

生成式人工智能如何影响学生发展

——基于31项实验与准实验研究的元分析

□胡钦太 梁心贤 刘颜帆 王姝莉

摘要:生成式人工智能在教育领域的应用日益广泛,但其对学生发展的影响效果仍存在争议。通过对31项实验和准实验研究的元分析发现,生成式人工智能对学生认知、行为能力和情感态度方面的发展具有中等程度的正向影响。具体而言,生成式人工智能在认知层面能够加深学生对概念的理解;在行为能力层面能够提高学生的基本技能和能力,有效培养学生的问题解决能力和创造力等高阶思维能力;在情感态度层面能增强学生的价值认可,激发学习兴趣和态度动机,提升自我调节和监控,增强获得感和效能感。这些作用也受到实验周期、样本规模和教育阶段的调节影响,表现为:在实验周期上,中期实验对学生发展的促进效果最强,短期实验次之,长期实验最弱;在样本规模上,中等规模及大规模样本对学生认知发展方面的调节作用相对有限,而在行为能力和情感态度发展方面则展现出较强的调节效果;在教育阶段上,生成式人工智能在基础教育阶段对学生的认知和情感态度发展具有更为显著的调节作用。未来为推动生成式人工智能在教育教学中的有效应用,应不断创新融合教学模式,研发教育专属大模型;加强工具使用培训,预防对技术的过度依赖;注重组织群体研讨,构建“师—生—机”多元反馈机制;全面融合数据分析,不断深化情境教学。

关键词:生成式人工智能;学生发展;认知;行为能力;情感态度;元分析

中图分类号:G434 **文献标识码:**A **文章编号:**1009-5195(2025)02-0083-09 doi10.3969/j.issn.1009-5195.2025.02.009

基金项目:2023年度国家社会科学基金重大项目“数字化促进中国教育现代化路径策略研究”(23VRC076);2024年华南师范大学课外科研种子培育项目“思政引领的信息技术课程资源网站开发与案例应用”(24JXKC02)。

作者简介:胡钦太,博士,教授,博士生导师,广东工业大学计算机学院(广东广州 510006);华南师范大学教育信息技术学院(广东广州 510631)。梁心贤、刘颜帆,硕士研究生,华南师范大学教育信息技术学院(广东广州 510631)。王姝莉,博士研究生,华南师范大学教育信息技术学院(广东广州 510631)。

一、问题提出

生成式人工智能在教育中的应用日益受到关注。为加速培养适应未来社会的创新型人才,国务院早在2017年发布的《新一代人工智能发展规划》中就明确指出要“推动人工智能在教学、管理、资源建设等全流程应用”(中华人民共和国中央人民政府,2017),探索人工智能在促进学生个性化发展和创新能力提升方面的巨大潜力。各个国际组织亦对人工智能的最新发展予以积极回应,如美国高等教育信息化协会(EDUCAUSE)发布的《2024地平线报告:教与学版》指出生成式人工智能对高等教育领域产生了深远影响(EDUCAUSE,2024),联合国教科文组织(UNESCO)发布的《学生人工智能能力框架》更是明确提出了学生在人工智能时代必须具备

的知识、技能和价值观(UNESCO,2024)。面对生成式人工智能带来的机遇与挑战,我国应深入探索智能时代的教育规律、人才培养模式和人工智能应用策略,构建以人工智能为引领、学生发展为核心的创新教育生态系统,推动基础教育高质量发展,培养适应人工智能时代的新型人才(柯清超等,2024)。

目前,ChatGPT、DeepSeek、文心一言、Kimi等生成式人工智能工具在教育领域中已得到广泛应用,其强大的内容生成能力与个性化交互特性,不仅为学生提供了丰富的学习资源和个性化学习路径,还极大地激发了他们的创造力和想象力(张线,2023)。相较于以往的人工智能技术,生成式人工智能凭借其强大的数据支撑、算力资源和模型架构,能以拟人化的对话形式提供响应(和文斌等,2024),并重塑传统的教学框架和互动模式。已有研究指

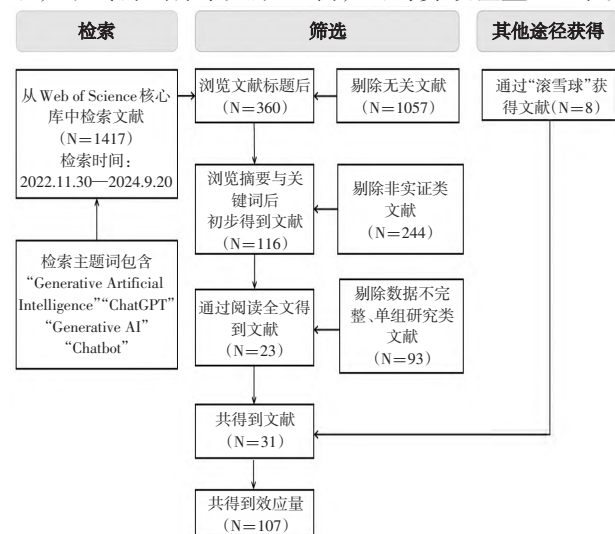
出,将生成式人工智能工具运用到教学中,有助于促进学生在学业成绩(Anyanwu et al., 2024)、问题解决能力(Lee et al., 2024)、创造力(Habib et al., 2024)等方面的发展。此外,相关研究也证实了ChatGPT的类人式对话和积极性反馈对于减轻学生考试焦虑及其他消极情绪具有积极作用(Habib et al., 2024)。但也有研究表明,生成式人工智能在实际应用过程中存在一定的局限性,其对于切实增强学生问题解决能力的效果并不显著(Liao et al., 2024),甚至出现了将ChatGPT引入教学实践后未能提升学生学业成绩的情形(Bhatia et al., 2024)。还有研究表明,采用生成式人工智能辅助的协作学习与传统的协作学习效果相比并无明显差异(Escalante et al., 2023)。由此可见,当前学界对生成式人工智能是否有利于学生发展尚未达成一致结论。有鉴于此,本研究将依托国际SSCI期刊中针对此类议题所发表的实验和准实验研究成果,运用元分析方法,探讨生成式人工智能对学生发展的影响。研究具体包括三个问题:(1)生成式人工智能能否促进学生在认知、行为能力、情感态度方面的发展,其具体效果如何?(2)哪些因素(如教育阶段、样本规模和实验周期)会对其效果产生调节作用?(3)不同因素对生成式人工智能促进学生认知、行为能力和情感态度方面的影响是否存在差异?

