

人工智能赋能教育评价改革研究——以 ChatGPT 为例

Research on the Reform of Educational Evaluation Empowered by Artificial Intelligence: A

Case Study of ChatGPT

杨谨瑜^{1*}, 熊芳庆¹¹ 华南师范大学教育信息技术学院

*1758370764@qq.com

【摘要】 在数字化教育浪潮中，传统教育评价机制的弊端日益突出，而人工智能技术为其变革带来希望，ChatGPT 的出现更是引发广泛关注。本研究以 ChatGPT 为切入点，深入探究其对教育评价改革的影响。从结果、过程、增值、综合评价四个维度探索创新路径，针对改革面临的挑战，提出提升评价主体素养、重塑价值导向、构建保障体系等应对策略，助力二者融合，推动构建科学全面的教育评价体系。

【关键词】 人工智能；ChatGPT；教育评价；评价改革

Abstract: In the digital education era, the flaws of traditional education evaluation are more obvious. AI technology offers hope for reform, and ChatGPT has drawn much attention. This study focuses on ChatGPT to explore its influence on education evaluation reform. It looks for innovative ways in outcome, process, value-added, and comprehensive evaluations. Facing reform challenges, it suggests enhancing evaluators' literacy, reshaping value orientation, and building a security-ethics system, aiming to integrate ChatGPT and education evaluation for a better evaluation system.

Keywords: Artificial Intelligence, ChatGPT, Educational Evaluation, Evaluation Reform

1. 引言

教育评价是教育质量的“指挥棒”，深刻影响教育改革走向。在数字化浪潮中，传统以考试为主的评价机制弊端凸显，无法满足个性化与终身学习需求。人工智能技术的兴起为教育评价革新带来机遇，正重塑教育评价生态，助力解决教育过度功利化和破除“五唯”问题。ChatGPT 作为人工智能技术发展历程中的一座新里程碑，有望打破教育评价变革的技术瓶颈，重塑评价范式和价值理念。但目前如何将其优势转化为推动教育评价优化的动力仍需研究。本研究聚焦 ChatGPT 赋能教育评价变革，探索其创新路径、挑战及应对策略，为二者协同发展提供参考。

2. 文献综述

当前，人工智能与教育评价的研究主要集中在两大领域（彭波等人，2021）。理论层面，郭元祥等人（2023）提出通过革新评价观念、创新机制及推进数字化转型来推动教育评价改革；吴林静等人（2023）倡导利用 AI 技术达成“学评融合”，以促进学生全面发展与元认知能力提升；李毅等人（2024）明确 ChatGPT 如何驱动教育评价在理念、主体、内容和方式上的变革。在技术实践层面，科大讯飞凭借人工智能开发的在线智能评卷系统，大幅提升高等教育纸笔测试评估效率与准确性；Jobin 等人（2019）基于机器学习构建智能系统，能自动评

估学习成效并预测趋势；吴立宝等人（2021）指出人工智能可采集课堂多模态数据，应用于课堂多种分析及教学评价体系构建，为课堂教学评价提供新视角。综上所述，人工智能在教育评价领域已从理论走向实践，生成式人工智能有望重塑教育评价模式与价值文化，改变教育评价的整体格局（曹渡帆等人，1993）。然而，现有研究对生成式人工智能赋能教育评价改革缺乏系统探讨。因此，本研究聚焦 ChatGPT，深入剖析其推动教育评价变革创新路径、面临挑战及应对策略，以促进二者协同发展。

3.ChatGPT 赋能教育评价改革的评价路径创新

《深化新时代教育评价改革总体方案》明确指出了结果、过程、增值、综合评价四种评价类型。在 ChatGPT 等生成式人工智能的支持下，这四种评价类型也将得到全面赋能，让结果评价更科学、过程评价更智慧、增值评价更便捷、综合评价更完善。

3.1. 结果评价：精准衡量成效，驱动目标达成

传统结果评价多依靠标准化测试或量化打分，存在明显弊端。标准化测试和评估量表的制定未充分顾及个体差异，统一的评价标准难以精准衡量其孩子的实际发展水平。在内容分析时，它能精准分析学生的考试作答、课程结业论文等。例如面对历史事件分析论文，可快速判定学生对历史知识要点的掌握情况，还能剖析论证逻辑与观点创新性。在学术论文评估时，ChatGPT 可识别引用规范与否，判断研究方法运用是否恰当，并对比同类成果衡量论文价值，为导师评价提供客观参考。同时，依托智能考试评卷系统，ChatGPT 能实现自动命题、组卷、批阅，并进行精准的数据与分析，深度衡量了教育成效，为结果评价的提升与发展提供坚实保障。

