

人工智能辅助形成性评价应用于高中英语词汇教学的研究

—以外研社必修 3 Unit4 Science and technology 为例

金岚¹ 刘峻利¹

(1.太原师范学院, 山西 晋中 03061)

摘 要:随着 AI 技术发展, 其在教育领域应用渐广, 尤其为解决高中英语词汇教学中记忆效果差、评价单一、反馈滞后等问题提供了可能。本文采用文献分析与理论探讨法, 聚焦 AI 辅助形成性评价在该教学场景的应用, 剖析教学现存问题与形成性评价理论基础, 阐述 AI 与形成性评价的契合点, 并从设计原则、应用模式、实施步骤提出可行策略, 以期为一线教师提供新思路与可操作方案。研究表明, 该模式可增强词汇学习的个性化、及时性与可持续性, 但推广受技术、师资、观念等挑战制约, 需在实践中持续优化。

关键词:人工智能; 形成性评价; 高中英语; 词汇教学; 教学创新

DOI: doi.org/10.70693/rwsk.v1i9.1424

一、引言

随着 AI 技术在教育领域的深度应用, 智能测评、个性化推荐等功能为破解高中英语词汇教学痛点提供了新路径^[1]。高中英语词汇教学既需满足《普通高中英语课程标准 (2017 年版 2020 年修订)》“掌握 3500 个左右核心词汇并灵活运用”的要求, 还需依托“人与自我”“人与社会”“人与自然”主题语境实现意义建构, 但当前教学中, 学生机械背诵导致记忆碎片化、评价侧重终结性测试忽视过程、教师反馈滞后等问题突出^[2]。例如外研社必修 3 Unit4“Science and technology”单元, 学生虽能拼写“artificial intelligence”等词汇, 却难在“科技发展利弊”语境中准确应用, 传统评价亦无法捕捉此类深层问题。

本研究聚焦该单元“innovation, algorithm”等主题词汇群, 采用文献分析与理论探讨法, 梳理 AI 与形成性评价的融合路径, 构建“主题导向的 AI 评价模式”, 旨在解决主题词汇应用评价缺失等痛点, 为一线教师提供“技术+主题”融合的教学方案, 同时丰富 AI 教育应用的学科适配理论, 推动词汇教学从“孤立记忆”转向“主题关联的深度学习”, 助力英语学科核心素养落地。

二、核心概念与理论基础

(一) 核心概念界定

本研究聚焦教育场景中与词汇教学评价、个性化指导直接相关的 AI 技术, 具体包括: 智能测评系统 (自动生成词汇试题、批改并统计错误类型)、个性化推荐技术 (依据学生答题数据推送词根词缀讲解、例句练习等资源)、数据追踪技术 (实时记录词汇记忆量、复习频率等轨迹, 为形成性评价提供动态支撑), 核心是实现“数据采集-分析-反馈-优化”的智能化评价闭环。

本研究中, 形成性评价^[3]是贯穿高中英语词汇教学全流程 (课前、课中、课后)、以促进词汇能力发展为核心的动态评价, 区别于传统终结性测试: 课前通过小测评定位基础, 课中实时反馈调整教学节奏, 课后错题复盘跟踪进步。评价不仅关注记忆准确率, 更分析错误原因 (拼写混淆、语境误用等) 并提供改进建议, 最终实现“以

作者简介:金 岚 (2002-) 女, 硕士 研究方向学科教学 (英语);

刘峻利 (2000-) 女, 硕士 研究方向学科教学 (英语);

通讯作者:金岚

评促学、以评促教”。以《普通高中英语课程标准》为导向, 教学内容涵盖课标 3500 个核心词汇及语境应用、搭配规则、文化内涵等, 目标分三维度: 认知维度(掌握音形义)、应用维度(读写说中正确使用)、素养维度(通过词汇理解中西方文化差异)。词汇教学需与听、说、读、写技能训练结合, 形成性评价指标亦需围绕此综合目标设计。

(二) 理论支撑

建构主义学习理论^[4]认为学习是学生主动建构知识的过程, 受个体经验与情境影响, 这为形成性评价提供核心依据^[5]: 一是评价需关注词汇学习“建构过程”(如理解生词、关联旧词的思维轨迹), 而非仅统计正确率; 二是主张个性化支持, 为 AI 定制差异化评价方案提供支撑。实践中体现为两场景: ①知识建构: AI 针对“Science and technology”单元推“主题词汇联想任务”(如从“environment”延伸“pollution, protect”, 关联“科技助环保”主题), 引导构建词汇网络。②个性化支持: AI 通过初始测评识别短板(如薄弱“科技过程类词汇”), 推送差异化资源(实验步骤短视频)。

