

人机协同下小学语文习作智能评价系统的实施路径探索

李梦丹, 胡思言, 张荣丽

成都大学师范学院, 四川 成都

收稿日期: 2026年9月3日; 录用日期: 2026年10月6日; 发布日期: 2026年10月12日

摘 要

本研究旨在为小学语文习作智能评价系统构建一条人机协同的实践路径。现有智能评价系统在情感态度、价值导向等柔性维度上识别能力有限, 也缺少清晰的人机协作流程。为此, 论文提出“机评-人评-协同反馈”三阶递进模型: 机器承担错别字、语法等结构化任务; 教师负责创意、情感等需要经验判断的内容; 协同反馈整合双方结果, 通过分层呈现和“引导-修改-再评-跟进”的循环, 推动学生持续修改与提升。在此基础上, 模型允许教师根据教学目标、学生水平和文体类型灵活调整人机任务权重与边界, 以增强对不同教学情境的适应性; 同时, 研究也讨论了模型在实际推行中可能面临的挑战, 包括教师技术培训、不同信息化条件学校的差异化适配、学生使用引导策略及负面反馈预案, 从技术、教学和实施层面提出相应保障措施。研究从这些层面论证了该路径的可行性, 为一线教学提供具有弹性和可操作性的实践参照。

关键词

人机协同, 小学语文, 习作评价, 智能评价系统, 实施路径

Exploring the Implementation Path of an Intelligent Evaluation System for Chinese Composition in Primary Schools in the Context of Human-Machine Collaboration

Mengdan Li, Siyan Hu, Rongli Zhang

College of Teachers, Chengdu University, Chengdu Sichuan

Received: September 3, 2026; accepted: October 6, 2026; published: October 12, 2026

Abstract

This study aims to construct a practical pathway for human-machine collaboration in the intelligent evaluation system for Chinese composition writing in primary schools. Existing intelligent evaluation systems have limited capacity to identify soft dimensions such as emotional attitudes and value orientations, and they also lack a clear collaborative workflow between humans and machines. To address this, the paper proposes a three-stage progressive model of “machine evaluation - human evaluation - collaborative feedback”: machines handle structured tasks such as spelling and grammar checking; teachers take charge of content that requires experiential judgment, including creativity and emotional expression; and collaborative feedback integrates the results from both sides, using layered presentation and a cycle of “guidance - revision - re-evaluation - follow-up” to promote students’ continuous revision and improvement. On this basis, the model allows teachers to flexibly adjust the task weights and boundaries between humans and machines according to teaching objectives, student proficiency levels, and genre types, thereby enhancing its adaptability to diverse teaching contexts. At the same time, the study also discusses potential challenges in actual implementation, including teacher training, differentiated adaptations for schools with varying levels of information technology infrastructure, strategies for guiding students in using the system, and contingency plans for negative student responses, and it proposes corresponding safeguards from technical, pedagogical, and implementation perspectives. The study demonstrates the feasibility of this pathway from these dimensions, offering a flexible and operable practical reference for front-line teaching.

Keywords

Human-Machine Collaboration, Chinese Composition in Primary Schools, Composition Evaluation, Intelligent Evaluation System, Implementation Path

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 现实困境：习作评价面临的难题

2023 年以来，国内多款智能作文批改系统进入中小学课堂，AI 辅助习作评价正在从概念走向实践。然而，技术入场并未自动解决习作评价的痛点——教师批改负担重、反馈周期长、评价标准主观的问题依然突出[1][2]。现有智能评价系统虽能完成错别字识别、语法纠错等基础工作，却始终难以触及习作评价中更核心的内容：情感是否真实、创意是否独特、价值导向是否正确。

1.1. 系统能做什么、不能做什么

王舒[1]在构建 AIGC 辅助写作技术的“素材收集 - 创作支持 - 评价反思”闭环教学系统的研究中，梳理了多款 AIGC 工具在小学语文习作中的应用情况，指出豆包、文心一言等工具可协助生成教学案例和素材，出口成章、笔神作文等则兼具语法纠错与智能评分功能。从上述工具的实践应用来看，其反馈多为通用性建议，在个性化指导方面仍有不足。吴东超[3]在探讨 AI 赋能教学评价策略时发现，AI 系统在语言纠错上准确率较高，能根据学生的常见错误自动生成针对性训练内容。

