

智能批改系统对提升小学高段作文教学效率的实证分析

胡东元

宁夏中卫市海原县西安镇中心小学 宁夏 中卫 755208

摘要: 随着人工智能技术的快速发展,智能批改系统(Intelligent Grading System, IGS)在基础教育领域的应用日益广泛。本研究聚焦于小学高段(四至六年级)作文教学,通过准实验设计,探讨智能批改系统对提升教师教学效率和学生写作能力的实际效果。研究选取S市两所小学共240名学生及12名语文教师作为研究对象,实验组使用“智文通”智能批改系统辅助教学,对照组采用传统人工批改方式。研究周期为一学期(16周),通过前后测写作成绩、教师工作日志、课堂观察记录及师生访谈等多种数据收集手段,综合评估智能批改系统的应用成效。研究发现:实验组学生在语言规范性、结构完整性等维度上显著优于对照组($p < 0.05$);教师批改时间平均减少42.3%,教学反馈频率提高2.1倍;学生对写作兴趣和修改意愿显著增强。然而,系统在情感表达、创意评价等方面仍存在局限。研究建议在“人机协同”模式下优化智能批改系统的功能设计,并加强教师数字素养培训,以实现作文教学效率与质量的双重提升。

关键词: 智能批改系统;小学高段;作文教学;教学效率;人机协同;实证研究

引言

作文教学是小学语文课程的重要组成部分,不仅关乎学生的语言表达能力,更涉及思维发展、情感体验与文化理解。《义务教育语文课程标准(2022年版)》明确指出,小学高段(四至六年级)应“能写简单的记实作文和想象作文,内容具体,感情真实”,并强调“重视引导学生在自我修改和相互修改的过程中提高写作能力”。然而,现实教学中,教师面临批改负担重、反馈滞后、个性化指导不足等困境^[1]。近年来,以自然语言处理(NLP)和机器学习为基础的智能批改系统逐渐进入教育场景。此类系统可自动识别语法错误、词汇搭配、篇章结构等问题,并提供即时反馈。然而,其在小学高段作文教学中的实际效果尚缺乏系统性实证研究。尤其在中文语境下,作文评价不仅关注语言形式,更强调思想情感与文化内涵,这对AI系统的语义理解能力提出更高要求。因此,本研究旨在通过实证方法,探究智能批改系统是否能够有效提升小学高段作文教学效率,并分析其优势与局限,为教育数字化转型提供实践依据。

1 文献综述

1.1 智能批改系统的发展与类型

智能批改系统起源于20世纪90年代的英语作文自动评分(Automated Essay Scoring, AES)研究,代表性系统如ETS开发的e-rater。近年来,随着深度学习技术的发展,中文智能批改系统如“批改网”“作文纸条”“智学网”等逐步成熟。根据功能划分,现有系统可分为三类:(1)基础语法检查型(如错别字、标点纠正);(2)结构分析型(如段落逻辑、首尾呼应);(3)内容评价型(如

主题契合度、情感倾向)。目前多数系统仍以第一、二类为主,第三类因涉及主观判断而准确率较低。

1.2 教学效率的内涵与测量

教学效率通常指单位时间内教学投入与产出的比值。在作文教学中,可从教师维度(批改时间、反馈质量)和学生维度(写作成绩、修改行为、学习动机)综合衡量。已有研究多关注单一指标,缺乏多维整合分析。

2 研究设计

2.1 研究问题

本研究围绕三个核心问题展开:

智能批改系统是否能显著提升小学高段学生的作文成绩?

该系统如何影响教师的批改负担与教学反馈行为?

师生对智能批改系统的接受度与使用体验如何?

2.2 研究对象

研究选取S市两所办学水平相当的公办小学(A校与B校)作为实验基地。每所学校选取四年级两个平行班,共计4个班级,240名学生参与研究,其中男生124人,女生116人,平均年龄为10.3岁。参与教师共12人,教龄介于5至20年之间,均具有小学高段语文教学经验。研究采用准实验设计,将A校设为实验组,B校为对照组,确保两组在前测作文成绩、教师职称结构、班级学生人数等关键变量上无统计学显著差异($p > 0.05$),从而增强内部效度。