自适应学习理论^[6]核心是“依学习者特征动态调整学习内容与路径”, 与 AI 特性高度契合^[7], 为 AI 辅助评价提供逻辑框架: AI 先通过初始测评将学生分基础层、提高层、精通层, 再推适配任务, 且实时分析数据动态调整难度。实践中体现为: 基础层推“词汇认知任务”(如匹配“algorithm-算法”), 提高层推“句型应用任务”(如用“invent”补全句子), 精通层推“短文写作任务”(如写“AI 助环保”短文)。基础层连续 3 次正确率达 90%自动升级, 精通层出错则推语境资料, 实现“评价-学习-再评价”循环。

间隔重复记忆理论^[8]核心是“依遗忘速度动态调整复习间隔”, 新学词汇间隔短、掌握词汇间隔长, 为 AI 评价内容设计提供科学依据。实践中体现为: AI 针对“Science and technology”单元词汇, 结合记忆数据算“最佳复习间隔”——如“biotechnology”首次正确率 60%→1 天后推短文检测, 二次 85%→3 天后推词义匹配, 三次 95%→7 天后推口语练习, 强化记忆持久性。

三、AI 与高中英语词汇形成性评价的契合性

形成性评价以“过程性、个性化、可持续性”为核心特征, AI 在数据处理、动态调整、长期追踪等方面的优势, 可精准满足这些需求, 二者在高中英语词汇教学中深度契合。

(一) 满足过程性需求: 实时追踪与即时反馈

从满足过程性需求来看, 形成性评价的核心诉求是打破“终结性测试”的局限, 聚焦词汇学习的动态过程, AI 技术通过实时数据追踪与即时反馈构建了过程性评价的技术闭环。传统词汇教学中, 教师仅能通过课堂提问、课后作业了解局部情况, 难以捕捉“词汇记忆频率”“错题思考时长”等细节数据, 而 AI 可依托数据采集技术, 精准记录学生词汇学习全轨迹, 比如学生练习“accommodate”时多次漏写中间的“m”, AI 会自动标记该拼写薄弱点并关联后续评价重点, 让评价直接基于过程数据生成结论。同时, 传统评价中教师批改需 1-2 天, 反馈滞后易强化错误记忆, AI 则能在学生完成测试后 10 秒内生成结果, 还会针对错误类型推送解析, 如因词义理解偏差出错时展示多语境例句、因搭配模糊出错时列出常见搭配, 契合形成性评价“及时干预”的需求, 帮助学生在遗忘临界点前修正错误。

(二) 满足个性化需求: 定制方案与指导

在满足个性化需求方面, 高中学生的词汇基础、学习风格存在显著差异, 形成性评价需针对个体差异提供个性化方案, AI 的自适应与数据建模能力恰好能实现这一目标。评价方案定制上, AI 通过初始测评构建学生“词汇能力画像”, 为基础薄弱学生设计侧重拼写、词义匹配的基础认知类评价, 题目难度控制在课标核心词范围, 为基础较好学生设计侧重词汇填空、短文写作应用的评价, 引入课标拓展词与复杂语境, 还会根据学习风格调整形式, 如为视觉型学习者推送图文结合的测试、为听觉型学习者提供听力辨词任务, 解决传统评价“一刀切”的问题。学习指导个性化上, 传统评价中教师反馈多为“加强词汇记忆”等笼统建议, AI 则能结合评价数据定位具体问题并关联资源, 比如学生时态相关词汇应用频繁出错时推送语法微课、对文化相关词汇理解有偏差时推荐中西方文化对比材料, 让形成性评价真正成为“促进学习”的工具。

(三) 满足可持续需求: 档案建立与模型迭代

从满足可持续需求来讲, 形成性评价需伴随学生词汇学习全过程, 实现长期监测与持续优化, AI 的档案管理与模型迭代能力为此提供了技术支撑。学习档案建立上, 传统教学中学生的词汇学习记录分散在作业本、试卷中, 难以系统留存分析, AI 可构建动态更新的“学生词汇成长档案”, 整合“每月测试正确率变化”“薄弱词汇类型

演变”等数据，既为教师提供学情分析依据，也让学生直观看到自身成长，增强学习动力。评价模型迭代上，AI会通过积累的教学数据优化评价体系，分析评价任务的区分度与效度，比如调整表述歧义的题目、优化难精准衡量的评价指标，还能根据课标更新纳入新增主题词汇，形成“数据驱动-模型迭代-效果提升”的循环，确保形成性评价始终与教学目标、学生需求保持同步。

