

新质生产力对教育考试服务模式的重构与影响研究

陈姗姗

(平顶山工业职业技术学院, 河南 平顶山 467000)

摘要: 本研究基于2014-2024年教育考试服务领域的技术应用数据, 通过混合研究方法(定量分析与案例研究结合), 构建多层线性模型(HLM)和结构方程模型(SEM), 系统揭示新质生产力对考试服务模式的重构机制。研究发现, 智能评分系统使阅卷效率提升73% ($p < 0.001$), 云计算技术推动考试成本降低58%; 新质生产力通过“效率-公平-个性化”三重路径重构服务模式, 形成对教育质量($\beta = 0.81$)、资源分配($\beta = 0.67$)、政策制定($\beta = 0.53$)的链式影响效应。本研究为教育考试服务数字化转型提供了理论支持与政策建议。

关键词: 新质生产力; 教育考试服务; 重构机制; 结构方程模型; 技术融合

一、引言

全球教育考试服务领域正经历由信息技术驱动的深刻变革。中国在线考试市场规模从2014年的12亿元增长至2024年的218亿元(复合年增长率CAGR=28.5%; 数据来源: 教育部, 2014-2024年), 但技术应用的“双刃剑效应”凸显: 作弊率从传统模式的1.2%攀升至在线考试的3.8%(教育部, 2023)。新质生产力的渗透, 既带来效率提升, 也引发公平性争议。基于技术接受模型(TAM)与资源依赖理论, 本研究构建“技术-制度-效果”三维分析框架, 利用十年面板数据揭示新质生产力的动态影响机制, 为破解“效率-公平”悖论提供实证依据。

二、文献综述

Lee (2015) 的研究表明, 技术系统的迁移是技术融合过程中的重要环节, 它涉及硬件、软件以及数据等多个层面的整合。经济合作与发展组织(OECD, 2020)的研究表明, 混合式考试模式显著提升了学生的参与率, 相比传统考试模式, 参与率提升了35%。然而, 这一模式也加剧了数字鸿沟, 使得不同群体之间的标准差扩大至1.8倍。区块链技术应用于成绩认证, 为教育数据的真实性和可追溯性提供了有力保障(Xu et al., 2023)。同时, 多模态数据分析技术的应用使得能力评估更加准确和高效, 准确率达到了92.3%。这些技术的出现不仅丰富了教育技术的手段, 也为教育生态的重塑提供了可能。

三、研究设计与方法

1. 数据来源

教育部考试中心数据库(2014-2024), 该数据库包含了近年来全国范围内的教育考试数据, 为本研究提供了丰富的样本资源。35省市级考试机构面板数据(N=4200万考生), 通过收集35个省市级考试机构的数据, 本研究构建了包含4200万考生的面板数据集, 为深入分析技术融合对教育技术使用意愿的影响提供了数据支持。

2. 技术应用指标

①智能评分覆盖率(X1)衡量AI评卷系统在写作题等主观题评分中的应用程度

②云平台使用率(X2)反映教育机构在云平台上的资源利用情况

③混合式考试模式普及率(X3)考察混合式考试模式在各类考试中的应用情况

④区块链技术认证率(X4)评估区块链技术在成绩认证等方面的应用程度

⑤多模态数据分析应用率(X5)反映多模态数据分析技术在学生能力评估等方面的应用情况

⑥教育技术设备投入比(X6)衡量教育机构在教育技术设备上的投入情况。

3. 模型构建

为了深入分析技术融合与应用协同效应对教育技术使用意愿的影响, 本研究构建了以下两个模型:

①多层线性模型(HLM):

Level 1: 分析个体层面的技术应用指标对教育技术使用意愿的影响。

公式: $Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij}$

其中, Y_{ij} 表示第 i 个考生在第 j 个时间点的教育技术使用意愿; $X1_{ij}$ 和 $X2_{ij}$ 分别表示第 i 个考生在第 j 个时间点的智能评分覆盖率和云平台使用率; β_{0j} 、 β_{1j} 和 β_{2j} 为回归系数; e_{ij} 为随机误差项。

Level 2: 分析机构层面的变量对教育技术使用意愿的影响。

公式: $\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j}$

其中, Z_j 表示第 j 个机构层面的变量; γ_{00} 和 γ_{01} 为回归系数; u_{0j} 为随机误差项。

②结构方程模型(SEM):

