CMA). Based on the actual situation, this study adopts CMA 3.0 as the main technical analysis tool to carry out relevant scientific research. In meta-analysis research, commonly used effect size indicators include Cohen's d and Hedges' g. [10] In this study, some literature had a sample size of less than 20, so Hedges' g was used as the effect size indicator. When the effect value is around 0.8, it indicates a significant and significant impact; When the effect value is around 0.5, it indicates a moderate degree of influence; When the effect value is around 0.2, it is generally considered that the impact is relatively small and the effect is not significant.

3 Research results and analysis

3.1 Publication bias test

In meta-analysis, the selected sample must represent the overall research population in the field, so the publication bias testing is necessary, which often uses funnel plot and fail-safe N to comprehensively evaluate. [11] In the funnel plot of Figure 2, it can be seen that the majority of the effect sizes of the 70 independent studies are located in the upper half of the funnel plot and distributed on both sides of the average effect value, indicating that the possibility of publication bias in this study is very small. This coefficient emphasizes that when the meta-analysis results have statistical significance, the larger the coefficient value, the lower the likelihood of conclusion reversal, the lower the probability of publication bias occurring. In this study, the loss of safety factor was 2630, far exceeding 210 (i.e. 5) × 40+10) indicates that there is little possibility of publication bias in this study. Based on the comprehensive judgment of the above two evaluation methods, the results of this study are reliable and can draw stable conclusions.

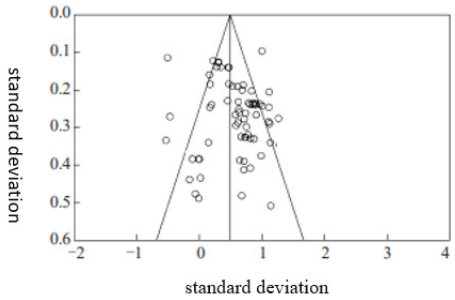

Figure 2 Funnel plot of effect value validity distribution

3.2 Heterogeneity test

Heterogeneity testing is mainly used to measure the degree of variation in effect values in research, with the aim of determining whether independent research results are mergeable. [12] In this study, Q and I2 were mainly used for testing, and the results showed that Q was 562.412 ($p<0.001$), indicating heterogeneity among the study samples; The value of I2 is 91.286%, which is greater than 75%, indicating that 90% of heterogeneity is due to true differences in effect values, and only 10% is caused by systematic errors. The publication time, country of origin, sample size, and research subjects of the sample literature included in the study may be various factors that lead to heterogeneity.[13] The I2 value reflects the proportion of heterogeneity in the total variation of the effect quantity, and the larger the value, the stronger the heterogeneity.[14] To ensure the validity and reliability of the research results, this study adopts a random effects model for analysis to effectively eliminate the influence of heterogeneity. below.

3.3 The overall effect of artificial intelligence technology on student learning outcomes

The sample size, mean, and standard deviation of the experimental group, control group, or pre and post test data of 50 studies (including 70 effect sizes) were imported into the meta-analysis software CMA3.0. The overall effect test of the impact of artificial intelligence technology on student learning effectiveness is shown in Table 1. The overall effect value Hedges' g is 0.825, and $p<0.001$, indicating that artificial intelligence technology can significantly improve students' learning outcomes.

Table 1 The overall effect test of the impact of artificial intelligence on learning outcomes

Effect model	Number of effects	effect size			two-tailed test	
		Hedges'g	standard error	95% confidence interval	Z-Value	P-Value
random effect	70	0.825	0.064	(0.526,0.982)	10.058	0

3.4 Effect analysis of artificial intelligence technology on each dimension of students' learning effect

According to Table 2, the effect value Hedges' g of artificial intelligence technology on the cognitive level of learning effectiveness is 0.907, reaching a statistically significant level

($p<0.001$, effect value greater than 0.8), indicating that artificial intelligence technology has a significant impact on the cognitive level of learning effectiveness. The effect value Hedges' g of non-cognitive level is 0.774, reaching a statistically significant level ($p<0.05$), There is no significant difference in the cognitive and non- cognitive effects of artificial intelligence technology on student learning. The research results indicate that artificial intelligence technology has a high improvement effect on both cognitive and non-cognitive learning outcomes of students.

Table 2 Testing the effects of artificial intelligence on various dimensions of learning effectiveness

learning effect	Number of effects	effect size			two-tailed test		Between -group effect
		Hedges'g	standard error	95% confidence interval	Z-Value	P-Value	
cognitive level	35	0.907	0.103	(0.526,0.982)	7.652	0	Q=0.635 P=0.326
Non-cognitive level	35	0.774	0.095	(0.465,0.882)	7.022	0	

The impact of artificial intelligence technology on various dimensions of cognitive and non-cognitive levels is shown in Table 3. At the cognitive level, the learning gain effect value is 0.911, and the cognitive ability effect value is 0.986, both of which are statistically significant ($p<0.001$), indicating that artificial intelligence technology has a positive impact on students' cognitive ability and learning gain. The effect value of problem-solving ability is 0.624 ($p<0.05$), indicating that artificial intelligence technology has a moderate positive impact on students' problem-solving ability. The impact of artificial intelligence technology on students' creative thinking and spatial ability is not statistically significant ($p>0.05$). At the non-cognitive level, the effect value of learning participation is 0.932, the effect value of learning attitude is 0.881, the effect value of learning motivation is 0.803, and the effect value of learning will is 0.796, all of which are statistically significant ($p<0.05$), indicating that artificial intelligence has a positive and significant impact on student learning effectiveness in all four aspects, with the greatest impact on learning participation. The effect value of learning interest is 0.557, which is statistically significant ($p<0.001$), indicating that artificial intelligence technology has a moderate positive promoting effect on learning interest.