3.2. 过程评价：智慧监控进度，优化实施路径

传统过程性评价以人为预设标准评判教学与学习，评价主观色彩浓，评价者的个人倾向易使结果产生偏差。ChatGPT 在过程监控中扮演着智能助手角色。在线上学习场景，它实时接入平台，即时捕捉与深度分析学生各种学习行为。例如，学生线上学习时，ChatGPT 发现大量学生反复观看某一知识点且提问集中，迅速判断该章节为难点并反馈，助力教学。反过来学生也可以进行个性化地学习，例如在超星学习通的人工智能模块上，学生可以根据自己的学习进度和兴趣，通过知识图谱自主选择学习路径，直接点击自己薄弱的知识点，进行定向学习。在小组合作项目中，它能剖析成员交流记录，评估任务执行等情况。如通过分析学生的线上讨论记录，清晰呈现成员表现，为教师评价提供客观依据。此外，借助可穿戴设备、传感器及非接触式感知技术等多种数据收集工具采集多模态数据并整合多模态数据，能够更加智慧地优化过程性评价。

3.3. 增值评价：探索价值增长，激发创新潜力

增值评价关注学生学业进步幅度，重视学习起点与历程。在基础构建环节，ChatGPT 运用智能采集技术收集学生多维度纵向数据。如借助 ChatGPT 收集学生三年来的考试成绩、课堂表现、社团活动参与情况等数据，精准勾勒学生成长轨迹。在模型构建环节，ChatGPT 依据教育场景需求定制专属模型，并能够兼顾学生个体差别，深度挖掘学生的增值潜力，构建贴合的评价模型。在结果呈现环节，ChatGPT 将学生成长信息可视化。例如 ChatGPT 把学生学科成绩提升、综合素质发展等数据转化为折线图、雷达图，家长和教师能直观了解学生成长节奏，摆脱单纯分数比较，聚焦学生自身成长突破，为教育决策提供直观依据，开辟学生学习价值增长新路径。

3.4. 综合评价：构建全面体系，实现多元发展

ChatGPT 能够整合多维度评价数据，为学生绘制全景画像。在数据采集环节，ChatGPT 借助大数据与区块链技术搭建统一平台。该技术能够驱动校园内的智能设备，全方位采集学生课堂互动、作业完成情况、社团活动参与度等数据。通过整合，避免了传统评价中仅靠成绩评定的片面性，为全面评价提供丰富、真实的数据支撑。深入分析阶段，ChatGPT 运用数据挖掘与学习分析技术全面剖析学生综合发展态势，例如可以利用 ChatGPT 分析学生多学期课程成绩、实践项目成果及科研参与情况，清晰揭示评价结果共性与个性，精准预判发展潜力与未来走向，实现从单一知识考核迈向多元能力评估跨越。在结果呈现方面，ChatGPT 能够为每个学生生成涵盖学业成绩走势、兴趣特长发展的可视化画像，家长和教师能直观了解学生综合表现，便于因材施教，助力学生多元发展。

4.ChatGPT 赋能教育评价改革面临的挑战

4.1. 主体智能素养不足

ChatGPT 支持的先进教育评价模式，要求评价主体具备高水平智能素养，如人机协同和批判性思维能力。然而，当前中小学教师智能教育知识与技能不足，实际应用生疏（郑智勇等人，2023）。而且，不同年龄层教师智能素养差异大，中老年教师难以借助 ChatGPT 开展教学评价，青年教师相对熟练，加剧了数字鸿沟。另外，由于 ChatGPT 技术新、颠覆性强，学生、教师和教育行政管理人员对其认识有限，这进一步导致评价主体智能素养欠缺，严重阻碍了 ChatGPT 在教育评价改革中的应用，急需解决。

