

DOI: 10.3969/j.1001-7232.2024253

高校英语写作应用 AIGC 自动评价实效性的对比研究

韩芳婷¹ 毛景泽² 凌淑珍¹

(1. 西北农林科技大学语言文化学院, 陕西杨凌 712000; 2. 西北农林科技大学人文社会发展学院, 陕西杨凌 712000)

摘要:在教育数字化转型及人工智能生成内容兴起的大背景下,本研究立足“基础英语写作”课程开展数字化教学实践与改革,以课程所发布的描述文写作、说明文写作和议论文写作任务评分数据为依据,运用SPSS 29.0版本对批改网、ChatGPT与教师3种不同方法的评分进行统计分析,采用描述统计、正态性检验、相关性分析和回归性分析等方法。研究表明,ChatGPT评分与其他两者显著相关,其均值介于教师评分与批改网评分之间,且在评价维度和反馈形式上独具特色。通过对师生的访谈和对学生的调查问卷分析发现,教师对AI技术融入写作课堂基本持积极态度,并认为ChatGPT是对传统的“教师+批改网”模式的完善与提升;学生对英语写作评价认可程度为教师评价>ChatGPT评价>批改网评价。此外,还需要加强对学生如何使用AI工具的有效指导。建议应提升教师数字素养、推动数字课堂变革、实施多元教学,以推动AIGC在英语作文自动评价中的有效应用。

关键词:教育数字化;人工智能生成内容;英语写作自动评价;批改网;ChatGPT

进入21世纪,随着互联网、大数据、人工智能等一系列新型科技的突破与网络用户的持续增长,教育数字化成为顺应时代潮流、国家战略的重要举措。2019年10月《中共中央关于坚持和完善中国特色社会主义制度 推进国家治理体系和治理能力现代化若干重大问题的决定》中提出:“发挥网络教育和人工智能优势,创新教育和学习方式,加快发展面向每个人、适合每个人、更加开放灵活的教育体系,建设学习型社会。”^[1]党的二十大报告进一步提出“推进教育数字化,建设全民终身学习的学习型社会、学习型大国”的要求^[2]。2023年中共中央政治局第五次集体学习强调“教育数字化是我国开辟教育发展新赛道和塑造教育发展新优势的重要突破口”,要进一步“推进数字教育,为个性化学习、终身学习、扩大优质教育资源覆盖面和教育现代化提供有效支撑”^[3]。2023年以ChatGPT为代表的AIGC(artificial intelligence generated content,人工智能生成内容)的涌现,为包括教育在内的社会各个领域的发展带来新的动能。

为积极贯彻落实党的“十四五”规划与党的二十大报告对数字化转型的指示和要求,高校积极开展人工智能融入课程教学的应用探索。本研究在“基础英语写作”课程教学实践中融入人工智能(artificial intelligence, AI),对以ChatGPT为代表的AIGC自动评价、以批改网为代表的传统自动评价和课堂授课教师的评价模式进行对比分析,探究AIGC在农林高校英语写作自动评价中的作用,为英语写作类课程教学数字化改革提供参考。

一、批改网英语写作自动评价的应用现状

在AIGC大规模投入使用之前,国内高校英语写作自动评价基本是以批改网作文自动评改系统为基础。批改网自2011年投入应用以来得到了广泛的关注。批改网是以自然语言处理技术和语料库技术为基础,通过分析学生作文和标准语料库之间的距离,对学生的英语书面作文进行即时评分,并提供改善建议和-content分析结果的本土化的在线自动作文批改系统。

批改网在写作自动评价方面具有简单易操作、反馈及时、逐句点评、可反复修改提交等优势,因此在英语写作教学中得以广泛应用。许多高校教师会在批改网上注册班级并发布写作任务,让学生在平台上完成作文。学生提交作品后,批改网即时自动生成打分和逐句的评价,直接反馈给学生。学生可基于批改网的反馈进行多次编辑修改,直至满意再最终提交。同时在批改网平台,当学生提交的作文达到一定数量,即可形成专属英语学习者的语料库,这为很多学者开展教学研究提供了便利条件。

然而,随着教学任务量增大、教学理念更新、新时代的数字化需求,批改网也逐渐显露出其不足之处。批改网自动评价体系以评分为导向,存在机械化弊端。①过分强调对作文基础项目的评价,即词汇的高阶性、语法的正确性、句法的复杂多变性等,因此,会出现学生为了不断地刷高分而刻意堆砌生僻词、长难句,从而忽略实际的语用效果。②缺少对作文核心内容与思想的评价,缺少对情感态度、价