二、研究过程

1. 文献检索与筛选

研究设定Web of Science数据库中“SSCI期刊”为检索范围,鉴于ChatGPT3.5之前版本的生成式人工智能技术在性能上存在局限,而当前研究主要聚焦探索ChatGPT3.5之后涌现的大语言模型及其相关产品对学生发展的影响,故将检索时间限定在2022年11月30日(即ChatGPT3.5发布时间)至2024年9月20日,同时以“Generative Artificial Intelligence”“ChatGPT”“Generative AI”“Chatbot”为关键词进行检索。

为探索生成式人工智能对学生发展的影响,研究设定了如下文献筛选标准:(1)研究主题为生成式人工智能对学生发展的影响;(2)研究对象包括基础教育和高等教育阶段的学生;(3)纳入的文献必须有实验组与对照组的比较,单组前后测实验不纳入本研究的考察范畴;(4)文献的数据完整,须包含可计算效应量的信息。具体的筛选过程如图1所示,最终得到样本文献31篇,可计算效应量107个。



2. 特征值编码

研究不仅提取了纳入文献的基础信息(包括文章名、作者、期刊、发表时间等),还对31篇文献进行了特征值编码处理。编码由两位研究生分别独立进行,Cohen Kappa系数为0.821,表明编码结果可信。编码内容包括教育阶段、实验周期、样本规模、作用效果等。综合来看,此次纳入元分析的31篇文献,共计有107个效应量,包括基础教育和高等教育两个阶段,且以高等教育阶段的研究为主(其中包括4项以研究生为对象的研究)。样本规模介于11~260人,其中居于30~100人的研究最多。实验时长介于1~16周,其中1个月以内的研究最多。特征值编码结果如表1所示。

表1 特征值编码结果

教育阶段		样本规模		实验周期	
特征	效应数	特征	效应数	特征	效应数
基础教育	32	30以下	31	1个月以内	54
		30~100人	67	1~4个月	31
高等教育	75	100人以上	9	4个月以上	22

国内诸多学者将布鲁姆的目标分类理论划分为认知、行为能力和情感态度三大维度(王映朝等, 2018; 谢静等, 2023)。本研究通过对效应量的特征值编码结果进行分析,从中共提取出14个作用效果,并将这些作用效果归入认知、行为能力、情感态度三个维度,以进一步探究生成式人工智能对学生发展的影响。其中,认知维度涵盖学生对知识点概念的理解、所涉及的学业成绩及认知能力;行为能力维度主要包括学生在使用生成式人工智能过程中行为技能的提升,涉及基本技能和能力、计算思

维和算法思维、批判性思维、问题解决能力、创造力以及协作能力等;情感态度维度主要包括学生的价值认可、学习兴趣、态度动机、自我调节与监控,以及获得感与效能感。

三、研究结果

1.发表偏倚检验

为确保研究结论的有效性,在进行元分析之前,研究基于107个效应量进行了发表偏倚检验,得到图2所示的漏斗图。由图2可知,纳入的107个效应量相对均匀地分布在漏斗两侧,故初步判断本研究样本不存在发表偏倚。为进一步确认发表偏倚情况,研究还基于CMA3.3工具,采用失安全系数(Fail-Safe Number, Nfs)进行了检验。结果显示,Nfs的值为2892,远大于 $5k+10$ (k 是研究纳入分析的效应量)的容许界限。此外,研究还采纳Egger线性回归检验与Begg秩相关检验两种方法探究了发表偏倚问题,结果显示,Egger线性回归检验的 p 值为0.263,Begg秩相关检验的 p 值为0.083,均大于0.05。以上结果一致表明,本研究中的样本不存在发表偏倚现象,可以进行元分析。

