

GenAI 支持人机协同学习的效果与优化策略研究——基于 25 项实验和准实验的元分析

A Study on the Effects and Optimization Strategies of GenAI-Supported Human-Machine

Collaborative Learning: A Meta-Analysis

夏天驰¹, 王娟¹

¹ 江苏师范大学智慧教育学院

*2020240567@jsnu.edu.cn

*wjuan8@163.com

【摘要】 研究采用元分析系统评估生成式人工智能（GenAI）在人机协同学习中的应用效果，探讨了不同条件下的效应差异。基于多项实验与准实验研究，结果显示 GenAI 能显著提升学习成效，其改善效果统计学意义显著。研究发现，学科领域、知识类型、干预时长、学习方式、角色设定及教学模式均对效应存在调节作用，各领域与知识类型提升趋势一致，但效应强度受学科特性与任务复杂度影响；短期干预成效更明显，而长期干预可能因新颖性降低或信息过载而波动；群体协作及角色化互动可进一步优化学习效果。本研究为教育实践和未来教育技术设计提供了实证依据。

【关键词】 生成式人工智能；人机协同学习；学习效果；元分析

Abstract: This study employs meta-analysis to systematically evaluate the effectiveness of Generative Artificial Intelligence (GenAI) in human-machine collaborative learning and to examine variations in its effects under different conditions. Based on multiple experimental and quasi-experimental studies, the results indicate that GenAI significantly enhances learning outcomes, with the improvement being statistically significant. The study finds that variables such as discipline domain, knowledge type, intervention duration, learning mode, role setting, and instructional model moderate these effects: while the trend of improvement is consistent across various disciplines and knowledge types, the strength of the effect is influenced by disciplinary characteristics and task complexity; short-term interventions yield more pronounced improvements, whereas long-term interventions may fluctuate due to diminished novelty or information overload; and collaborative learning along with role-based interactions can further optimize outcomes. This study provides empirical evidence for educational practice and the design of future educational technologies.

Keywords: Generative Artificial Intelligence (GenAI), Human-Machine Collaborative Learning, Learning Effectiveness, Meta-Analysis

1. 前言

生成式人工智能（GenAI）在教育中的应用日益广泛，但对人机协同学习的影响尚存争议。一些研究表明，GenAI 能够提升学习成效，例如 Hwang 等（2023）研究发现，采用 GenAI 的协同学习在知识建构和成绩上优于传统方法。然而，部分研究指出 GenAI 可能增加认知负荷或影响自主学习能力（Forero & Herrera-Suárez, 2023）。为此，研究通过元分析方法系统评估 GenAI 对人机协同学习的总体影响，并探讨不同学科、任务类型及学习方式的差异。

2. 研究方法与过程

传统文献分析方法主要是定性分析，存在一定的主观性，很难给出定量的结论（夏凌翔，

2005)。元分析按照严格的程序，运用统计手段整合多项定量研究，客观评价以往研究结果（蔡金亭，2012）。本研究整合国内研究，并严格遵循格拉斯(Glass)等人提出的元分析程序评估 GenAI 对学习效果的影响。

2.1 文献检索

基于中国知网、万方等中文数据库，以“生成式人工智能+学习效果”等关键词构建检索式，初筛文献 5124 篇。纳入标准包括：

- 1) 基于 GenAI 的人机协同学习实验或准实验；
- 2) 包含与传统学习的成效对比；
- 3) 提供均值、标准差等可量化数据，最终纳入 25 项研究（图 1）。