四、AI 辅助形成性评价的应用策略

在高中英语词汇教学中应用 AI 辅助形成性评价，需以科学的设计原则为指引，构建覆盖教学全流程的应用模式，并遵循清晰的实施步骤，才能充分发挥 AI 技术优势，切实提升词汇教学效果。不同学生背景（年龄、性别、水平）在 AI 辅助词汇学习中的表现可能不同，例如 Oman 的一项研究发现这些变量对效果有显著影响^[9]。

（一）应用设计原则：为策略落地划定核心准则

遵守目标导向原则，紧扣“Science and technology”单元主题词汇教学目标，将“掌握主题词汇的音形义”“能在科技语境中灵活应用”“理解科技词汇的文化内涵（如‘sustainable development’的全球共识）”拆解为具体评价指标（如表1），确保 AI 评价任务围绕指标展开。

表1 评价指标

评价维度	具体指标	AI 评价任务示例
词汇认知	主题词汇拼写正确率、词义匹配准确率	拼写“nanotechnology”；匹配“algorithm”与“算法”
		用“experiment”完成“Students _____ in the lab yesterday.”；在“科技环保”短文中补全“data analysis helps _____ pollution.”
词汇应用	主题词汇在句型、短文中的应用正确率	判断“sustainable development only refers to environmental protection”的正误，并说明理由
文化理解	对“sustainable development”等词汇的文化内涵解读准确性	

任务设计上，AI 系统提供多样化选择：学生可按学习节奏选“5-10 题碎片化短测试”或“20-30 题系统性长测试”，也可自主选择文字答题、听力辨词、图片联想等评价形式。结果呈现上，AI 以学生易懂方式展示数据（如折线图呈现词汇正确率变化、标签化标注薄弱点，并开放自主查看错题解析、调整复习计划功能，推动学生从“被动接评”转向“主动参与评价与改进”。

遵守数据安全原则，针对 AI 评价涉及的学生学习数据，需建立严格安全机制：工具选择优先合规平台，仅采集词汇学习相关数据（答题记录、记忆轨迹），不收集无关个人信息。数据采用加密存储，设置分级访问权限（如教师仅查看本班数据），且仅用于词汇学习分析与评价优化，不向第三方泄露。需明确告知学生与家长数据用途及保障措施，获取知情同意以消除隐私顾虑。

遵守实用性原则，AI 评价需贴合教学实际，降低师生负担。工具选择界面简洁、操作便捷的系统：教师端支持“一键布置任务”“自动生成学情报告”，学生端通过微信小程序、轻量化 APP 参与，无需复杂下载。评价频率与时长适配课堂节奏，课前测评≤5 分钟，课中每 15-20 分钟穿插 3-5 题实时检测，课后复习灵活安排，避免占用过多教学或学习时间，保障策略长期落地。

（二）应用模式：构建覆盖教学全流程的评价体系

在课前预习阶段，教师通过科大讯飞智学网提前 2 天推送“科技主题基础测评”（10 题/5 分钟），AI 实时生成班级报告（如“65% 学生未掌握‘robot’动词用法”）与个人诊断（如“薄弱点：sustainable development 词义”）。教师依报告调整课堂重点，学生针对性预习薄弱词汇。

在课中学习阶段，教师通过智学网小程序分时段推送任务：新词汇讲解后推 3 道巩固题（5 分钟），AI 反馈正确率（<70% 重讲，≥90% 进下一环节）。课堂练习推应用型任务，AI 标记典型错误、生成答题热力图，教师集中纠错并对耗时超 2 分钟学生一对一指导。

在课后复习阶段, AI 生成个性化方案: 已掌握词汇 (如 computer) 按间隔推送对话/短文检测, 薄弱词汇 (如 sustainable development) 推目标匹配、造句等任务。每周生成学情报告 (如 7 天前词汇留存率 82%), 教师提醒薄弱词汇复习, 学生调整计划。