2.3 研究工具

本研究采用多元工具组合以确保数据三角验证。核心干预工具为“智文通”V3.2智能批改系统,该系统由合

作教育科技公司提供,具备错别字自动识别、病句修正建议、段落逻辑分析、主题偏离预警等功能,并以可视化方式(如颜色高亮、弹窗提示、星级评分)向学生呈现反馈^[2]。为科学评估学生写作水平,研究团队依据《小学语文课程标准(2022年版)》及本地教研室通用标准,编制了包含五个维度的作文评分量表,总分50分,具体涵盖内容切题、语言规范、结构完整、情感真实与创意表达。该量表经三位资深教研员独立试评,Cronbach's α 系数达到0.87,表明具有良好的内部一致性。此外,研究还通过教师工作日志持续记录每周批改耗时、反馈形式及教学调整情况,并辅以半结构化访谈提纲,分别面向教师与学生收集深度质性数据。

2.4 实验流程

整个研究周期为一学期,共16周。第1周进行前测,所有学生完成统一命题作文《一次难忘的经历》,由三位教研员在不知分组情况下进行盲评。第2至15周为干预阶段:实验组学生通过平台在线提交作文,系统自动生成初评报告,教师在此基础上进行二次批注,重点补充系

统难以判断的情感表达、立意深度等内容,形成“人机双反馈”;学生随后查看综合反馈并进行修改。对照组则维持传统教学流程,即学生纸质提交作文,教师全人工批阅,并通过面批或集体讲评方式进行反馈。第16周进行后测,题目为《我学会了_____》,评分方式与前测一致。期间,研究团队同步收集前后测成绩、12位教师共192份工作日志、8次课堂观察记录以及24份师生访谈转录文本。

2.5 数据分析方法

定量数据采用SPSS 26.0软件进行处理。首先通过独立样本t检验比较实验组与对照组在后测各维度得分上的差异;其次运用配对样本t检验分析各组内部前后测的变化幅度。对于质性数据,则采用主题分析法(Thematic Analysis),由两名研究者独立编码,经讨论达成共识后提炼核心主题,以揭示师生对智能批改系统的深层体验与认知。

3 研究结果

3.1 学生作文成绩分析

表1: 学生作文成绩

维度	实验组 (M±SD) 前测	实验组 后测	对照组 前测	对照组 后测	组间后测差值 (t, p)
内容切题	7.2±1.1	8.5±0.9	7.1±1.2	7.8±1.0	t = 4.32, p < 0.001
语言规范	6.8±1.3	8.9±0.8	6.9±1.4	7.4±1.1	t = 7.65, p < 0.001
结构完整	7.0±1.2	8.7±0.9	7.0±1.3	7.6±1.0	t = 6.88, p < 0.001
情感真实	7.5±1.0	8.2±0.9	7.4±1.1	8.0±1.0	t = 1.98, p = 0.049
创意表达	6.5±1.4	7.3±1.2	6.6±1.5	7.1±1.3	t = 1.21, p = 0.227
总分	35.0±4.2	41.6±3.1	35.0±4.5	37.9±3.8	t = 6.03, p < 0.001

结果显示,实验组在语言规范、结构完整、内容切题三个维度提升显著($p < 0.001$),情感真实维度略有提升($p = 0.049$),创意表达无显著差异($p = 0.227$)。总分提升6.6分,显著高于对照组的2.9分。

3.2 教师教学效率变化

教师工作日志数据显示,智能批改系统显著优化了教师的时间分配。实验组教师平均每篇作文的批改时间从前测阶段的12.5分钟降至7.2分钟,降幅达42.3%。这一节省出的时间被重新投入到更高价值的教学活动中:83%的实验组教师表示,他们得以设计更多微写作任务、录制个性化讲解视频,或组织小组互评活动。同时,反馈频率大幅提升——实验组学生平均每篇作文获得2.3次反馈(系统初评+教师精评),而对照组仅为1.1次。这种高频、即时的反馈机制不仅加速了学生的认知调整,也增强了其写作过程中的参与感与掌控感。

3.3 师生使用体验

质性访谈揭示了师生对智能批改系统的复杂态度。多数教师肯定其在“减负增效”方面的价值,一位教师提到:“系统帮我快速揪出错别字和病句,我可以把精力集中在引导学生如何把‘妈妈送伞’这样的日常小事写出新意和温度。”然而,也有教师指出系统建议过于模板化,例如频繁提示“建议加入心理描写”,却未提供具体方法或范例,导致学生知其然不知其所以然。学生方面,低年级学生普遍对系统的视觉化反馈(如红色波浪线、星级图标)感到新奇且易于理解,认为“改起来特别快”。但部分学生也表达了困惑,如一位学生坦言:“系统说我‘主题不突出’,但我真的不知道怎么改才叫突出。”这反映出当前系统在提供可操作性指导方面仍有提升空间。