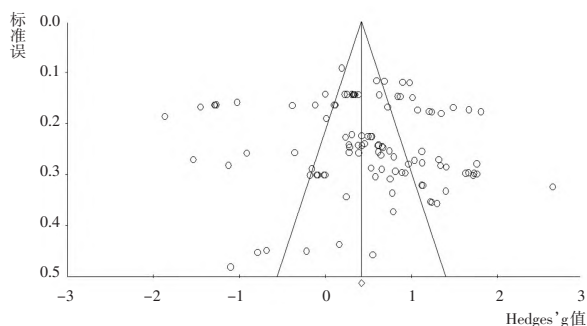


图2 发表偏倚检测漏斗图

2.异质性检验

异质性检验旨在探讨研究个体间差异对因变量的影响情况,其目的在于检验不同研究之间是否具备可整合性或一致性(张天嵩,2015)。本研究运用固定效应和随机效应两种模型对纳入文献进行分析,通过参考 Q 检验和 I^2 检验的统计数值,以确定最合适的模型,保障分析结果的精确性和可信度。异质性检验结果如表2所示,其中 $I^2=92.9\%$

且 $p=0.000<0.1$,表明研究间存在异质性。鉴于此,本研究采用随机效应模型进行分析(Borenstein et al., 2021)。

3.结果分析

(1)生成式人工智能对学生发展的整体效应

效应值(Effect Size, ES)是用于衡量不同组别间的差异大小或变量间关系强度的重要指标,其大小反映了效应的强度。依据Cohen提出的标准(Cohen, 1988),当 $ES<0.2$ 时,视为小效应;当 $0.2\leq ES<0.8$ 时,视为中等效应;当 $ES\geq 0.8$ 时,视为大效应。本研究的合并效应值为0.477($p<0.01$,见表2),属于中等效应范畴,且具备统计学显著性,这表明生成式人工智能在影响学生发展方面发挥了较为明显且正向的推动作用。同时,结合随机效应模型中95%置信区间为[0.330, 0.624]可知,生成式人工智能对学生发展的影响具有一定的稳定性,其产生的效应是实验干预的直接结果,而非偶发性事件。综上,生成式人工智能在影响学生发展方面展现出显著且可量化的效果,其效应值达到了中等程度。

(2)生成式人工智能对学生发展三大维度及其子维度的效应

为了进一步揭示生成式人工智能对学生发展的影响,研究从认知、行为能力、情感态度三个维度展开分析(见表3)。其中,认知维度的总体效应值为0.296($p<0.05$),表明生成式人工智能对学生认知发展的影响较小。具体来看,其对概念理解(Hedges' $g=0.758$, $p<0.05$)存在中等程度的影响,而对学业成绩(Hedges' $g=0.089$, $p>0.05$)和认知能力(Hedges' $g=0.363$, $p>0.05$)并没有显著的促进作用。在行为能力维度,其合并效应值为0.536($p<0.001$),表明生成式人工智能对学生行为能力的发展具有显著的中等程度的影响。具体而言,其对基本技能和能力(Hedges' $g=0.386$, $p<0.05$)、问题解决能力(Hedges' $g=0.607$, $p<0.05$)、创造力(Hedges' $g=0.677$, $p<0.001$)存在中等程度的显著影响,但是对计算思维和算法思维(Hedges' $g=0.281$, $p>0.05$)、批判性思维

表2 异质性检验结果

模型	效应数	合并效应值 (Hedges' g)	95%置信区间		双尾检验		异质性检验			
			下限	上限	Z值	p值	Q	df	p值	I^2
固定效应	107	0.420	0.381	0.458	21.484	0.000	1497.032	106	0.000	92.9%
随机效应	107	0.477	0.330	0.624	6.363	0.000				

表3 生成式人工智能对学生发展三大维度的作用效果分析

维度	分类	效应数	效应值	95%置信区间		合并效应值	95%置信区间	
				下限	上限		下限	上限
认知	概念理解	7	0.758*	0.173	1.342	0.296*	0.022	0.570
	学业成绩	17	0.089	-0.270	0.448			
	认知能力	4	0.363	-0.329	1.055			
行为能力	基本技能和能力	8	0.386*	0.020	0.751	0.536***	0.380	0.692
	计算思维和算法思维	4	0.281	-0.176	0.738			
	批判性思维	4	0.636	-0.052	1.324			
	问题解决能力	4	0.607*	0.010	1.204			
	创造力	10	0.677***	0.458	0.896			
	协作能力	2	0.350	-0.453	1.153			
情感态度	价值认可	7	1.364***	1.120	1.607	0.549***	0.330	0.825
	学习兴趣	4	0.661**	0.157	1.166			
	态度动机	25	0.698***	0.487	0.909			
	自我调节和监控	3	0.320***	0.157	0.483			
	获得感和效能感	8	1.091***	0.595	1.587			

(注: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, 下同。)

(Hedges' $g=0.636$, $p>0.05$) 和协作能力 (Hedges' $g=0.350$, $p>0.05$) 的作用效果并不显著。在情感态度维度, 其合并效应值为0.549 ($p<0.001$), 表明生成式人工智能对学生发展具有显著的中等程度的促进作用。具体来看, 生成式人工智能能够显著提升学生的价值认可度 (Hedges' $g=1.364$, $p<0.001$) 及获得感和效能感 (Hedges' $g=1.091$, $p<0.001$), 对学生的学习兴趣 (Hedges' $g=0.661$, $p<0.01$) 和态度动机 (Hedges' $g=0.698$, $p<0.001$) 具有中等程度的作用效果, 但对自我调节和监控 (Hedges' $g=0.320$, $p<0.001$) 的作用效果较小。