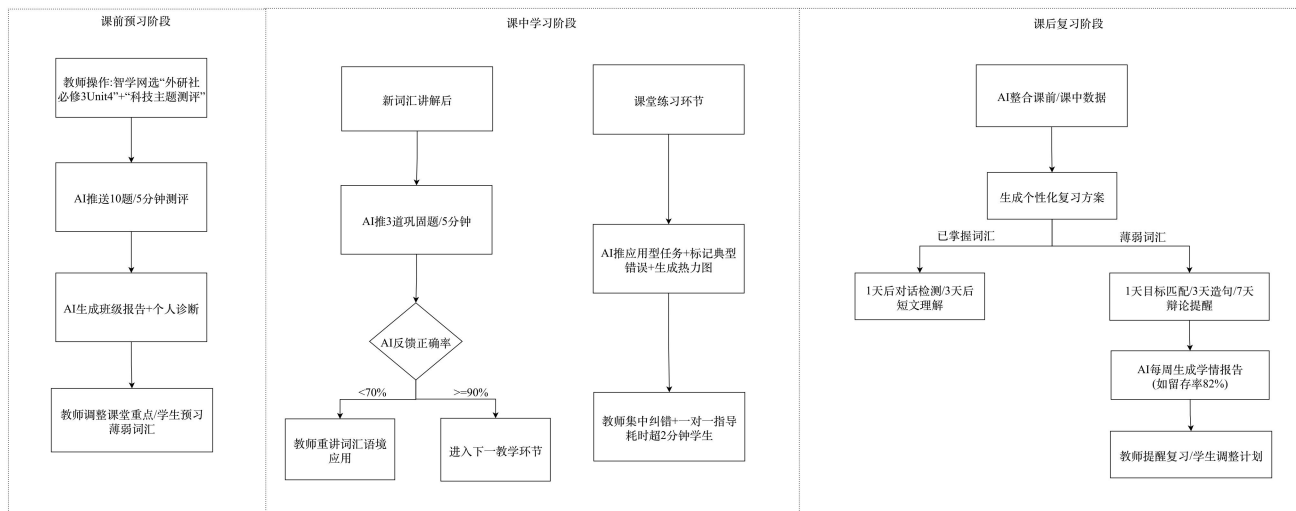


图 1 应用模式

(三) 实施步骤: 确保策略有序落地与优化

在前期准备阶段, 需要夯实基础, 为后续应用做好铺垫。AI 工具的选型应优先考虑与外研社教材同步的科大讯飞智学网。该平台内置“必修 3 Unit 主题词汇库”, 支持“自定义评价指标”和“生成主题词汇学情报告”, 且操作简洁高效, 教师端仅需三步即可完成任务布置, 学生端则可通过微信扫码参与。为确保顺利实施, 教师需接受两次培训: 第一次重点讲解工具的基础操作, 包括导入主题词汇库和查看学情报告, 并通过实操演示如何在 1 分钟内推送课前测评; 第二次则侧重数据解读与教学调整, 结合“班级主题词汇正确率 80% 但应用率仅 50%”的案例, 指导教师如何优化教学策略, 如增加应用类任务。同时, 还需对学生开展一次培训, 演示“扫码参与评价—查看错题解析—接收复习任务”的完整流程, 帮助其熟练掌握工具使用方法。

中期实施阶段采用“分阶段、循序渐进”的方式, 避免一次性全面推广带来的适应难题。第一阶段 (1-2 周) 主要在课后复习中应用 AI 评价, 让师生初步熟悉操作与模式; 第二阶段 (3-4 周) 逐步增加课堂中的实时检测应用, 重点实践“实时检测—动态调整”的教学配合; 第三阶段 (5 周及以后) 全面应用课前、课中、课后三阶段评价模式, 形成完整的教学闭环。在实施过程中, 教师需每周收集师生反馈, 例如“AI 任务难度是否合适”“操作是否便捷”, 并及时解决应用中的问题, 如调整任务难度、简化操作步骤, 从而保证实施的顺畅推进。

后期优化的核心是“以数据反馈为依据, 不断改进应用策略”。教师应分析 AI 生成的教学数据, 例如班级整体词汇学习效果 (学期初与学期末词汇测试平均分对比、各阶段任务正确率变化)、AI 评价对教学效率的提升情况 (如教师批改作业时间减少比例)。若数据显示某一阶段评价效果不佳 (如课后复习评价的词汇留存率未提升), 则需分析原因 (如任务难度过高或复习间隔不合理), 并相应调整策略 (如降低难度、优化复习间隔)。同时, 应结合师生的长期反馈, 全面了解 AI 辅助形成性评价的优势与不足。例如, 学生可能认为 AI 反馈详细但希望增加更多互动形式, 教师则可能认为学情报告有价值但期待更多个性化教学建议。基于这些反馈, 可以推动工具功能优化 (如增加词汇游戏化评价形式), 并在学情报告中附带教学建议, 以确保 AI 辅助形成性评价持续适配教学需求, 发挥最大效用。

