

人工智能技术应用于创造力评估的实证研究系统综述

A Systematic Review of Empirical Studies on the Application of Artificial Intelligence

Technology in Creativity Assessment

边欣民^{1*}, 王蓉¹, 陈鹏¹, 江怡美¹

¹ 首都师范大学教育学院

* 1806656749@qq.com

【摘要】 人工智能技术与教育政策为创造力评估带来机遇。基于近十年 21 篇实证研究的系统综述发现：研究数量持续增长，集中于教育、计算机类期刊，美国研究较丰富，技术呈现智能化、自动化及多模态融合趋势；研究方法以实验与定量为主，应用领域涵盖艺术、编程等，聚焦高等教育。教育理论支撑薄弱，仅间接体现社会学习与建构主义理论，技术实现优先于教育规律的现象显著。建议进一步推动场景化技术整合和多模态数据治理、推动教育人工智能理论与应用的双向协同。未来需深化理论与技术协同机制，强化跨文化比较研究以提升实践价值。

【关键词】 人工智能技术；创造力评估；实证研究；系统性文献综述

Abstract: Artificial intelligence and education policy synergistically enhance creativity assessment. A systematic review of 21 empirical studies (2014-2024) reveals: (1) Growing U.S.-led research in education/computer journals, driven by intelligent automation and multimodal integration trends; (2) Dominance of experimental and quantitative methods in arts/programming fields, focusing on higher education; (3) Limited educational theory integration, with only implicit connections to social learning/constructivism. It is recommended to advance context-driven technological integration and multimodal data governance frameworks, while fostering bidirectional synergy between educational artificial intelligence theories and applied practices. Future directions emphasize theory-technology alignment and cross-cultural comparative studies.

Keywords: Artificial Intelligence, Creativity Assessment, Empirical Research, Systematic Review

1. 引言

人工智能技术（AI）的快速发展促使教育领域发生深刻变革。《中国教育现代化 2035》等政策强调技术与教育深度融合(国务院, 2019)，为教育评价改革提供机遇。创造力作为 21 世纪核心技能(Riadi et al., 2022)，其评估的科学性和普适性关乎教育质量提升。传统方法如自陈问卷、他人评分等存在效度局限(钟柏昌等, 2022)。人工智能技术通过多模态数据处理捕捉创新行为，构建自动化评估与个性化反馈机制，为创造力评估带来新契机。目前关于人工智能在创造力评估中应用的研究已取得一定进展，尚缺乏对其相关文献的系统性梳理。因此，本研究旨在系统梳理人工智能技术在创造力评估领域的应用现状，重点探索以下问题：

- (1) 人工智能技术应用于创造力评估的实证研究主要呈现哪些特征和发展趋势？
- (2) 人工智能技术应用于创造力评估的实证研究有哪些常见研究方法、具体应用领域？
- (3) 人工智能技术应用于创造力评估的实证研究是否应用了教育理论或策略？

2. 研究设计

2.1. 文献检索

本研究采用系统性文献综述方法,选取 Web of Science、IEEE Xplore 和 ScienceDirect3 个数据库进行检索,主题检索条件为(“artificial intelligence” OR “machine learning” OR “deep learning” OR “generative AI” OR “neural networks” OR “ChatGPT”)AND (“creativity assessment” OR “creative evaluation” OR “innovation assessment” OR “divergent thinking” OR “creative thinking” OR “Creativity”)AND (“education” OR “learning” OR “teaching” OR “student performance”)。检索时段覆盖 2014 年 1 月至 2024 年 12 月,通过数据库与“滚雪球”方式共获得文献 944 篇。

2.2. 文献纳入和排除标准

为确保文献与研究主题的高相关性,本研究制定了以下纳入和排除标准:(1)剔除重复论文,仅保留无重复的期刊论文;(2)仅纳入实证研究;(3)仅保留主题明确涉及人工智能技术在创造力评估中的应用的文献;(4)聚焦教育领域内的研究。

2.3. 文献筛选与编码

本研究遵循 PRISMA 规范开展文献筛选(Page et al., 2021),具体见图 1,通过三轮筛选保证研究精准性:(1)剔除重复与非期刊文献;(2)阅读文献标题和摘要排除 221 篇非实证、主题偏离及非教育领域文献;(3)全文精读排除 12 篇同类不合规文献。最终纳入 21 篇文献进行编码,涵盖作者、年份、期刊、研究类型、方法、应用领域、理论及策略等维度。