综上, 生成式人工智能能够促进学生在认知、行为能力和情感态度三个维度的正向发展, 表现为其对行为能力和情感态度两个维度具有中等程度的促进作用, 但对认知维度的促进效果偏弱。该研究结果表明, 在运用生成式人工智能进行教学的过程中, 一方面要持续发挥其促进学生行为能力和情感态度的积极作用, 另一方面要通过方法创新来增强

生成式人工智能对学生认知发展的影响。

(3) 实验周期对学生发展的影响

为探索不同实验周期对学生发展的影响, 本研究将实验周期分为短期 (1个月以内)、中期 (1~4个月) 和长期 (4个月及以上) 三个类别进行分析。整体而言, 不同实验周期对学生发展的调节作用存在显著差异 (QB=18.132, $p<0.001$)。其中, 短期 (Hedges' $g=0.352$, $p<0.001$) 和中期 (Hedges' $g=0.869$, $p<0.001$) 实验中生成为式人工智能对学生发展具有显著影响, 而长期 (Hedges' $g=0.243$, $p>0.05$) 实验中生成为式人工智能对学生发展的影响没有达到显著水平 (见表4)。

表4 不同实验周期对学生发展三大维度的影响效果差异

实验周期	维度	效应数	效应值	95%置信区间		异质性检验(I^2)	合并效应值
				下限	上限		
短期	认知	16	-0.100	-0.449	0.249	90.25%	0.352***
	行为能力	15	0.535***	0.384	0.687	66.79%	
	情感态度	23	0.536***	0.285	0.786	86.93%	
中期	认知	9	0.988***	0.574	1.402	81.07%	0.869***
	行为能力	8	0.639**	0.157	1.120	87.10%	
	情感态度	14	0.913***	0.669	1.157	80.46%	
长期	认知	3	0.274*	0.052	0.496	26.04%	0.243
	行为能力	9	0.378	-0.088	0.844	83.39%	
	情感态度	10	0.098	-0.812	1.007	98.51%	

具体来看, 在认知维度, 短期实验 (Hedges' $g=-0.100$, $p>0.05$) 中生成式人工智能对学生的认知发展具有轻微的负面影响但并不显著。随着实验周期的延长, 中期实验中生成为式人工智能对学生认知发展的效应值达到显著水平 (Hedges' $g=0.988$, $p<0.001$), 表明其对学生认知发展具有较大助益。而在长期实验中, 其效应值下降至0.274, 表明其仍然能对学生的认知发展产生正面影响, 但影响却较小。与此同时, 异质性检验结果表明, 短期实验的异质性较高 ($I^2=90.25\%$), 而长期实验的异质性较低 ($I^2=26.04\%$), 这说明短期实验中各研究之间的差异较大, 而长期实验的研究更趋于一致。在行为能力维度, 无论是短期实验 (Hedges' $g=0.535$, $p<0.001$), 还是中期实验 (Hedges' $g=0.639$, $p<0.01$), 生成式人工智能对学生行为能力的发展均产生了显著的正面影响, 而

长期实验 (Hedges' $g=0.378$, $p>0.05$) 中则表现出并不显著的影响。在情感态度维度,中期实验中生成式人工智能对学生情感态度发展的影响效应 (Hedges' $g=0.913$, $p<0.001$) 要大于短期实验 (Hedges' $g=0.536$, $p<0.001$)。然而,在长期实验 (Hedges' $g=0.098$, $p>0.05$) 中,生成式人工智能对学生情感态度的效应值较低且并不显著,而异质性却极高 ($I^2=98.51\%$),这表明长期使用生成式人工智能可能会对学生的情感态度产生复杂的影响。

总体而言,在探索生成式人工智能对学生发展的影响时,短期实验的效果呈现出较为不稳定的现象,尤其在认知维度上甚至出现了轻微的负面影响;中期实验普遍表现出较强的正面效应,不仅在促进学生认知发展上大有裨益,还显著促进了学生行为能力与情感态度的提升;然而,进入长期实验阶段,尽管某些正面影响依旧存在,但其显著性普遍减弱,特别是在情感态度方面,长期应用生成式人工智能可能会对学生的情感态度产生复杂影响并带来潜在挑战。

(4) 样本规模对学生发展的影响

为考察不同样本量情况下生成式人工智能对学生发展的影响差异,研究将样本分为小规模(30人以下)、中等规模(30~100人)、大规模(100人以上)三类别。分析结果显示,不同样本量下生成式人工智能对学生发展的影响效果存在显著差异 (QB=16.681, $p<0.001$) (见表5)。

表5 不同样本规模对学生发展三大维度的影响效果差异

样本量	维度	效应数	效应值	95%置信区间		异质性检验 (I^2)	合并效应值
				下限	上限		
小规模	认知	9	0.295	-0.116	0.705	71.71%	0.386***
	行为能力	13	0.310	-0.068	0.688	81.76%	
	情感态度	9	0.615**	0.191	1.040	67.58%	
中等规模	认知	18	0.316	-0.071	0.703	93.71%	0.483***
	行为能力	14	0.529***	0.345	0.713	68.87%	
	情感态度	36	0.545***	0.218	0.873	96.33%	
大规模	认知	1	0.191*	0.011	0.372	0.00%	0.736***
	行为能力	5	0.872***	0.739	1.004	12.64%	
	情感态度	2	0.642***	0.479	0.804	0.00%	