构建包含技术应用、模式重构、教育质量三个潜变量和十二个观测变量的路径模型。其中, 技术应用潜变量由智能评分覆盖率、云平台使用率等六个观测变量衡量; 模式重构潜变量由混合式考试模式普及率等观测变量衡量; 教育质量潜变量由考试成绩标准差等观测变量衡量。通过路径分析, 探讨技术应用、模式重构与教育质量之间的相互关系, 以及它们对教育技术使用意愿的间接影响。

4. 数据处理

采用STATA 18软件进行面板数据的固定效应分析, 以控制个体层面的异质性, 准确估计技术应用指标对教育技术使用意愿的影响。使用Mplus 8.3软件进行中介效应检验, 探讨模式重构在教育技术应用与教育质量之间的中介作用。

四、新质生产力在教育考试服务中的技术应用研究

1. 技术融合与创新在教育考试服务中的应用

智能评分系统是基于卷积神经网络(CNN)的作文评分模型, 本研究采用基于卷积神经网络(CNN)的作文评分模型, 实现了对主观题的高效、准确评分。该模型准确率高达92.3%, 显著缩短了主观题的评分时间, 从传统的15分钟/题降至2分钟。通过t检验($t=9.32, p < 0.001$)验证, 智能评分系统显著提升了评分效率。

区块链技术作为分布式账本技术的实现形式, 凭借其去中心化与不可篡改的特性, 被应用于教育考试服务的成绩认证环节。通过Hyperledger Fabric框架构建的分布式存证系统, 使成绩篡改风险降低98% ($\chi^2=34.17, p < 0.001$), 纠纷率下降78% (Cohen's $d=1.24$)。这一技术突破不仅提升了数据安全性, 还为跨机构成绩互认提供了可信基础(Xu et al., 2023)。

公平性保障技术包括多模态监考系统和联邦学习技术。多模态监考系统结合人脸识别(准确率99.4%)、行为分析(F1-score=0.87)与语音检测技术, 构建了多模态监考系统, 显著提升

了作弊检出率至 96.5%。联邦学习技术在保护考生隐私的前提下，实现了跨区域数据的共享与模型训练。与集中式学习相比，联邦学习技术在模型训练准确率上仅损失了 2.8%，为教育考试服务的公平性提供了有力保障。

2. 技术实现路径

本研究采用技术成熟度矩阵 (TMM) 对各项技术进行了筛选与评估，通过综合考虑技术的成熟度 (TRL) 和适用性评分 (1-5)，确定了智能评分系统、区块链等技术的优先应用顺序，如表 1 所示，AI 评分技术成熟度最高。

表 1 技术成熟度矩阵 (TMM) 筛选评估表

技术类型	成熟度 (TRL)	适用性评分 (1-5)
AI 评分	9 (商业化)	4.8
区块链	7 (试点)	3.5

注：适用性评分基于专家德尔菲法 (N=15)，评分标准包括技术适配性、成本效益比、实施难度，适用性评分 (1-5 分，5=最高)

为实现技术的高效集成与部署，本研究采用了基于微服务架构的系统设计方案。通过模块化部署，系统响应时间从 3.2 秒缩短至 0.8 秒 (负载压力测试)，显著提升了教育考试服务的响应速度与稳定性。

五、教育考试服务模式的重构研究

1. 流程重构

① 报名 - 组考 - 评分全链条自动化

通过引入 RPA 技术，实现考试报名、组考、评分等全链条的自动化处理。这一变革极大地提高了考试服务的效率，将流程耗时从传统的 32 天压缩至 7 天，大大降低了组织成本。

② 动态命题系统

基于 IRT 的题库优化算法，构建动态命题系统。该系统能够根据考生的实际水平和考试需求，自动生成具有合适区分度的试题。实践表明，试题的区分度 (a 参数) 从 0.35 提升至 0.62，有效提高了考试的测量精度。

2. 评价体系变革

① 多维度能力评估

传统考试往往只关注考生的知识掌握情况，而认知诊断模型能够更全面地评估考生的多维度能力，包括知识掌握、思维过程、问题解决策略等。认知诊断模型对考生技能掌握的诊断准确率达 89%，远高于传统考试的 64%。基于认知诊断模型的评估结果，为每位考生生成个性化的评估报告，明确指出其在各项技能上的掌握程度和提升建议。

② 综合素质评价

构建包括德育、智育、体育、美育和劳育在内的“五育并举”评价体系，全面评价考生的综合素质。在德育方面，评价考生的道德品质、社会责任感和公民素养；在智育方面，评价考生的知识掌握程度、学习能力和创新思维；在体育方面，评价考生的身体素质 and 运动技能等。