4.2. 技术至上人文衰减

ChatGPT 给教育评价带来机遇，却也存在过度依赖风险，会削弱评价主体地位。在实际操作中，人们因惰性或盲目崇技，过度依赖其模板和方案，减少对教育本质的思考。教育评价需要人文关怀，而 ChatGPT 缺乏情感感知和道德判断，无法给予情感关怀和个性化引导。更严重的是，过度依赖会使评价内容失衡，人们易忽视学生灵感、理想信念和创造性思维等关键要素，导致教育评价偏离核心宗旨，陷入“异化”困境。

4.3. 安全伦理双重隐忧

ChatGPT 在教育评价应用中面临数据安全与伦理隐患。数据安全上，它在更新时收集海量数据，增加泄露风险，且系统可能有漏洞，易遭黑客攻击，导致师生信息泄露，影响教学秩序。伦理方面，训练数据繁杂，可能使生成内容有歧视或偏见，影响对学生的公正评价，损害学生心理健康。此外，它还可能被诱导生成违规内容，让学生借此作弊，破坏教育评价公平性与真实性。若不解决，将影响人才培养质量。

5.ChatGPT 赋能教育评价改革的应对策略

5.1. 智赋于人，培育铸魂：全方位提升评价主体智能素养

ChatGPT 的广泛应用对教育评价主体智能素养构成挑战，提升教育主体智能素养刻不容缓。精准化培训是解决关键。构建分层培训体系，为中老年教师设置 ChatGPT 基础课程，给青年教师等开展进阶培训。开发专属学习资源库，涵盖多样学习资料，鼓励参与在线学习社区。学生培养方面，基础教育渗透人工智能知识，高等教育深化相关课程。在职教师通过继续教育与校本培训实现常态化提升。同时，建立智能素养评价指标体系，定期评估并与绩效考核、学业评定挂钩，对表现突出的师生予以表彰，如设智能教育创新奖，激励提升素养与创新应用。

5.2. 技术为翼，人文为魂：重塑教育评价价值导向

在教育评价领域，强化人文教育理念、合理运用 ChatGPT 等人工智能工具极为重要。强化人文理念，重点是提升教师人文素养，通过教育哲学等专题培训，让教师领会教育人文本质，在教学及运用 ChatGPT 时融入人文元素。评价学生作品关注创新思维、情感表达及文化理解，利用 ChatGPT 时突出人文视角。同时，设置批判性思维训练环节，教师引导学生质疑、验证 ChatGPT 答案，审视其输出，避免盲目依赖，实现人文教育与人工智能工具的有机融合。

5.3. 伦理为基，安全为盾：构建安全伦理保障体系

ChatGPT 等人工智能技术在教育应用中，数据安全与伦理问题亟待解决。数据安全方面，技术上，升级防护措施，用加密算法保障数据保密，设多重身份验证防未授权访问，强化备份恢复。管理上，制定规范，明确数据全流程规则，设监督小组审计。伦理规范与监管机制也需完善，多方联合制定准则，规范教育评价应用，提供操作指南。建立动态监管与反馈机制，设监管机构审查数据与评价过程，搭建反馈渠道收集问题，依反馈调整策略，推动 ChatGPT 在教育评价中健康应用，助力教育良性发展。

参考文献

- 李毅,郑鹏宇,&张婷.(2024).ChatGPT 赋能教育评价变革的现实前提、作用机理及实践路径.现代远距离教育(03),9-17.doi:10.13927/j.cnki.yuan.20240612.001.
- 吴立宝,曹雅楠,&曹一鸣.(2021).人工智能赋能课堂教学评价改革与技术实现的框架构建.中国电化教育(5),94-101.
- 吴林静,高喻,&涂凤娇,等.(2023).数据驱动的多维融合在线协作会话可计算模型及应用研究.现代教育技术,33(04),101-110.
- 郑智勇,肖林,&宋乃庆.(2023).中小学教师智能教育素养的现状分析与提升路径---基于 11703 位教师的实证调查.教育发展研究(4),30-39.
- 郭元祥,叶飞,&刘志军,等.(2023).中国式基础教育现代化的理论思考与实践路向.中国电化教育(01),68-80.
- 曹渡帆,&朱德全.(2023).新时代教育评价改革的数字正义何以可能.电化教育研究,44(12),21-27+34.
- 彭波,王伟清,&余剑波.(2021).人工智能视域下教育评价改革何以可能(06),1-15.
- Jobin,A.,Ienca,M.,&Vayena,E.(2019).The Global Landscape of AI Ethics Guidelines.Nature Machine Intelligence(1),389-399.