具体而言,在认知维度,小规模(Hedges' $g=$

0.295, $p>0.05$)和中等规模样本(Hedges' $g=0.316$, $p>0.05$)的研究虽然显示出一定的正面影响,但未达到显著水平,而大规模样本(Hedges' $g=0.191$, $p<0.05$)的研究则显示出较为显著的正面效应。在行为能力维度,大规模 (Hedges' $g=0.872$, $p<0.001$) 和中等规模样本 (Hedges' $g=0.529$, $p<0.001$) 的研究显示出显著的正面效应,但小规模样本 (Hedges' $g=0.310$, $p>0.05$) 的研究却表现出并不显著的正面效应。在情感态度维度,生成式人工智能在小规模 (Hedges' $g=0.615$, $p<0.01$)、中等规模 (Hedges' $g=0.545$, $p<0.001$) 和大规模样本 (Hedges' $g=0.642$, $p<0.001$) 上均表现出显著的正面影响,且效应量均较高,这表明无论样本大小,生成式人工智能对学生情感态度的发展均具有较为明显的促进效果。

综上,生成式人工智能在不同样本规模下对学生发展的影响呈现出一定的差异性。在认知维度,尽管各样本量均表现出一定的正面影响,但相较于其他维度,其影响效应相对较小,且仅在大规模样本研究中达到显著性。在行为能力维度,随着样本量的增加,生成式人工智能的正面效应逐渐增强,且大规模样本研究展现出最为显著的正面影响。在情感态度维度,无论样本规模大小,生成式人工智能均表现出明显的正向影响,且在不同样本量中均保持一致性和显著性。

(5) 教育阶段对学生发展的影响

整体而言,生成式人工智能对基础教育 (Hedges' $g=0.659$, $p<0.001$) 的影响效应大于对高等教育 (Hedges' $g=0.393$, $p<0.001$) 的影响效应,且不同教育阶段的生成式人工智能对学生发展的整体影响不存在显著差异 (QB=3.612, $p>0.05$)。

由表6可知,在认知维度,生成式人工智能对基础教育阶段学生的认知提升具有较强的正面影响 (Hedges' $g=0.781$, $p<0.001$);相比之下,其对高等教育阶段学生的认知影响较弱且不显著 (Hedges' $g=0.095$, $p>0.05$),但高等教育阶段表现出的异质性较高 ($I^2=90.95\%$),说明生成式人工智能在高等教育阶段的应用过程中可能受到更多不稳定因素的影响。在行为能力维度,生成式人工智能对基础教育阶段 (Hedges' $g=0.414$, $p<0.001$) 和高等教育阶段 (Hedges' $g=0.554$, $p<0.001$) 的学生发展都表现出显著的正面影响。在情感态度维度,生成式人工智能对基础教育

(Hedges' $g=0.753$, $p<0.001$) 和高等教育阶段 (Hedges' $g=0.470$, $p<0.01$) 的学生发展均产生了积极效应, 且对基础教育的影响效果更佳。

表6 不同教育阶段对学生发展三大维度的影响效果差异

教育阶段	维度	效应数	效应值	95%置信区间		异质性检验(I ²)	合并效应值
				下限	上限		
基础教育	认知	8	0.781***	0.341	1.220	87.86%	0.659***
	行为能力	10	0.414***	0.290	0.537	16.27%	
	情感态度	14	0.753***	0.386	1.120	92.76%	
高等教育	认知	20	0.095	-0.232	0.423	90.95%	0.393***
	行为能力	22	0.554***	0.322	0.786	83.83%	
	情感态度	33	0.470**	0.119	0.821	95.92%	

综上所述,生成式人工智能对基础教育阶段学生的认知、行为能力和情感态度的影响整体上要强于高等教育阶段。基础教育阶段学生在认知和情感态度上更易受到生成式人工智能的影响,而生成式人工智能对高等教育阶段学生的行为能力具有更显著的正面影响,但在认知和情感态度上的影响则较为不一致。这些差异可能与不同教育阶段学生的认知发展水平、学习动机和个体差异有关。

四、研究结论与启示

1. 研究结论

第一,生成式人工智能对学生的认知、行为能力和情感态度三个维度的发展均具有正向促进作用。整体来看,生成式人工智能可以在中等程度上促进学生的发展。从学生发展的三大维度来看,生成式人工智能对学生的认知、行为能力和情感态度均产生了显著的积极影响。其中,行为能力和情感态度方面的效应值较高,属于中等效应,而在认知方面的效应值略低,属于较小效应,但三者均具有统计显著性。

具体而言,在认知层面,生成式人工智能能够加深学生对概念的理解,但在提升认知能力及学业成绩方面,其效果相对有限,这一发现与先前研究的结论相符 (Al Kahf et al., 2023; Escalante et al., 2023; Canonigo, 2024)。这可能因为在人机互动场景下,生成式人工智能快速生成内容的特点大大缩短了学生在问答环节中的自主思考和探索时间 (Civit et al., 2024),进而影响了学生的知识建构及认知能力发