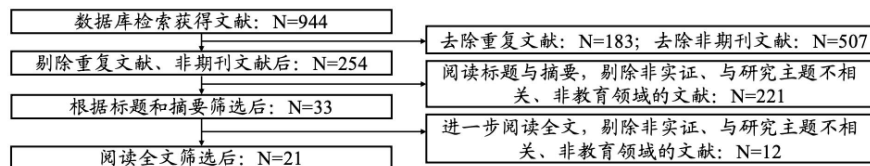


图 1 PRISMA 流程图

3. 结果分析

3.1. 基于文献计量分析的实证研究现状

3.1.1. 文献发表与引用情况

从发表时间来看(图 2a),AI 技术在创造力评估中的应用自 2016 年起得到学术界的持续关注,并呈逐步增加态势。在发表期刊分布方面(图 2b),多数发表于教育类期刊(62%),高于综合类期刊(24%)与计算机科学类(14%),凸显教育评估技术转型需求。在第一作者分布方面,涵盖 8 个国家,分别是美国(10 位)、中国(4 位)、以色列(1 位)、澳大利亚(1 位)、波兰(1 位)、德国(1 位)、韩国(1 位)、印度(1 位)。美国学者是 AI 技术应用于创造力评估实证研究的主要群体。文献引用方面,Organisciak 等人(2023)文章引用次数最高,为 109 次,其余文献低引现象主要源于领域初期理论体系不完善、主题专业化程度高及新成果发表周期短等客观因素。

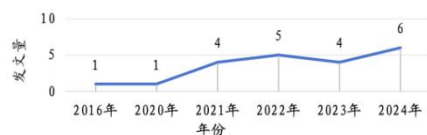


图2a 发文量

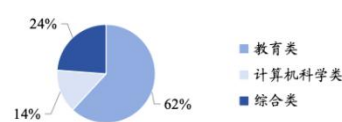


图2b 期刊分布

图 2 文献发表概况

3.1.2. 研究方法特征与应用领域

从研究范式来看(图 3a),当前以实验研究为主(18 篇),辅以调查研究(6 篇)和案例研究(3 篇),数据分析方法(图 3b)主要采用定量分析(17 篇),体现技术验证的实证

倾向。研究领域聚焦艺术（3 篇）等学科，研究层次方面（图 3c），高等教育（8 篇）因评估目标明确成为核心场景。成人教育受技术适配性制约仅 3 篇，暴露研究场景分布失衡问题。

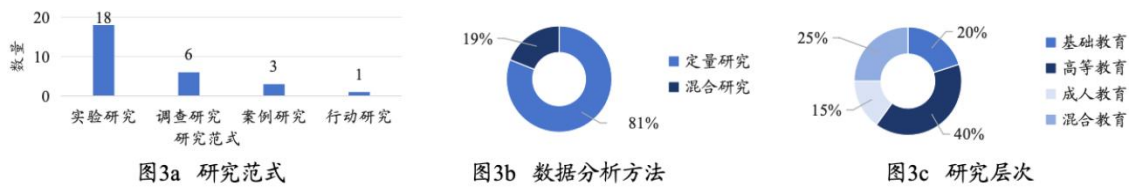


图 3 研究方法特征

3.2. 人工智能技术应用于创造力评估的实证研究趋势

3.2.1. 技术融合与评估效能的协同创新

AI 技术正通过多元路径融入创造力评估。已有研究通过整合神经网络与层次分析法(Fan et al., 2022)、多模态解析技术(Kovalkov et al., 2021), 突破传统评估的维度局限。TCT-DP 图像评分系统(Cropley et al., 2024)及文本创造力深度语义模型(Sung et al., 2022)等自动化工具的应用, 实现评估效率与资源利用率的提升。

3.2.2. 智能化与个性化在创造力评估中的深度融合

AI 在创造力评估中呈现智能化与个性化融合趋势。Fan 等(2022)通过分析学习行为与偏好数据构建个性化学习路径, 在艺术教学中实现差异化资源推荐; Hadas 等(2024)利用大语言模型的上下文感知能力实现创造力回答的智能评分, 其创新点识别技术的持续优化能有效提升评估智能化与个性化水平。

3.3. 人工智能技术应用于创造力评估的教育理论与策略运用

教育理论层面, 现有研究虽未直接援引教育理论, 但实践映射出社会学习、建构主义等理论内核。Ali 等(2021)通过儿童-机器人交互实验中的行为模仿机制验证社会学习理论; Hadas 等(2024)构建的大语言模型对话环境则体现建构主义学习特征; Sung 等(2022)开发的 C-CRAFT 评估系统基于多元智能理论实现创造力多维诊断。教学策略层面, 技术融合推动教学策略创新。Hadas 等(2024)结合启发式教学与生成反馈突破思维定式; Fang 等(2024)运用物联网与生成对抗网络优化创作决策流程; Lee 等(2024)整合项目式与支架式策略支持创意实践。