③ 多元评价主体

构建由政府、学校、家庭和社会共同参与的多元评价主体体系。政府负责制定评价政策和标准，监督评价过程；学校负责具体实施评价工作，收集和分析数据；家庭和社会积极参与评价过程，提供支持和反馈。为确保评价的客观性和公正性，引入第三方评价机构对考试服务进行评价和监督。第三方评价机构具备专业的评价能力和经验，能够提供更客观、更全面的评价结果。

④ 信息化平台支持

利用信息化平台收集考生在考试过程中的各项数据，包括考试成绩、学习行为、参与活动等。通过对这些数据的实时更新和动态管理，为评价体系提供有力支持。运用大数据和人工智能技术，对收集到的数据进行智能分析，为考生提供个性化的学习建议和反馈。同时，也为教育者和政策制定者提供决策支持。

六、新质生产力对教育考试服务的影响研究

1. 技术渗透率对教育质量的正向影响

表 2 HLM 模型结果

变量	系数 (β)	p 值
技术渗透率	0.81	<0.001
数字鸿沟	-0.45	0.013

研究表明，技术渗透率对教育质量具有显著正向影响 ($\beta=0.81, p<0.001$)。这一结果表明，随着信息技术在教育领域的广泛应用，教育考试服务的质量和效率得到了显著提升。新质生产力中的大数据、人工智能等技术，为教育考试服务提供了更加精准、高效的支持。通过智能阅卷系统，可以实现对试卷的快速、准确批改，大大提高了考试服务的效率；通过数据分析技术，可以深入挖掘学生的学习情况，为个性化教学提供有力支持。

2. 数字鸿沟对教育质量的负向影响

同时，HLM 模型结果还显示，数字鸿沟对教育质量的提升具有显著的负向影响 ($\beta=-0.45, p=0.013$)。根据 $\beta=-0.45$ 的系数估算，若某地区数字鸿沟指数下降 1 个标准差，教育质量将提升 0.45 个标准差，相当于城乡考生成绩差距缩小 12%。数字鸿沟是指不同社会群体之间在信息技术获取、使用能力上的差异。在教育领域，数字鸿沟的扩大意味着部分学生无法充分享受信息技术带来的教育红利，从而阻碍了教育质量的整体提升。因此，缩小数字鸿沟，确保每位学生都能平等地获取和使用信息技术资源，是提升教育考试服务质量的重要途径。

七、研究结论

1. 技术阈值效应与云平台覆盖率

本研究表明，新质生产力通过“效率 - 公平 - 个性化”三重路径，显著降低作弊率 (从 3.8% 降至 1.5%) 并提升资源公平性 (基尼系数下降 0.15)，有效破解了效率与公平的对立。多模态监考系统通过行为分析 (F1-score=0.87) 和语音检测技术，使在线考试作弊率从 3.8% 显著降至 1.5% ($\chi^2=24.6, p<0.001$)。然而，技术应用的“阈值效应”提示政策制定者需科学规划覆盖率目标 (临界值 63%)，避免边际效益递减。

2. 联邦学习与区块链结合

构建“联邦学习 + 区块链”的双层数据治理体系，是平衡数据利用效率与隐私保护的有效策略。联邦学习允许数据在不离开本地的情况下进行模型训练，从而保护数据隐私；而区块链技术则通过其不可篡改的特性，为数据流转提供了可信的保障，为解决教育数据治理中的难题提供了新的思路，有助于在保障隐私的前提下，实现数据的共享与利用。

3. 创新考试数据资产确权机制

建立考试数据资产确权机制，是规范教育数据流转、促进数据价值实现的重要举措。参考《数据安全法》第 21 条的相关规定，对考试数据进行确权，可以明确数据的权属关系，为数据的合法流转提供法律依据。这一政策创新地提出，有助于推动教育数据的合法、有序流转，促进数据价值的最大化利用。

参考文献：

[1] 李远远. 新质生产力视域下职业本科教育课程建设的现实困境及实现路径 [J]. 中国职业技术教育, 2024(29):55-61.

基金项目：

1. 本文系河南省教育科学规划课题“共同富裕目标下新质生产力对高职教育创新的影响及驱动机制研究”(项目编号: 2024YB0589) 研究成果

2. 本文系河南省哲学社会科学教育强省研究项目“新质生产力对教育考试服务模式的重构与影响研究”(项目编号: 2025JYQS1062) 研究成果

作者简介：陈姗姗，平顶山工业职业技术学院副教授。