展,由此导致其在促进学生认知发展层面的成效不够显著。在行为能力层面,生成式人工智能除了能提高学生的基本技能和能力外,还能有效培养学生的问题解决能力和创造力等高阶思维能力,这与 Essien 等人的研究结论相一致 (Essien et al., 2024)。可能的原因在于,生成式人工智能所生成的内容并非绝对准确无误,仍需要学生用辩证与审慎的态度去分析与评估,因而有助于促进其问题解决能力的发展。在情感态度层面,生成式人工智能对学生的价值认可、学习兴趣、态度动机、自我调节和监控、获得感和效能感方面都存在正向显著影响。这可能是因为生成式人工智能倾向于“表扬和鼓励”的机制,有别于教师直接指出问题并提供改进建议的传统反馈。换言之,生成式人工智能善于通过积极的方式引导学生,学生在接收到正面反馈后,其学习兴趣与自我效能感更易被激发。然而,生成式人工智能在自我调节和监控方面却表现出较弱的效应,原因可能是此类工具大多采取即问即答模式,缺乏内置的监督和强化机制,这种模式对学生的自我调控能力要求较高,若学生把控不好则容易对其产生依赖 (Kasneci et al., 2023)。因此,后续研究应不断深入研发符合学生发展规律的教育大模型。

第二,生成式人工智能对学生的认知、行为能力、情感态度发展具有促进作用,且受实验周期、样本规模、教育阶段等调节变量的影响。

在实验周期上,中期实验对学生发展的促进效果最强,短期实验次之,长期实验最弱。这可能是由于生成式人工智能虽具有强大的文本处理与分析能力,但初学者并不熟悉此类工具,因而短期内未能充分发挥其效用 (郑泉, 2024)。随着时间的推移,学生逐渐掌握了此类工具的使用方法,并将其有效融入学习过程中,因而也产生了积极的影响。但在长期使用过程中,学生可能存在滥用或误用等行为,并逐渐对工具产生依赖,因此又可能削弱学生独立思考问题的能力 (钟柏昌等, 2024)。

在样本规模上,研究发现其在学生认知发展方面的调节作用相对有限,而在推动学生行为能力和情感态度发展方面却展现出较强的调节效果,在中等规模及大规模样本研究中的影响效果尤为突出。具体来看,就行为能力和情感态度而言,大规模样本的促进效果要显著优于小规模样本和中等规模样本。而且大规模样本对学生发展的各个方面均具有显著的正向影响,其中对行为能力的促进效果最

强,对情感态度的影响次之,对认知的影响最弱。可能的原因在于,在小规模教学过程中,学生面临同伴反馈稀缺和教师及时指导不足的问题,再加上生成式人工智能所提供的反馈具有随机性,且难以确保输出内容的质量,因而可能导致学生在缺乏系统性指导的情况下,不加甄别地接受工具输出的反馈。相比之下,在大规模教学情境中,社会互动可以显著增强学生的学习效果(Vygotsky, 1978, pp.79-91),而且学生的行为和情感态度也更易受到同伴行为的影响。

在教育阶段上,相较于高等教育,生成式人工智能在基础教育阶段对学生的认知和情感态度发展具有更为显著的中等程度的影响。这是因为生成式人工智能凭借其庞大的知识库与极快的响应速度,能够针对概念理解和知识点内容提供答案及实例,方便学生理解,而这些学习内容恰恰与基础教育的知识积累和学业成绩相关。相比之下,高等教育的内容更为复杂多元,仅凭掌握个别概念难以全面理解其内容,更无法培养学生的实践能力并促进其高阶思维的发展。有鉴于此,探索与不同教育阶段相匹配的教学模式、教育大模型及教学智能体等,对未来教育的高质量发展具有重大意义。

2.研究启示

第一,创新融合教学模式,研发教育专属大模型。一方面,可通过模拟人类在回答问题时的逻辑推理步骤(Wei et al., 2022),或采用自问自答的教学策略(Lee et al., 2024),或通过问题链的形式引导学生逐步完成学习,为学生提供更多的思考空间,并根据学生回答的结果提供针对性的反馈指导。另一方面,建议研发符合我国国情和学生认知发展规律的教育专属大模型。此类模型可通过设置知识强化机制及测验题目,帮助学生在理解概念的同时有效内化知识,同时需严格把关知识反馈的准确性,以确保学生掌握的内容科学合理。此外,教师要引导学生学会批判性地看待人工智能的反馈和建议,帮助他们理解人工智能的局限性,特别是在创新和复杂问题解决方面。教师还应营造一个有吸引力和激励性的环境,将生成式人工智能视为一种辅助性教学手段(Gamage et al., 2023),借助其高效性,持续激发学生的学习热情和内在动力(Crawford et al., 2023),进而增强学生的自我效能感。

第二,加强工具使用培训,预防对技术的过度依赖。一方面,在应用生成式人工智能之前应对教

师和学生进行必要的培训,如在教学过程中分阶段设定学习目标,从初级操作逐步过渡到指令语句再到熟练应用,帮助教师和学生逐步掌握工具功能,以避免因不熟悉工具使用而影响学习效果。另一方面,在实际教学中教师应关注技术依赖可能带来的潜在风险(Habib et al., 2024),注重引导学生建立良好的技术使用习惯(郑永红等, 2024),鼓励学生在使用人工智能前先进行自我思考,将工具反馈作为一种学习参考而非唯一答案,保持学习的主动性,以避免在长期使用过程中过度依赖技术工具而丧失自主学习和独立思考的能力。此外,相关教育管理部门还应出台生成式人工智能的规范化制度,清晰界定人工智能工具的角色及应用准则,有效防范工具的滥用,确保其合理、有序地融入教育实践。