4 讨论

4.1 智能批改系统的优势:效率与规范性提升

本研究证实,智能批改系统能显著提升作文教学效

率,尤其在语言规范与结构分析等客观维度。这与NLP技术在词法、句法层面的成熟应用一致。系统提供的即时反馈符合“形成性评价”理念,促使学生在“提交—反馈—修改”循环中主动建构知识,契合建构主义学习理论。

4.2 局限:主观评价的“天花板”

系统在情感真实与创意表达维度效果有限,反映出当前AI在语义深层理解上的瓶颈。作文不仅是语言技能,更是生命体验的表达^[3]。如一位学生写道:“爸爸的手像枯树枝”,系统标记为“比喻不当”,而教师解读为“体现对父亲辛劳的心疼”。此类文化隐喻与情感张力,尚难被算法精准捕捉。

4.3 “人机协同”模式的必要性

研究支持“AI做初筛,教师做升华”的协同路径。教师可将节省的时间用于:(1)设计高阶写作任务(如辩论稿、书信体);(2)开展个性化面批;(3)组织同伴互评。这种分工既发挥AI的效率优势,又保留教师的人文引领作用。

4.4 数字鸿沟与伦理风险

需警惕技术依赖导致的“评价窄化”——学生可能为迎合系统评分标准而放弃个性化表达。此外,农村学校若缺乏设备与网络,可能加剧教育不公。因此,技术推广需配套基础设施与教师培训。

5 建议与研究局限

5.1 实践建议

为充分发挥智能批改系统的潜力,建议从系统、教师与机制三个层面协同推进。在系统设计上,应进一步丰富“情感关键词库”与“创意表达案例库”,提升语义理解深度,并提供分层修改建议,满足不同水平学生的需求。在教师发展方面,亟需开展“AI辅助写作教学”专题研修,帮助教师掌握数据解读、反馈整合与教学重构的能力,同时鼓励建立“AI误判案例集”用于集体教研^[4]。在教学机制上,学校应将智能批改纳入校本教研体系,定期评估其适用性,并开发“人机协同”教案模板,明确双方职责

边界。此外,政府应加大对薄弱学校的技术投入,开发轻量化离线版本,确保教育公平。

5.2 研究局限与展望

本研究存在一定局限。样本局限于城市优质学校,未来可拓展至城乡结合部或乡村小学,考察不同教育生态下的适用性;干预周期仅为一学期,长期效应有待追踪;研究未区分不同作文体裁(如记叙文、说明文、应用文)下系统的表现差异。后续研究可结合眼动追踪、学习分析等前沿技术,深入探究学生与AI互动过程中的认知负荷与策略选择,为智能教育的精细化发展提供更坚实的科学依据。

6 结语

本研究通过为期一学期的准实验,系统验证了智能批改系统在小学高段语文教学中的应用价值。结果表明,该系统能有效提升学生的语言规范性与结构完整性,显著减轻教师批改负担,并增强学生修改意愿。然而,其在情感表达与创意评价等主观维度上存在明显局限,无法替代教师的专业判断与人文引领。理想的作文教学不应是“人被机器取代”,而应是“人与机器共舞”,构建一个融合智能初评、教师精评、学生自评与同伴互评的多元反馈生态。

参考文献

- [1]徐海宾.小学语文作文智能批改系统在人工智能技术支撑下的应用效能分析[J].中华活页文选(教师版),2025,(20):40-42.
- [2]明世超,潘昌雷.智能批改在小学高段语文作文教学中的应用[J].新课程教学(电子版),2023,(02):160-161.
- [3]黄凤仪.AI赋能下的小学语文作文“教—学—评”一体化开展实践[J].读写算,2025,(30):58-60.
- [4]张智敏.人工智能助力小学语文作文教学:策略与实践案例分析[C]//广东教育学会.广东教育学会2025年度学术讨论会论文集(一).茂名市化州市南盛街道中心学校;,2025:1751-1754.