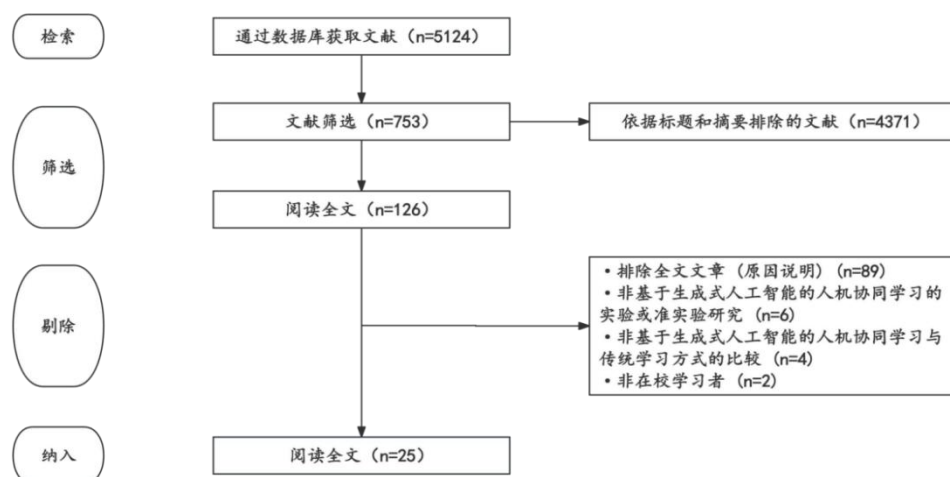
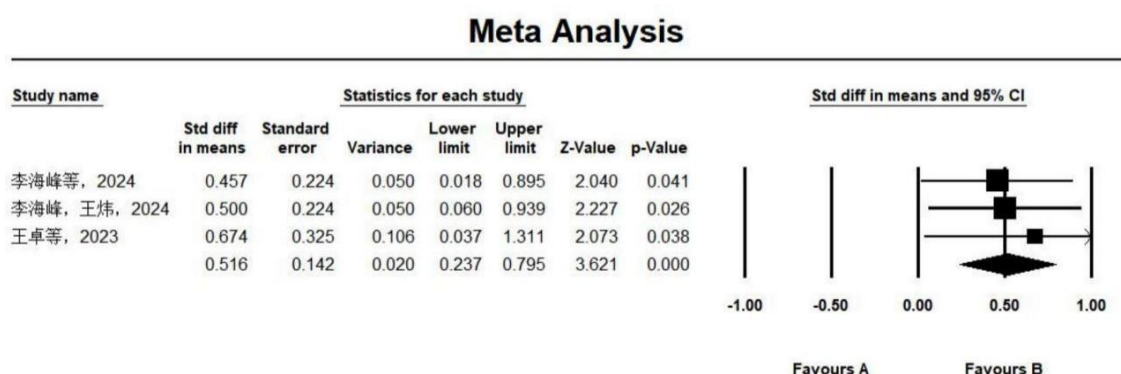


图 1 文献数据采集流程

2.2 特征值编码与数据分析



使用CMA软件整合25项研究的效应值，通过森林图（如图2）显示GenAI的总体效应值为0.516（95% CI=[0.237,0.795]， $p=0.000$ ），各研究效应值均显著。异质性检验（ $Q=156.005$ ， $I^2=86.539\%$ ）表明样本间差异显著，故采用随机效应模型。

图2 GenAI 对学习成效总体影响森林图（节选部分）

2.3 发表偏倚检验与异质性检验

失安全系数（ $Nfs=726$ ）及 Egger（ $p=0.727$ ）、Begg（ $p=0.554$ ）检验表明，发表偏倚风险较低。异质性分析（ $I^2=86.539\%$ ）支持采用随机效应模型校正结果。如图 3 所示。

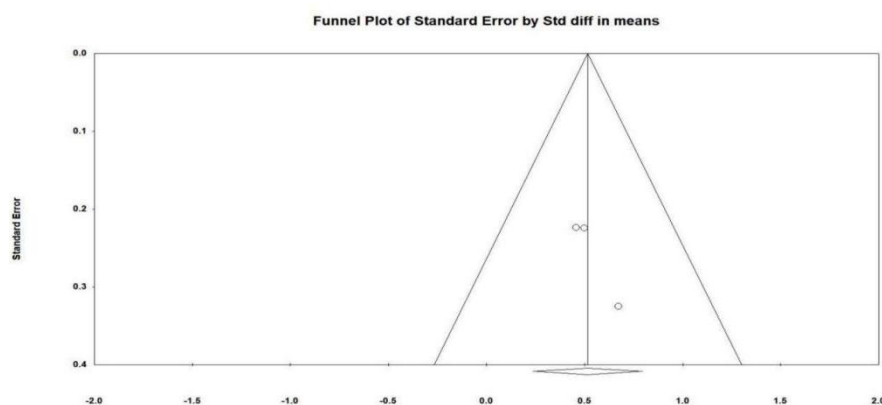


图3 样本数据发表偏倚检测漏斗图

由于纳入的分析样本在研究设计、样本特征与测量工具等维度存在差异，为提高分析结果的科学性和准确性，本研究采用博伦斯坦（Borenstein et al., 2009）开发的分析模型，通过异质性检验选取模型处理数据间的差异性，确保元分析的结果接近无偏估计。选取分析模型依据是Q统计量和I²统计量。Q统计量的检测结果为156.005 (df=21, $p < 0.001$), I²为86.539%。这表明样本间存在较显著的异质性（Higgins et al., 2003）。为有效处理样本间效应值的差异，本研究选择随机效应分析模型校正结果。