价值观为代表的抽象主题的评判。③过于侧重通过衔接词的使用来判断文章的逻辑性,对文章的立意是否明确、逻辑是否合理、观点是否具有新意难以进行合理判断。④评价存在标准化趋势,缺乏个性化评价,无法对个人独立的写作水平给出个性化建议。因此,不少学者在研究中虽然都肯定了批改网作为英语写作自动评价系统的优势,但同时也指出,目前批改网仍无法替代教师在学习者写作过程中扮演的重要角色,不能忽视教师的作用^[4-6]。数字化时代背景下,英语写作自动评价研究只聚焦于批改网显然是落后且片面的,需要融入 AIGC 和大语言模型进行对比研究以深入了解各自的作用与成效。

二、AIGC 背景下英语写作自动评价的实效性分析

为探究 AIGC 背景下的英语写作自动评价实效性,本研究依托农林高校英语写作课堂教学,分析人工智能在当前英语写作评分、评价方面的实际应用情况。本研究使用的评分数据来源于西北农林科技大学英语专业“基础英语写作”课程作文评分,其中包括描述文写作(descriptive writing)、说明文

写作(expository writing)和议论文写作(argumentative writing)3次不同体裁写作任务的评分。笔者于2023年春季学期教学课堂上共搜集225份有效个案。本研究使用IBM公司开发的SPSS 29.0版本对批改网、教师、ChatGPT共3种方法的评分进行分析,采用描述统计、正态性检验、相关性分析与回归性分析方法。

(一) 研究结果统计分析

首先进行描述统计,对3次作文评改分数进行简单的数理统计(见表1)。笔者观察到批改网的平均评分始终高于另外两项,方差较小,数据波动较小;教师评分存在的极端值较为突出,方差较大,数据波动较大;ChatGPT评分在平均值、方差方面处于中间位置,其极端值出现次数少,方差适中,数据波动幅度适中。从偏度和峰度的数据方面,计算Z值(偏度或峰度的统计值除以标准误),并以此判断偏离程度。 $|Z| < 1.96$ 为轻微偏离, $1.96 \leq |Z| < 3$ 为中等偏差, $|Z| \geq 3$ 为显著偏差。从表1的Z值可知,在全部分数统计中,教师评分相对趋向于正态分布,更符合英语作文最终分数分布情况,而批改网评分和ChatGPT评分则与正态分布的最优情况有明显偏差。

表1 3种评价方法评分的描述统计

	CN	Min.	Max.	AV	SD	Var.	SKEW(偏度)			KURT(峰度)		
							ST	SE	Z 值	ST	SE	Z 值
全部批改网评分	225	60.0	94.0	85.073	5.537 8	30.668	-1.357	0.162	-8.38	2.112	0.323	6.54
全部教师评分	225	53.0	94.0	76.947	9.798 7	96.015	-0.524	0.162	-3.23	-0.353	0.323	-1.09
全部 ChatGPT 评分	225	62.0	93.0	82.071	5.875 5	34.522	-1.290	0.162	-7.96	1.720	0.323	5.33
描述文批改网评分	75	60.0	92.5	81.327	6.671 6	44.510	-0.630	0.277	-2.27	0.186	0.548	0.34
描述文教师评分	75	53.0	93.0	73.213	11.081 1	122.792	-0.051	0.277	-0.18	-0.864	0.548	-1.58
描述文 ChatGPT 评分	75	62.0	89.0	79.720	7.155 7	51.204	-0.942	0.277	-3.40	0.005	0.548	0.01
说明文批改网评分	75	74.0	93.0	86.840	4.273 7	18.265	-1.125	0.277	-4.06	1.284	0.548	2.34
说明文教师评分	75	53.0	90.0	75.600	8.818 0	77.757	-0.426	0.277	-1.54	-0.131	0.548	-0.24
说明文 ChatGPT 评分	75	62.0	93.0	82.613	5.321 8	28.321	-0.956	0.277	-3.45	1.790	0.548	3.27
议论文批改网评分	75	74.5	94.0	87.053	2.948 4	8.693	-1.278	0.277	-4.61	3.773	0.548	6.89
议论文教师评分	75	60.0	94.0	82.027	6.926 2	47.972	-0.764	0.277	-2.76	0.618	0.548	1.13
议论文 ChatGPT 评分	75	71.0	90.0	83.880	3.969 3	15.756	-1.182	0.277	-4.27	1.584	0.548	2.89