Table 3 Analysis of the Effects of Artificial Intelligence on Learning Effectiveness in Various Dimensions

learning effect		Number of effects	effect size			two-tailed test	
			Hedges'g	standard error	95% confidence interval	Z-Value	P-Value
cognitive level	learning gain	12	0.911	0.137	(0.662,1.216)	7.011	0
	problem-solving ability	5	0.624	0.206	(0.245,1.028)	2.627	0.005
	creative thinking	5	0.302	0.322	(0.374,1.122)	0.782	0.357

	spatial ability	6	0.285	0.398	(0.145,1.062)	1.226	0.113
	cognitive competence	7	0.986	0.316	(0.497,1.134)	3.521	0
Non-cognitive level	learning interest	5	0.557	0.126	(0.216,0.828)	4.320	0
	learning attitude	5	0.881	0.321	(0.207,1.012)	2.624	0.006
	learning motivation	15	0.803	0.207	(0.307,1.143)	4.337	0
	Participation in learning	5	0.932	0.215	(0.662,1.307)	5.102	0
	Will to learn	5	0.796	0.176	(0.406,1.213)	3.988	0

3.5 The impact of different moderating variables on student learning outcomes

Firstly, the research divides the application types of artificial intelligence technology into educational robots, wearable technology, intelligent educational games and intelligent learning systems. As it can be seen from Table 4, the effect value of educational robot is 0.886, the effect value of wearable technology is 0.864, the effect value of intelligent educational game is 0.912, and the effect value of intelligent learning system is 0.942. All effect values are greater than 0.8 and are statistically significant ($p < 0.001$). It shows that four different types of artificial intelligence applications have positive and positive effects on students' learning. The inter-group effect $Q=4.532$ did not reach the statistically significant level ($p>0.05$), indicating that there was no significant difference between different types of artificial intelligence technology application in promoting the learning effect of students.

Table 4 The Effect Test of Different Types of Artificial Intelligence on Learning Effectiveness

learning effect	Number of effects	effect size			two-tailed test		Between-group effect
		Hedges'g	standard error	95% confidence interval	Z-Value	P-Value	
Educational robot	23	0.886	0.103	(0.524,1.102)	5.543	0	Q=4.532 P=0.275
wearable technology	15	0.864	0.095	(0.495,0.897)	7.216	0	
Intelligent educational game	22	0.912	0.112	(0.547,1.126)	6.758	0	
Intelligent learning system	10	0.942	0.176	(0.507,0.962)	4.112	0	

Secondly, the sample literature included in the study is divided into four stages: preschool, elementary, middle, and university. According to the results in Table 5, the effect value during the university stage is 1.021, and the effect value during the preschool stage is 0.989, both of which are greater than 0.8 and reach a statistically significant level ($p<0.05$). This indicates that the impact of artificial intelligence technology on learning outcomes during the university and preschool stages has reached a positive and significant effect; The effect value in primary school stage is 0.677, and the effect value in secondary school stage is 0.634, both of which

have statistical significance ($p<0.001$), indicating that artificial intelligence technology has a moderate positive impact in primary and secondary school stages. The intergroup effect $Q=10.26$ reached a statistically significant level, indicating significant differences in the learning effectiveness of artificial intelligence technology among students of different age groups.

Table 5 The Effect Test of Artificial Intelligence on the Learning Effectiveness of Students in Different Stages

learning effect	Number of effects	effect size			two-tailed test		Between-group effect
		Hedges'g	standard error	95% confidence interval	Z-Value	P-Value	
preschool	10	0.989	0.103	(0.447,1.126)	2.336	0	Q=10.26 P=0.024
primary school	15	0.677	0.105	(0.488,0.861)	5.172	0	
middle school	20	0.634	0.074	(0.452,0.633)	6.483	0	
university	25	1.021	1.002	(0.307,4.322)	2.226	0	

Thirdly, this study divides the disciplinary fields applied by artificial intelligence into four disciplines: humanities and social sciences, natural sciences, cultural arts, and interdisciplinary studies. The results are shown in Table 6, with an inter-group effect of $Q=36.00$, reaching a statistically significant level ($p<0.05$), indicating significant differences in the learning effects of artificial intelligence technology on different disciplines. The impact on the natural science discipline is the greatest, with an effect value of 1.165, followed by a cross disciplinary effect value of 0.916. The effect values of both disciplines are above 0.8, and both have statistical significance ($p<0.001$), indicating that artificial intelligence has a positive and significant effect on the learning of natural science and cross disciplinary integration. The effect value of humanities and social sciences is 0.732, ranging from 0.5 to 0.8, and both have statistical significance ($p<0.0001$), indicating that artificial intelligence technology has a moderate positive impact on the learning effect of humanities and social sciences. The effect value of artificial intelligence technology on cultural and artistic disciplines is only 0.342 ($p>0.05$), which is not statistically significant.