第三,注重组织群体研讨,构建“师一生一机”多元反馈机制。一方面,要通过引导群体讨论,营造浓厚的学习氛围。已有研究表明,学生在同一时空下的集体学习能显著增强其社会临场感(胡勇, 2013)。因此,教师应构建一种促进思维启迪与互动驱动的人机协同群体学习机制,不断激发学生的问题探索欲望,推动多视角的集体讨论与观点交流,从而更有效地促进学生的全面发展。另一方面,要构建“师一生一机”多元反馈机制。生成式人工智能更依赖语言模式而不是逻辑推理,致使其生成的内容具有一定的随机性和不可预测性(Joshi et al., 2024)。为弥补这个缺陷,教师应及时为学生提供额外的学习支架。例如,教师可通过个别辅导、及时反馈来弥补生成式人工智能的不足,以确保学生的学习需求能够得到充分满足。此外,还可采用小组合作模式深化学习成效。例如,教师可以实施分组教学,合理设计生成式人工智能辅助教学的任务,促进学生之间的高频互动。这不仅能避免学生过度依赖技术,还有助于学生在合作中培养互助精神,并激发其创新思维。

第四,全面融合数据分析,不断深化情境教学。一方面,生成式人工智能通过整合数据分析结果,可以针对不同问题为学生提供个性化的反馈,实现类人互动(Lu et al., 2024)。例如,在课堂教学中,生成式人工智能通过迅速解析学生的学习数据,可以实现即时反馈,有助于教师灵活调整教学策略,推动学生深度学习。另一方面,生成式人工智能通过生成与情境相关的反馈,并根据学生的反应及时调整教学策略来支持学生进行个性化学习。

这符合维果茨基的社会文化理论,即强调将学生引导至其“最近发展区”(Vygotsky, 1978, pp.84-91)。此外,还应开发定制化的领域模型以增强其适配性。该类模型需整合学科专业知识与实际应用场景,确保与课程内容紧密对接,从而有效解决生成内容与教育需求不匹配的问题。例如,在医学教育中,生成式人工智能可以通过模拟病例分析过程来帮助学生理解复杂的诊断流程,并结合学生的反馈情况为其提供个性化指导,从而不断提升其学习的深度。

参考文献:

- [1]和文斌,范占江,郑浩等(2024).ChatGPT 赋能 CSCL:未来图景与突围路径[J].现代教育技术,34(4):37-46.
- [2]胡勇(2013).在线学习过程中的社会临场感与不同网络学习效果之间的关系初探[J].电化教育研究,34(2):47-51.
- [3]柯清超,米桥伟,鲍婷婷(2024).生成式人工智能在基础教育领域的应用:机遇、风险与对策[J].现代教育技术,34(9):5-13.
- [4]王映朝,周开济,兰春梅等(2018).考试绩效自评量表的初步编制[J].心理技术与应用,6(3):177-192.
- [5]谢静,陈梦娇,钱佳(2023).课后服务如何影响中小学生学习——基于国外31项实验与准实验研究的元分析[J].教育研究与实验,(5):126-136.
- [6]张绒(2023).生成式人工智能技术对教育领域的影响——关于 ChatGPT 的专访[J].电化教育研究,44(2): 5-14.
- [7]张天嵩(2015).高级 Meta 分析方法:基于 Stata 实现[M].上海:复旦大学出版社:304-306.
- [8]郑泉(2024).生成式人工智能的知识生产与传播范式变革及应对[J].自然辩证法研究,40(3):74-82.
- [9]郑永红,王辰飞,张务伟(2024).生成式人工智能教育应用及其规制[J].中国电化教育,(5):114-119.
- [10]中华人民共和国中央人民政府(2017).国务院关于印发新一代人工智能发展规划的通知[EB/OL].[2024-10-18].https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm.
- [11]钟柏昌,刘晓凡(2024).生成式人工智能何以、以何生成教育[J].电化教育研究,45(10):12-18,27.
- [12]Al Kahf, S., Roux, B., & Clerc, S. et al. (2023). Chatbot-Based Serious Games: A Useful Tool for Training Medical Students? A Randomized Controlled Trial[J]. PLoS One, 18(3):e0278673.
- [13]Anyanwu, C. C., Ononiwu, P. N., & Isiozor, G. N. (2024). Comparative Impact of Whatsapp Chatbot Technology and Glaser's Teaching Approaches on the Academic Performance of Education Economics Students in Tertiary Institutions in Nigeria[J]. Education and Information Technologies, 29:23801-23816.
- [14]Bhatia, A. P., Lambat, A., & Jain, T. (2024). A Comparative Analysis of Conventional and Chat-Generative Pre-Trained Transformer-Assisted Teaching Methods in Undergraduate Dental Education[J]. Cureus, 16(5):e60006.
- [15]Borenstein, M., Hedges, L. V., & Higgins, J. P. T. et al. (2021). Introduction to Meta-Analysis[M]. New Jersey: John Wiley & Sons:3-7,99-101,200-203.
- [16]Canonigo, A. M. (2024). Levering AI to Enhance Students' Conceptual Understanding and Confidence in Mathematics[J]. Journal of Computer Assisted Learning, 40(6): 3215-3229.
- [17]Civit, M., Escalona, M. J., & Cuadrado, F. et al. (2024). Class Integration of ChatGPT and Learning Analytics for Higher Education[J]. Expert Systems, DOI:10.22541/au.170995060.04086473/v1.
- [18]Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences[M]. New York: Erlbaum Associates:20-40.
- [19]Crawford, J., Cowling, M., & Allen, K. A. (2023). Leadership Is Needed for Ethical ChatGPT: Character, Assessment, and Learning Using Artificial Intelligence (AI)[J]. Journal of University Teaching & Learning Practice, 20(3).
- [20]EDUCAUSE (2024). 2024 EDUCAUSE Horizon Report (Teaching and Learning Edition)[EB/OL]. [2024-10-18]. <https://library.educause.edu/resources/2024/5/2024-educause-horizon-report-teaching-and-learning-edition>.
- [21]Escalante, J., Pack, A., & Barrett, A. (2023). AI-Generated Feedback on Writing: Insights into Efficacy and ENL Student Preference[J]. International Journal of Educational Technology in Higher Education, 20:57.
- [22]Essien, A., Bukoye, O. T., & O' Dea, X. et al. (2024). The Influence of AI Text Generators on Critical Thinking Skills in UK Business Schools[J]. Studies in Higher Education, 49(5): 865-882.
- [23]Gamage, K. A., Dehideniya, S. C., & Xu, Z. et al. (2023). ChatGPT and Higher Education Assessments: More Opportunities than Concerns?[J]. Journal of Applied Learning and Teaching, 6(2):358-369.
- [24]Habib, S., Vogel, T., & Anli, X. et al. (2024). How Does Generative Artificial Intelligence Impact Student Creativity?[J]. Journal of Creativity, 34(1):100072.
- [25]Joshi, I., Budhiraja, R., & Dev, H. et al. (2024). ChatGPT in the Classroom: An Analysis of Its Strengths and Weaknesses for Solving Undergraduate Computer Science Questions[C]// Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1, OR, Portland: 625-631.
- [26]Kasneci, E., Seßler, K., & Küchemann, S. et al. (2023). ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education[J]. Learning and Individual