系统能做好语言纠错，不等于能做好完整的习作评价。具体来说存在三重矛盾。

第一，评价标准错位。现有系统的设计重心在语言形式层面，对习作中的情感表达和价值取向只能

依赖预设规则来判断。但刘茂勇[2]的研究指出, 习作评价要考虑内容、过程等多重因素, 尤其要体察情感。系统与教师在评价标准上的这种错位, 导致两者评判结果经常不一致, 系统难以给出让教师信服的评定。

第二, 反馈语言成人化。陈燕萍[4]在构建“三轮五步”人机协同评改模式的研究中发现, AI生成的批阅报告内容繁杂, 评语中的术语和表述超出部分小学生的理解水平。

第三, 教师不知如何整合。反馈不可靠且不易理解, 教师对系统的利用走向了两个极端——要么完全不用, 要么直接照搬AI评语作为最终结果。全彩霞[5]在探讨“AI+教师”协同模式下的习作高效指导策略时指出, 部分教师将AI评语直接作为最终评价, 忽视了对学生个性化表达的深度引导。无论哪种做法, 都未实现真正的人机协同。

1.2. 为什么需要清晰的协同路径

上述三重矛盾的共同根源在于: 系统的技术设计没有充分考虑小学语文习作评价的教育性、情境性和发展性特征。传统习作评价本身长期存在批改负担重、反馈滞后、标准主观的困境, 而人机协同恰能部分缓解这些难题: 机器可快速完成语言层面的基础批改, 直接减轻负担、缩短反馈周期; 标准主观的问题仍需教师复核把关。这正体现了人机各自的优势——机器擅长快速、一致地处理可量化任务, 人则擅长情感理解、创意判断和价值引领。但仅有优势互补还不够, 关键是要有一套可操作的路径, 把两种评价机制有序地组织起来、衔接起来。

基于此, 本研究聚焦以下问题: 如何构建一套适合小学语文习作教学的人机协同智能评价实施路径? 具体包括三个层面——人机协同评价的任务边界如何划分? 三阶递进模型的具体操作流程是什么? 该路径在技术、教学和实施条件层面是否具有可行性?

2. 优化路径: 三阶递进模型的设计

2.1. 模型的总框架

本路径的设计以人机协同理论为基础。方海光等[6]指出, 人机协同教育的核心在于人与机器在各自优势领域的动态任务分配与互补协作, 而非简单的替代关系。这一观点为本路径中机评与人评的任务划分提供了直接的理论依据——机器承担可量化的标准化评价任务, 教师负责需要理解和价值判断的柔性评价任务。在此基础上, 方海光等[7]进一步从共生理论的视角揭示了人机协同中人与机器的合作博弈关系, 强调人类教师应处于监督和主导地位, 这与本路径中“教师复核作为最终把关机制”的设计逻辑一致。

路径同时以过程性评价理论和最近发展区理论为支撑。过程性评价理论主张将评价嵌入学习全过程, 通过持续反馈促进学习改进[8]。《义务教育语文课程标准(2022年版)》[9]也明确提出, 过程性评价应“关注学生个体的进步幅度”, 评价要“呈现学生的关键认知表现”。这一理念决定了本路径不是“一评了之”, 而是通过二次评价和个性化跟进形成闭环。最近发展区理论强调反馈的及时性和针对性对学生发展的重要性[10], 并指出学生的能力发展存在两种水平: 独立完成任务的现有水平和在帮助下可能达到的潜在水平, 两者之间的区域即为“最近发展区”[11]。本路径中反馈的分层呈现、修改策略的分层设计, 本质上都是在学生的最近发展区内搭建“支架”, 帮助学生从现有水平迈向潜在水平[12]。

基于上述理论, 本路径对机评和人评任务的划分逻辑是: 机评处理“有明确规则、可量化”的维度, 如错别字识别、语法检测、结构判断; 人评承担“需要理解、判断、价值介入”的维度, 如创意评价、情感判断、价值导向把关。