Table 6 The Effect Test of Artificial Intelligence on the Learning Effectiveness of Students in Different Disciplines

learning effect	Number of effects	effect size			two-tailed test		Between-group effect
		Hedges'g	standard error	95% confidence interval	Z-Value	P-Value	
humanities & social sciences	16	0.732	0.118	(0.524,1.102)	5.543	0	Q=36.00 P=0.005
natural science	16	1.165	0.062	(0.495,0.897)	7.216	0	
culture and art	16	0.117	0.108	(0.547,1.126)	6.758	0.342	
Cross synthesis	22	0.916	0.164	(0.507,0.962)	4.112	0	

4 Research conclusion

1) The results showed that the overall effect value of artificial intelligence technology on students 'learning effect was 0.825, indicating that artificial intelligence technology could actively promote students 'learning effect, which was reflected not only in the cognitive level, but also in the non-cognitive level.

2) There was no significant difference between different types of technology application, which could effectively improve students 'learning effect. In addition, on the whole, artificial intelligence technology can have different degrees of positive impact on students' learning results under different classes and subject types.

3) Based on this, the paper puts forward the action path and improvement strategy of future artificial intelligence technology to promote students' learning effect from three aspects: policy design, teaching reform and technology promotion. The policy design should actively promote the deep integration of artificial intelligence and education. The teaching reform, man-machine collaboration leads the new way of artificial intelligence technology in the future. The technology promotion, to promote artificial intelligence multi-section, multi-disciplinary and differentiated application.

References

[1] Zafar A, Albidewi I. (2015) Evaluation study of ELGuide: A framework for adaptive e-learning. Computer Applications in Engineering Education, Vol 23(04):542-555.

[2] Özgen KORKMAZ. (2016) The Effect of Lego Mindstorms Ev3 Based Design Activities on Students' Attitudes towards Learning Computer Programming, Self-efficacy Beliefs and Levels of Academic Achievement. Modern Computing, Vol 4(04):994-1007.

[3] Julià C, Antolí J Ò. (2016) Spatial ability learning through educational robotics. International Journal of Technology and Design Education, Vol 26(02):185-203.

[4] Abbasi S, Kazi H. (2017) Measuring effectiveness of learning chatbot systems on student's learning outcome and memory retention. Asian Journal of Applied Science and Engineering, Vol 3(02): 251-260.

[5] Nugent G, Barker B, Grandgenett N. (2010) Impact of robotics and geospatial technology interventions on youth STEM learning and attitudes. Journal of Research on Technology in Education, Vol 42(04): 391-408.

[6] Sun Lihui, Zhou Liang. (2023) Development direction of generative artificial intelligence education policies for primary and secondary schools: an analysis based on Japan's "Application Guidelines for Generative Artificial Intelligence Education for Primary and Secondary School Students". China Electronic Education, Vol 11 (05):53-61.

[7] Chang R C, Chung L Y, Huang Y M. (2016) Developing an interactive augmented reality system as a complement to plant education and comparing its effectiveness with video learning. Interactive Learning Environments, Vol 24(06):1245- 1264.

[8] Calvo-Ferrer J R. (2017) Educational games as stand-alone learning tools and their motivational effect on L2 vocabulary acquisition and perceived learning gains. British Journal of Educational Technology, Vol 48(02):264-278.

[9] Zheng Minghua. Meta analysis software application and instance analysis. Beijing: People's Health Publishing House, 2013.
[10] Shan Junhao, Gong Lingling, Li Yu. (2019) The impact of educational robots on student learning outcomes: a meta-analysis based on 49 experimental or quasi experimental research papers. China Electronic Education, Vol 28(05): 76-83.
[11] Zheng L Q, Niu J Y, Zhong L. (2021) The effectiveness of artificial intelligence on learning achievement and learning perception: A meta-analysis. Interactive Learning Environments, Vol 46(01):1-15.
[12] Hu Dexin, Wang Yaorong. (2022) A study on the impact of artificial intelligence technology on student learning outcomes: a meta-analysis based on 40 experiments and quasi experiments from 2010 to 2022. Journal of Tianjin University (Social Sciences Edition), Vol 24(06): 493-502.
[13] Gu Xiaoqing, Li Rui, Li Shijin. (2023) What is the current situation of artificial intelligence education—— Evidence response to the results of artificial intelligence education competitions in primary and secondary schools. Open Education Research, Vol 29 (06): 82-91.
[14] Zhu Yongxin, Yang Fan. (2023) ChatGPT/Generative Artificial Intelligence and Educational Innovation: Opportunities, Challenges, and the Future. Journal of East China Normal University (Education Science Edition), Vol 41 (07): 1-14.