Differences, 103:102274.

[27]Lee, H. Y., Chen, P. H., & Wang, W. S. et al. (2024). Empowering ChatGPT with Guidance Mechanism in Blended Learning: Effect of Self-Regulated Learning, Higher-Order Thinking Skills, and Knowledge Construction[J]. International Journal of Educational Technology in Higher Education, 21:16.

[28]Liao, J., Zhong, L., & Zhe, L. et al. (2024). Scaffolding Computational Thinking with ChatGPT[J]. IEEE Transactions on Learning Technologies, 17:1628–1642.

[29]Lu, Q., Yao, Y., & Xiao, L. et al. (2024). Can ChatGPT Effectively Complement Teacher Assessment of Undergraduate Students' Academic Writing?[J]. Assessment & Evaluation in Higher Education, 49(5):616–633.

[30]UNESCO (2024). AI Competency Framework for

Students[EB/OL]. [2024–10–12]. <https://unesdoc.unesco.org/ark:/48223/pf0000391105>.

[31]Vygotsky, L. S. (1978). Mind in Society: The Development of Higher Psychological Processes[M]. Cambridge, MA: Harvard University Press.

[32]Wei, J., Wang, X., & Schuurmans, D. et al. (2022). Chain of Thought Prompting Elicits Reasoning in Large Language Models[C]// Proceedings of the 36th International Conference on Neural Information Processing Systems (NIPS' 22). NY, Red Hook: Curran Associates Inc.:24824–24837.

收稿日期 2024–11–08 责任编辑 刘选

Can Generative Artificial Intelligence Affect the Development of Student

——A Meta-Analysis of 31 Experiments and Quasi-Experiments Studies

HU Qintai, LIANG Xinxian, LIU Yanfan, WANG Shuli

Abstract: Generative Artificial Intelligence (GenAI) is increasingly applied in the field of education, yet its effects on student development remain controversial. A meta-analysis of 31 experimental and quasi-experimental studies reveals a moderate positive impact of GenAI on students' cognitive, behavioral and affective development. Specifically, GenAI can deepen students' understanding of concepts at the cognitive level. At the level of behavioral ability, it can improve students' basic skills and abilities, and effectively cultivate higher-order thinking skills such as problem solving, critical thinking and creativity. On the level of emotional attitude, it can enhance students' value recognition, stimulate their interest, attitude and motivation, improve self-regulation and monitoring, and enhance their sense of gain and efficacy. These effects are moderated by experimental duration, sample size, and educational stage. Medium-term experiments show the strongest promotion of student development, followed by short-term experiments, with long-term experiments exhibiting the weakest effects. While moderate and large sample sizes have limited moderating effects on cognitive development, they demonstrate stronger effects on the development of students' behavioral ability and emotional attitude. GenAI exhibits more significant moderating effects on the development of cognitive and emotional attitude in the stage of basic education. To promote the effective application of GenAI in education and teaching, it is essential to continuously innovate and integrate teaching models to develop specialized large models for education, strengthen training in tool usage to prevent overreliance on technology, foster collaborative group discussions to construct a diverse "teacher-student-machine" feedback mechanism, and comprehensively integrate data analysis to deepen contextual teaching.

Keywords: Generative Artificial Intelligence; Student Development; Cognition; Behavioral Ability; Emotional Attitude; Meta-Analysis