路径以“机评-人评-协同反馈”三大环节为核心框架, 整体流程(见表1): 学生提交作文→机评初

筛→教师复核→整合反馈→学生修改→二次评价→优化迭代。整个路径遵循“以生为本、人机互补、可操作、可迭代”四项原则。

Table 1. Overall implementation process
表 1. 整体实施流程

第一环节机评环节	第二环节人评环节	第三环节协同反馈环节
学生提交作文→机评初筛	教师复核	整合反馈→学生修改→二次评价→优化迭代

2.2. 环节一：机评——标准化维度的自动化评价

机评环节的定位是“初筛”和“辅助”，由系统完成可量化、有明确规则的标准化评价任务。

机评环节的设计依据是：这类任务具有明确的判断规则和客观的评判标准，可以通过算法实现自动化处理。具体分为四个维度：字词语法维度——系统建立小学阶段常见错别字库(如“在”与“再”、“的”“地”“得”混淆)，通过匹配算法识别并给出正确写法；识别句号、逗号、问号、感叹号等常见标点的缺失或误用；识别成分残缺、搭配不当、语序混乱等语病并提供修改示例。表达质量维度——统计作文中不同词语的总数、计算重复率、标注高级词汇；检测长短句搭配情况、提示“句式单一”风险。篇章结构维度——识别是否有开头段、中间段、结尾段。量化评分维度——按内容、结构、语言、书写规范四个维度分别评分，生成总分和各维度得分。

以上四个维度的核心是规则匹配和统计计算，属于典型的可编程操作，因此由机评承担。

机评案例：以一次五年级记叙文写作为例。学生提交作文后，系统自动完成以下检测：字词层面，检测到“的”“地”“得”混用 3 处，在原文中高亮标注并给出正确写法；语句层面，识别出两个成分残缺的病句，提供修改示例；篇章层面，检测到缺少结尾段，在原文末尾标注“建议补充总结或升华内容”；词汇层面，发现“开心”重复出现 5 次，提示可替换为“愉快”“兴奋”等词汇。系统生成初步评分报告后，教师在此基础上进行复核。

机评结果需经教师复核，不作为最终依据。

2.3. 环节二：人评——柔性化维度的教师复核

人评环节的定位是“复核”和“升华”，由教师对机评结果逐项复核，重点处理机评无法触及的内容。如果说机评解决的是“对不对”的问题，那人评解决的是“好不好”“对不对(价值层面)”的问题。具体从以下三个维度展开。

内容创意方面，教师从新颖性、独特性和“灵气”三个维度判断。新颖性看选材是否跳出常见套路——“扶老奶奶过马路”是套路化表达，“帮同学找到丢失的宠物”则相对新颖；独特性看是否有个人化的观察，如有学生写“妈妈的脚步声能分出轻重缓急，急的是上班要迟到，轻的是我生病时”，这种细节体现了真实的观察；灵气看是否有超出年龄的表达亮点。教师通读全文后，在系统中勾选创意等级并补充评语。

情感态度方面，教师从真实性和感染力两个维度判断。真实性看情感表达是否自然——学生写“我很感动”但说不出原因，属于空洞表达；写“妈妈半夜起来给我盖被子，我看见她眼睛红红的”，则具有真实感。教师对真情实感段落做绿色高亮标注，对虚假空洞内容做灰色标注并提示“可以写得更具体”。

价值导向方面，教师从是非判断和行为引导两个维度把关。是非判断看是否存在不正确的价值观(如“有钱就是了不起”)；行为引导看是否传递积极的行为导向(如“帮助同学是为了加分”需要纠正为“帮助别人自己也会快乐”)。教师对有问题的内容做红色标注并填写引导性评语。

机评结果修正——对机评误判，教师点击“修正”修改评分或标注；对机评漏判，点击“补充”添加发现的问题。系统记录所有修正操作，用于后续优化。教师的复核重点应放在学生“写了什么”和“为什么写”上，而非纠结于语言形式。

真实性和感染力无法通过算法判断，因为两者依赖对生活经验的理解和共情能力；价值导向的判断涉及道德判断和教育引导，必须由教师根据教育目标和社会主义核心价值观进行把关——这正是教师不可替代之处，也是人评环节在整个路径中不可省略的原因。

2.4. 环节三：协同反馈——分层呈现与闭环跟进

协同反馈环节的核心任务是将机评与人评结果整合后有效传递给学生，并引导其完成修改、形成完整闭环。具体来说，这一环节要依次解决三个问题：结果怎么整合？反馈怎么呈现？学生怎么跟进才能真正改进？