注: CN(case number), 个案量; Min.(minimum value), 最小值; Max.(maximum value), 最大值; AV(average value), 平均值; SD(standard deviation), 标准差; Var(variance), 方差; SKEW(skewness), 偏度; KURT(kurtosis), 峰度; ST(statistics), 统计值; SE(standard error), 标准误。

接下来运用非参数检验法,使用柯尔莫戈洛夫-斯米洛夫检验(Kolmogorov-Smirnov)和夏皮洛-威尔克检验(Shapiro-Wilk)对数据进行进一步的正态分布检验,观察到显著性都小于0.001,三者评分并不符合正态分布的理想结果(见表2)。

表2 3种评价方法评分的正态性检验

	(Kolmogorov-Smirnov)			(Shapiro-Wilk)		
	ST	df	Sig.	ST	df	Sig.
全部批改网评分	0.135	225	<0.001	0.893	225	<0.001
全部教师评分	0.098	225	<0.001	0.959	225	<0.001
全部 ChatGPT 评分	0.157	225	<0.001	0.897	225	<0.001

注: df为自由度,一般情况下Kolmogorov-Smirnov检验和Shapiro-Wilk检验的样本量直接决定自由度。

采用皮尔逊相关性分析(见表3),观察到ChatGPT评分与批改网评分、教师评分的相关系数分别为0.611、0.569,说明ChatGPT评分与另外两

表3 3种评价方法评分的相关性分析

		批改网评分	教师评分	ChatGPT评分
全部批改网评分	PC(皮尔逊相关性)	1	0.484*	0.611*
	Sig.(2-tailed)		<0.001	<0.001
全部教师评分	PC(皮尔逊相关性)	0.484*	1	0.569*
	Sig.(2-tailed)	<0.001		<0.001
全部 ChatGPT 评分	PC(皮尔逊相关性)	0.611*	0.569*	1
	Sig.(2-tailed)	<0.001	<0.001	

注: *表示在0.01级别(双尾),相关性显著。

项均呈强相关关系，关联性强，线性关系非常明显，在一定情况下可以使用 ChatGPT 评分用来拟合和预测批改网评分和教师评分。

相关分析是回归分析的基础和前提，回归分析是相关分析的深入和继续，在得到 ChatGPT 评分与批改网评分、教师评分都具有强相关关系的基础上，通过回归分析进一步探究并确定 ChatGPT 评分与另外两项在显著依存关系之上的数量关系，能够为之后开展英语写作融合评价、形成更为完善的英语作文评价体系提供参考。通过回归分析(见表 4)，可以得到 ChatGPT 评分与批改网评分的数量关系大致为批改网评分=0.576×ChatGPT 评分+37.841，教师评分=0.949×ChatGPT 评分-0.953。根据上述函数关系，对于学生来说，可以通过 ChatGPT 评分，简单预测和拟合之后的批改网评分和教师评分。对于教师教学来说，可以在批改作文时一定程度上参考 ChatGPT 评分，提前对作文分数进行大致预判。同时，教师在实践中，可以对比分析 AI 评分与教师自身评分的数值，进一步完善相关的函数关系，提高之后对作文分数评价的合理性。

表 4 以 ChatGPT 为常量的 3 种评价方法回归分析

		USC		SC	t	Sig.
		B	SE	Beta		
因变量 1 (批改网评分)	Con.(常量)	37.841	4.113		9.201	<0.001
	ChatGPT 评分	0.576	0.050	0.611	11.514	<0.001
因变量 2 (教师评分)	Con.(常量)	-0.953	7.555		-0.126	0.900
	ChatGPT 评分	0.949	0.092	0.569	10.337	<0.001

注：USC(standardized coefficients)，非标准化系数；SC(standardized coefficients)，标准化系数。

(二) 讨论

1. 以 ChatGPT 为代表的 AI 评价与教师评价的对比

从评价分数来看，ChatGPT 评分的平均值无论是在整体还是单次作文任务中都高于教师评分的平均值，在最值方面并未出现过多如教师评分中较为极端的分数。出现这一现象的原因一方面是 ChatGPT 评分相较于教师评分能够更充分地减少个人评价偏好给最终作文分数带来的极端影响，以智能式的更为客观的评判方式，极少受到外部因素干扰的优势得出作文分数；另一方面与教师相比，ChatGPT 拥有更为丰富的互联网评价数据与评价案例支撑下的评价体系，同时具有强大的分析已有案例以达到自适应的学习完善能力。