4. 研究结论与启示

4.1. 研究结论

第一, 人工智能在创造力评估领域研究自 2016 年起受到关注, 美国学界研究丰富且成果多发表于教育、计算机及综合期刊; 第二, 研究方法以实验和定量分析为主, 聚焦艺术教育、多学科融合等场景, 研究对象主要聚焦高等教育, 缺乏成人教育相关研究; 第三, 技术应用呈现智能化、自动化及多模态融合趋势; 第四, 人工智能技术与教育理论、教学策略的深度融合不足, 呈现出技术实现优先于理论支撑的特征, 制约了个性化需求响应能力。

4.2. 研究启示

4.2.1 推动场景化技术整合和多模态数据治理

人工智能技术需融入教育场景, 通过开发智能评估与可视化平台提升师生创新能力, 覆盖考试评价、智慧校园等场景。针对资源分配不均, 构建开源生态与低代码工具降低应用门槛。深化多模态数据融合, 通过动态权重算法解决数据异构与模态失衡, 以“技术降本-数据提质”推动评价精准化与系统智能化, 形成教育数字化转型的可持续闭环体系。

4.2.2 推动教育人工智能理论与应用的双向协同

教育人工智能需构建理论-技术双向协同体系。开发教育理论驱动的可解释评估工具，利用长期追踪技术验证其对创造力培养的实效；建立适龄化数据伦理标准；通过动态校准平衡技术精准度与教育复杂性，促进人工智能技术与教育规律深度契合。

4.3. 研究局限与展望

本研究存在三方面局限：文献筛选受数据库、语言限制，存在跨文化成果缺失风险；缺乏技术应用实证数据验证；技术迭代速度与区域评价标准差异可能削弱结论时效性及普适性。未来可扩展文献检索范围，融合技术案例解析 AI 效能；深化东西方教育场景比较研究，提升成果的实践参考价值。

参考文献

- 国务院. (2019). 中共中央、国务院印发《中国教育现代化 2035》. 新华社.
https://www.gov.cn/xinwen/2019-02/23/content_5367987.htm
- 钟柏昌, 龚佳欣. (2022). 学生创新能力评价:核心要素、问题与展望——基于中文核心期刊论文的系统综述. 中国远程教育, 9, 34 - 43, 68.
- Ali, S., Park, H. W., & Breazeal, C. (2021). A social robot' s influence on children' s figural creativity during gameplay. *International Journal of Child-Computer Interaction*, 28, 100234.
<https://doi.org/10.1016/j.ijcci.2020.100234>
- Cropley, D. H., Theurer, C., Mathijssen, A. C. S., & Marrone, R. L. (2024). Fit-For-Purpose Creativity Assessment: Automatic Scoring of the Test of Creative Thinking - Drawing Production (TCT-DP). *Creativity Research Journal*, 1 - 16.
<https://doi.org/10.1080/10400419.2024.2339667>
- Fan, X., & Zhong, X. (2022). Artificial intelligence-based creative thinking skill analysis model using human - computer interaction in art design teaching. *Computers and Electrical Engineering*, 100, 107957. <https://doi.org/10.1016/j.compeleceng.2022.107957>
- Hadas, E., & HersHKovitz, A. (2024). Using large language models to evaluate alternative uses task flexibility score. *Thinking Skills and Creativity*, 52, 101549.
<https://doi.org/10.1016/j.tsc.2024.101549>
- Kovalkov, A., Paasen, B., Segal, A., Pinkwart, N., & Gal, K. (2021). Automatic Creativity Measurement in Scratch Programs Across Modalities. *IEEE Transactions on Learning Technologies*, 14(6), 740 - 753. <https://doi.org/10.1109/TLT.2022.3144442>
- 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.
- Riadi, B., Perdana, R., Prasetya, R. A., & Prayogi, R. (2022). *The Empowerment of Critical and Creative Thinking (CCT) Skills in Indonesian Language Learning: A Case Study of Online Learning in Secondary School During the COVID-19 Pandemic*. 85 - 90.
<https://doi.org/10.2991/assehr.k.220102.011>
- Sung, Y.-T., Cheng, H.-H., Tseng, H.-C., Chang, K.-E., & Lin, S.-Y. (2022). Construction and validation of a computerized creativity assessment tool with automated scoring based on deep-learning techniques. *Psychology of Aesthetics Creativity and The Arts*.
<https://doi.org/10.1037/aca0000450>