第一个问题对应“交接机制”，解决的是两份评价结果如何合并为统一的报告——协同的前提是评价信息要先汇合，否则学生拿到的是两套互不相干的内容。第二个问题对应“反馈的分层呈现”，解决的是合并后的评价如何以学生能理解的方式输出——评价信息需要适合小学生的认知水平。第三个问题对应“修改引导策略”和“二次评价与个性化跟进”，解决的是学生拿到反馈后如何有效修改、修改质量如何保障——这是评价促进发展的最终体现。

反馈的分层呈现。反馈按三个维度分层呈现(见表 2)：

Table 2. Layered feedback presentation rules
表 2. 反馈分层呈现规则

分层维度	第一层	第二层	第三层	第四层
按问题类型	基础层(错别字、标点) ——红色标注 + 正确写法	语言层(病句、用词) ——黄色标注 + 修改建议	结构层(段落衔接、逻辑顺序)——蓝色标注 + 图示说明	内容层(创意、情感) ——绿色标注 + 鼓励性评语
按修改优先级	紧急修改——不修改影响理解	建议修改——修改后表达更流畅	鼓励尝试——不强制，进阶目标	—
按年级水平	低年级——图标 + 简单文字	中年级——图文结合 + 简短说明	高年级——文字说明为主	—

在分层反馈的基础上，还需要解决两部分评价结果的合并问题，避免学生收到两套互不相干的内容。具体做法是：系统完成机评后生成初步评价报告，教师在系统中逐项复核(认同/修正/补充)，系统自动合并机评与教师复核结果生成“双层评价报告”，推送至学生端。标注规则为：绿色标注对应教师认同的问题(建议修改)，橙色标注对应教师补充的问题(重点修改)，红色标注对应教师修正的问题(必改)。当机评与教师判断不一致时，以教师判断为准——教师最了解学生的实际发展水平，能够做出符合具体情境的专业判断。机评结果保留为“参考意见”呈现给学生，标注“机器建议：……老师意见：……”。至此，结果整合的问题解决之后，接下来要面对的是更关键的挑战：学生拿到这份报告后，是否愿意改、能否改好？

修改引导策略。修改是闭环机制中最关键的一环——评价结果只有通过学生的修改行为才能真正转化为写作能力的提升。修改引导策略的设计本质上是在学生的最近发展区内搭建“支架”，帮助学生从现有水平迈向潜在水平。具体策略如下：

第一，聚焦式修改。每次修改聚焦 1~2 类问题，避免一次性面对太多问题产生挫败感。这一策略将

复杂的修改任务分解为可逐步完成的子任务，降低认知负荷。

第二，示例式引导。对每一类问题提供“修改前－修改后”对比示例。例如：原文“我去公园玩很开心”改为“周六下午，我去中山公园玩，开心极了！”，让学生直观感受“具体”和“笼统”的差别。

第三，分层修改任务。基础层学生优先改错别字和标点；进阶层学生尝试改病句和用词；挑战层学生尝试优化结构和内容。系统根据学生水平自动推荐修改重点。

第四，修改记录与展示。制作“修改成长记录表”，记录每次修改的问题类型和数量，定期展示“最佳修改案例”，通过同伴激励激发修改热情。

修改策略的意义在于让学生有了修改的动力和方法，但修改后的质量如何、是否还有进步空间，需要二次评价和跟进机制来回答。

二次评价与个性化跟进。学生修改后重新提交，系统自动对比修改前后变化。教师复核修改质量，分三种情况处理：修改正确的点击“通过”并加分；修改错误的点击“再改”并再次提示；未修改的点击“提醒”，必要时单独沟通。二次评价后进入个性化跟进：对反复出现同类问题的学生，系统推送针对性微练习(3~5分钟)；对进步明显的学生发放“进步之星”勋章；对长期不修改或无提升的学生，教师单独沟通并一对一指导。

通过上述“引导－修改－再评－跟进”的闭环机制，每次评价都有机会转化为学生的实际进步。这一闭环的意义在于：评价不再是一次性的“打分”，而是持续推动学生改进的“引擎”，这正是过程性评价理念在习作教学中的具体落地。