五、应用挑战与展望

(一) 主要挑战: 多维度制约 AI 辅助评价落地

在技术层面, AI 对词汇应用能力评价易“机械判断”, 仅能检测拼写、语法错误, 难识别“词汇搭配生硬”“语境适配不当”等深层问题。多数 AI 工具词汇库更新滞后, 未纳入新课标新增主题词汇 (如“碳中和”“元宇宙”相关英文词汇), 与教学实际脱节。

在师资层面, 教师面临双重短板: 一是操作局限于“布置任务、看分数”, 无法熟练使用“自定义评价指标”等进阶功能; 二是难解读海量数据 (如词汇错误占比、复习间隔与留存率关联数据), 无法将“学生高频错动词短语”等数据转化为教学调整策略, AI 数据未发挥实效。

在社会与心理因素层面,教师支持、家长支持和学生的学习韧性在 AI 辅助词汇学习中具有重要作用^[10]。此外,游戏化形成性评价被证明能够显著提升 EFL 学习者的动机与参与度^[11]。

(二) 优化方向与未来展望

在技术升级上,需推动 AI 工具向“深度评价”与“精准适配”发展。一方面,通过自然语言处理技术迭代,提升 AI 对词汇应用能力的评价深度,例如引入“词汇搭配库”“语境适配模型”,让 AI 能识别词汇使用的灵活性与恰当性。另一方面,加强 AI 工具与教材、课标的协同开发,鼓励企业与教育部门合作,推出“教材同步版”AI 评价工具,确保评价内容与教学进度、课标要求高度契合。

在师资培训上,构建“技术+数据”双维度培训体系。学校可联合 AI 企业开展分层培训,基础层培训聚焦工具操作技巧,提升教师使用熟练度。进阶层培训侧重数据解读能力,通过案例教学(如“如何根据 AI 数据调整词汇复习计划”),引导教师将数据转化为教学策略,同时建立“教师技术互助小组”,让技术熟练的教师带动同伴共同提升。

在观念引导上,以“成果示范”打破认知误区。学校可选取部分班级开展 AI 辅助评价试点,定期展示试点班级的教学成果(如词汇测试平均分提升、学生复习效率提高),用实际效果消除教师的技术焦虑。同时,通过“AI 反馈与教师反馈对比”活动,让学生直观感受 AI 反馈的即时性、针对性优势,逐步建立对 AI 评价的信任,转变学习观念。

未来 AI 辅助高中英语词汇形成性评价可向“多技术融合”方向突破,其中 AI 与 VR(虚拟现实)技术的融合前景广阔。借助 VR 技术,可构建“沉浸式词汇学习场景”,例如在“购物场景”VR 环境中,学生通过与虚拟店员对话使用英语词汇,AI 实时评价学生词汇发音、搭配、语境应用的准确性,并即时推送纠错提示与拓展词汇。此外,AI 与大数据技术的深度融合,可实现“跨学科词汇关联评价”,例如在科学课英语词汇学习中,AI 通过分析学生数学、物理学科的英语词汇掌握情况,构建“学科词汇能力图谱”,为学生推荐跨学科词汇学习资源,推动词汇教学从“单一学科”向“综合素养”延伸。

六、结论与建议

(一) 研究结论

本研究以“外研社必修 3 Unit4 Science and technology”单元为载体,探索 AI 辅助形成性评价在高中英语主题词汇教学中的应用,核心价值有三:一是主题关联的个性化,AI 依单元主题词汇群定制评价方案(如“科技环保”子主题中,为基础层推认知任务、精通层推应用任务),破解传统评价“一刀切”问题,助力学生建立词汇与主题的深度关联;二是语境适配的及时性,AI 实时追踪学习数据,10 秒内生成反馈并推送语境资料,避免错误记忆强化,同时帮助教师即时调整教学节奏;三是素养导向的可持续性,AI 建立“主题词汇成长档案”(如学生“科技词汇”应用正确率从 50%升至 85%)实现长期监测,且评价融入“科技伦理”等内容,兼顾词汇能力提升与文化意识、社会责任感培养,推动学科育人。