3.结果探讨

3.1 学科领域差异对学习效果的调节作用

GenAI 支持的人机协同学习效果在不同学科领域呈现显著差异。社会科学领域因其强调情境模拟与人文交互特性，GenAI 能够通过复现社会脉络和提供类人对话显著提升学习效果；而自然科学、工程与医学领域因涉及复杂逻辑与深度问题求解，GenAI 在知识理解与生成上的局限性可能导致其成效相对有限。这一发现表明，教育技术开发者需依据学科特点优化 AI 功能设计，例如为自然科学领域强化逻辑推理模块，为社会科学领域增强情境生成能力，以实现技术与学科需求的精准适配。

3.2 知识类型与 GenAI 的适配性机制

GenAI 对陈述性知识（如事实、概念）和程序性知识（如技能、流程）的学习均具促进作用，但适配性存在差异。陈述性知识的学习易受 AI 生成内容的准确性干扰，开放数据库中的冗余或错误信息可能导致学习者认知混淆，尤其在陌生情境中易产生误解；而程序性知识的学习则可通过 AI 的即时反馈功能强化技能训练。因此，教师需在陈述性知识教学中引导学生批判性审视 AI 生成内容，并通过高质量数据输入优化知识库，同时在程序性知识学习中充分利用 AI 的交互优势。

3.3 干预时长的动态效应与教学策略优化

干预时长对学习效果的影響呈现动态变化特征。短期干预（如单次任务或短期课程）因技术新颖性和即时反馈优势显著提升学习动力与参与度，但长期干预可能因新颖性衰退、信息过载或注意力分散导致效果波动。教育实践者需设计模块化、阶段性的 GenAI 支持方案，避免单一化交互模式，并探索与传统教学手段结合的“间歇性强化”策略，以平衡技术应用的短期效能与长期可持续性。

3.4 群体协作与角色化交互的增效路径

群体协作学习模式能够显著优化 GenAI 的效能。学习者通过多样化提问与批判性讨论可提升 AI 生成内容的质量，而个体学习易因提问精准性不足或信息筛选能力有限陷入浅层交互。此外，为 GenAI 设定明确角色（如导师、协作者）可增强交互的精准性与情感连接，通过定制化

情境与多轮对话提升学习体验。教育者需推广小组协作模式，并根据教学目标设计角色化 AI 交互策略，例如将 AI 定位为“虚拟辩论对手”或“项目协作者”，以适配不同学习场景。

3.5 技术赋能与教育本质的辩证关系

GenAI 的技术赋能需与教育本质保持平衡。尽管 GenAI 能显著提升学习效率，但其过度依赖可能削弱学习者自主思考能力，且无法替代人类社交的情感支持与深层认知互动。未来研究需探索“人主导、机辅助”的混合教学模式，在教学设计中协调技术工具与人类教师的角色分工，同时关注伦理规范制定，防范技术异化风险，确保教育技术应用兼具效率性与人文关怀。

参考文献

- 王卓,马洋珍,杨现民,李康康(2023). ChatGPT 类阅读平台对研究生学术阅读能力的影响 [J]. 开放教育研究,(6):60-68.
- 李海峰,王炜(2023). 人机协同深度探究性教学模式——以基于 ChatGPT 和 QQ 开发的人机协同探究性学习系统为例 [J]. 开放教育研究,0(6):69-81.
- 和文斌,赵帅,阿不来提·瓦依提,塔卫刚 & 徐恩伟.(2024).基于生成式人工智能的人机协同学习更能提升学习成效?——基于 20 项实验和准实验的元分析.开放教育研究(05),101-111.doi:10.13966/j.cnki.kfjyyj.2024.05.010.
- 和文斌,范占江,郑浩,谭征,赵帅(2024). ChatGPT 赋能 CSCL:未来图景与突围路径 [J]. 现代教育技术,(4):37-46.
- 李海峰,王炜(2024). 人机争论探究法:一种争论式智能会话机器人支持的学生高阶思维能力培养模式探索 [J]. 电化教育研究,(3):106-112+128.
- 沈书生,祝智庭(2023). ChatGPT 类产品:内在机制及其对学习评价的影响 [J]. 中国远程教育,(4):8-15.
- Hwang,W. Y.,Nurtantyana,R.,Purba,S. W. D.,Hariyanti,U.,Indrihapsari,Y.,&Surjono,H. D. (2023).AI and recognition technologies to facilitate English as foreign language writing for supporting personalization and contextualization in authentic contexts[J]. Journal of Educational Computing Research,61(5):1008-1035.
- Borenstein,M.,Hedges,L. V.,Higgins,J. P. T.,& Rothstein,H. R.(2009). Introduction to Meta-Analysis[M]. Chichester,UK:Wiley:77-85.
- Higgins,J. P.,Thompson,S. G.,Deeks,J. J.,& Altman,D. G.(2003). Measuring inconsistency in meta-analyses[J]. BMJ British Medical Journal,327(7414):557-560.
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students[J]. Frontiers in Psychology, 14:1260843.