3. 路径的价值与展望

本路径在技术、教学和实施条件三个层面具有可行性。技术层面，批改网、作文评阅系统等现有产品在错别字识别、语法纠错、结构分析和评分等功能上已相对成熟，多数系统支持教师复核、手动修正和二次评价等操作；系统误判的，由教师复核把关；农村学校网络条件有限的，可用离线版或降低频次。教学层面，评价、反馈、修改本就是习作教学的固有环节，本路径只是优化了实现方式，教学结构没有改变；反馈分层设计兼顾了低中高年级的认知差异，各年级均可适用；教师可通过前期培训逐步适应，学生若过度依赖机器反馈，可通过“查看反馈前先自评”或安排“无反馈日”来调节。实施条件层面，教师需配备1台电脑或平板，学生端可使用机房电脑或班级平板(2~3人共用1台亦可)，具备稳定的网络连接，条件不足时可降低使用频次或先选取部分班级试点。

本路径的创新主要体现在四个方面：明确了人机评价的任务边界——可量化、有规则的任务由机评承担，需要理解、判断和价值介入的由人评把关；构建了“引导－修改－再评－跟进”的闭环机制，将修改作为评价的核心目的；提出了交接机制，明确了人机冲突的处理规则——以教师判断为准；设计了分层反馈呈现方式，避免“一刀切”的反馈模式。

本路径适用于小学中高年级(3~6年级)的记叙文、说明文等常规文体教学，低年级使用需简化反馈呈现并降低频次。有基本信息化条件的城镇和农村学校均可适用，诗歌、想象作文等创意类文体需教师更多介入。本研究目前为理论设计，尚未经过大规模实践检验，后续需通过实证研究验证效果，同时也需针对不同智能评价系统的功能差异进行对比验证。

本研究得出以下结论：人机协同评价是小学语文习作教学的可行方向，能够有效缓解传统评价中负担重、反馈慢、标准不一的问题。协同的关键在于明确任务边界——机评负责可量化任务，人评负责需要理解和价值判断的任务。本路径的核心价值在于明确了人机各自的任务边界，同时提供了分层反馈和闭环修改的具体做法。对一线教师的建议是：主动熟悉系统操作，在复核环节坚持主导地位，重视修改引导；对学校的建议是：提供硬件和系统支持，组织培训，将路径纳入教研计划；对系统开发者的建议

是：优化反馈语言的儿童化程度。后续可从三方面深化研究：开展实证研究检验路径效果，探索 AIGC 在习作评价中的更多应用，构建更精细的人机协同评价模型。

基金项目

本项目为成都大学省级大学生创业训练计划项目成果，项目名称为“人机协同背景下小学语文习作智能评价系统的实施路径与反馈效果分析”，项目编号为“S202611079018X”。

参考文献

- [1] 王舒. AIGC 辅助写作技术在小学中段习作教学中的应用——基于统编版小学语文教材的闭环教学系统构建[J]. 辽宁教育, 2025(7): 20-24.
- [2] 刘茂勇. 困境与突围: “习作评价量表”的多维审思[J]. 中小学课堂教学研究, 2022(12): 11-15.
- [3] 吴东超. 多向度习作: AI 赋能教学评价策略探究[J]. 江西教育, 2025(42): 16-19.
- [4] 陈燕萍. 三轮五步: 人机协同推动小学习作评改模式转型[C]//中国陶行知研究会. 陶行知教育思想理论研究与创新实践(2025) (上册). 2025: 382-387.
- [5] 全彩霞. “AI+ 教师”协同模式下学生习作高效指导的策略与实践[J]. 读写算, 2025(36): 166-168.
- [6] 方海光, 孔新梅, 李海芸, 等. 人工智能时代的人机协同教育理论研究[J]. 现代教育技术, 2022, 32(7): 5-13.
- [7] 方海光, 孔新梅, 刘慧薇, 等. 基于共生理论的人机协同教育主体合作博弈及其优化策略研究[J]. 电化教育研究, 2024, 45(1): 21-27.
- [8] 薛蕾. 小学语文教学中过程性评价的应用策略——以习作“我来编童话”为例[J]. 语文新读写, 2024(8): 86-88.
- [9] 中华人民共和国教育部. 义务教育语文课程标准(2022 年版) [S]. 北京: 北京师范大学出版社, 2022.
- [10] 维果茨基. 思维与语言[M]. 李维, 译. 杭州: 浙江教育出版社, 1997: 171.
- [11] 赵钰. 支架式教学法在小学语文阅读课堂中的应用[J]. 小学生(下旬刊), 2026(5): 103-105.
- [12] 李伟. 怎样做评价: 写作反馈阶段学生参与式评价支架的搭建与运用[J]. 福建教育, 2026(10): 73-77.