(二) 实践建议

教育管理部门应加大“主题适配型 AI 词汇评价工具”投入,联合企业开发“教材同步版”工具(如外研社全册主题词汇库),编写《AI 辅助高中英语主题词汇评价指南》,明确多主题评价指标设计、AI 数据解读方法。高中学校应构建“AI+主题词汇教学”常态化机制,将 AI 评价应用频率、学生词汇掌握提升率纳入教师考核,定期开展 AI 主题教学案例评选,完善数据安全制度(仅采教学相关数据、加密存储、分级授权)。一线英语教师应提升“主题-技术”融合能力,从“被动用 AI”转向“主动设计评价方案”(如“文化交流”主题中设计跨文化词汇对比任务),每两周开展 AI 评价反馈座谈会,依学生意见调整策略,形成“教师主导、AI 辅助、学生主动”闭环。

(三) 研究不足与未来方向

本研究在实证范围窄,仅以“Science and technology”单元为案例,未验证多主题适用性。此外,技术应用浅,AI 仅能评价词汇拼写、基础应用,难识别搭配恰当性、情感色彩,未充分结合自然语言处理技术。未来将开展多主题实证研究,选 3-5 个典型主题单元,对比实验班与对照班差异,量化策略普适性。二是深化技术融合,探索 AI 与 VR 结合(如“科技博物馆”VR 场景,实时评价词汇应用)^[12]。三是聚焦教师能力提升,设计“AI+教师指导”协同培训方案,强化“主题-技术”融合应用能力。

参考文献:

- [1] Alazemi A F T. Formative assessment in artificial integrated instruction: delving into the effects on reading comprehension progress, online academic enjoyment, personal best goals, and academic mindfulness[J]. *Language Testing in Asia*, 2024, 14(1): 44.
- [2] Zhang H, Ge S, Saad M R B M. Formative assessment in K-12 English as a foreign language education: A systematic review[J]. *Heliyon*, 2024, 10(10).
- [3] 高凌飏. 过程性评价的理念和功能 [J]. *华南师范大学学报(社会科学版)*, 2004, (06): 102-106+113-160.
- [4] Hein G E. Constructivist learning theory[J]. *Institute for Inquiry*, 1991, 14.
- [5] Gao R. The vocabulary teaching mode based on the theory of constructivism[J]. *Theory and Practice in Language Studies*, 2021, 11(4): 442-446.
- [6] Rachmad Y E. Adaptive Learning Theory[J]. 2022.
- [7] Tran K Q. The effects of integrating artificial intelligence into learning academic vocabulary among masters in English studies students[J]. *International Journal of AI in Language Education*, 2024, 1(1): 1-19.
- [8] Hintzman D L. Repetition and memory[J]. *Psychology of learning and motivation*, 1976, 10: 47-91.
- [9] Jomaa N, Attamimi R, Al Mahri M. Utilising artificial intelligence (AI) in vocabulary learning by EFL Omani students: The effect of age, gender, and level of study[C]//*Forum for Linguistic Studies*. 2024, 6(5): 171-186.
- [10] Khan M A, Kurbonova O, Abdullaev D, et al. Is AI-assisted assessment liable to evaluate young learners? Parents support, teacher support, immunity, and resilience are in focus in testing vocabulary learning[J]. *Language Testing in Asia*, 2024, 14(1): 48.
- [11] Zhang Z, Crawford J. EFL learners' motivation in a gamified formative assessment: The case of Quizizz[J]. *Education and Information Technologies*, 2024, 29(5): 6217-6239.
- [12] Tang L, Shen K, Le H. Facilitating learners' self-assessment during formative writing tasks using writing analytics toolkit[J]. *Journal of Computer Assisted Learning*, 2024, 40(6): 2822-2839.

Application of artificial intelligence-assisted formative assessment in high school

English vocabulary teaching

-----A Case Analysis of FLTRP Book 3 Unit 4“Science and technology”

Lan Jin¹, Junli Liu²

¹ *Tai Yuan Normal University, Jin zhong, China*

Abstract: With the development of AI technology, its application in education is becoming increasingly widespread, offering potential solutions to problems such as poor retention, limited evaluation, and delayed feedback in high school English vocabulary instruction. This article, drawing on literature analysis and theoretical discussion, focuses on the application of AI-assisted formative assessment in this context. It analyzes existing teaching issues and the theoretical foundations of formative assessment, elaborating on the intersection of AI and formative assessment, and proposes feasible strategies based on design principles, application models, and implementation steps, aiming to provide new ideas and actionable solutions for frontline teachers. Research has shown that this model can enhance the personalization, timeliness, and sustainability of vocabulary learning. However, its widespread adoption is hampered by challenges such as technology, teacher resources, and conceptualization, requiring continuous optimization in practice.

Keywords: artificial intelligence; formative assessment; high school English; vocabulary teaching; teaching innovation