从评价维度来看，教师评价围绕“基础英语写

作”课程规划，结合热点或高频话题布置作文任务，作文评价主要围绕与相应类型题材作文契合度与语言文化适配度进行打分。ChatGPT 从 5 个维度对作文进行评价，即内容连贯(content and coherence)、语言词汇(language and vocabulary)、组织结构(organization and structure)、创意与想象(creativity and imagination)、投入与影响(engagement and impact)。每个维度给出详细的文字评价与分数，最终汇总成总分。总体而言，ChatGPT 的评价更为多元全面，更聚焦对作文整体性的把握。然而，在与课堂结合度方面，教师评价的维度与课堂教学重难点的结合更加密切。

从评价反馈来看，学生面对教师评价更能表现出信服与认同，这与传统师生关系与互信模式有较大关联，学生期待由教师所代表的“教学的官方形式”给出的评价。同时，教师富有个人特色和针对性的评价言辞让学生感到更加亲切，进一步促进学生对教师评价的正反馈。而面对 ChatGPT 评价，学生首先表现的是惊讶和不信任，这源于学生对当前 AI 评价功能的不了解，与随之产生的对 AI 评价能力的怀疑。当然，这也与学生适应作文分数评价舒适区有关，对新的评价模式和评价体系的接受速度较为缓慢。

2. 以 ChatGPT 为代表的 AI 评价与批改网评价的对比

从评价分数来看，批改网评分的平均值无论是在整体还是单次作文任务中都高于 ChatGPT 评分平均值，且批改网评分的方差始终小于 ChatGPT 评分，高低极端值相差较小。这说明批改网在高分段的分布频次较大且集中。批改网与 ChatGPT 在英语写作自动评价方面是相似的，且批改网在一定程度上已经成为大部分作文教学使用的软件选择。十余年的数据积累和模型完善，促使批改网更能够适应中国高校英语作文评价的模式，当然也不可避免地存在评价分数虚高、作文刷分等情况。ChatGPT 在分数评价方面相较于批改网并没有长时间参与高校英语写作评价之中，数据相对有限，评分则倾向于以外国或互联网标准进行评价，因此作文分数相对较低。

从评价维度来看，批改网的评价维度是由单词、句子、篇章结构与内容相关 4 个模块构成，覆盖英语基础知识模块的评价。这对于从英语作文中拓展高级词汇量、提升语法与句法应用能力、增强围绕主题写作能力有着较大的帮助。当然这对

于更为纯粹的写作教学在情感态度价值观的评价方面仍有些许不足。首先, ChatGPT 的评价维度包括了批改网的评价维度; 其次, ChatGPT 的评价维度中虽然并未直接表明对情感态度价值观的评价, 但其创意与想象、投入与影响两个维度其实有关注到学生作文的情感态度部分, 且它对提升学生作文的创造力和创新度有一定的促进作用。

从评价反馈来看, 学生在使用批改网评价作文后, 可以即刻根据批改网所给出的包括作文格式、单词拼写与替换、语法使用等方面进行完善, 并以此来提高作文分数。这是一个良性的且始终带有正反馈的提高作文基础水平的方式。学生纠正错误, 得到的正反馈是提升作文分数, 极大地提高了学生修改完善作文的积极性, 对作文质量提升有较为明显的成果。但不少同学也发现批改网的反馈存在机械化的弊端, 过度强调高阶词、短语、长难句等的运用而忽略写作的实际效果。ChatGPT 的模式是问答互动式, 在整体评价之后, 需要学生与 AI 进行提问互动。ChatGPT 通过人机交互、人机对话克服了批改网在评价上存在的机械性问题, 可以呈现出更为严谨、准确、科学的反馈与评价, 指导学生有针对性地进行修改。在与 AI 的交互中, 学生相当于接受了一次沉浸式、个性化的智能教学, 通过 AI 给出的反馈对作文进行修改完善, 得到作文水平真正的提升。

3. 师生对以 ChatGPT 为代表的 AI 评价模型的评价

为进一步了解师生对于 AI 融入英语写作评价的看法, 促进数字化赋能高校英语写作“智能+”教学模式的构建, 本研究对参与该课程改革的师生进行了调研。

1) 教师访谈。选取 3 名不同国籍、不同教龄的英语写作教师作为访谈对象, 分别于学期前、学期中、学期末开展了 3 次访谈, 以深入探究教师在教学过程中融入 ChatGPT 所持态度、真实体验、面临的困境及期望与建议等。访谈中, 3 名教师均表示 AI 赋能教学改革是当前教改的重要方面, 他们对以 ChatGPT 为代表的 AI 技术融入写作课堂基本持积极态度, 并认为 ChatGPT 自动评价的融入是对传统“教师+批改网”模式的完善与提升。在访谈过程中, 在实际应用环节, 英语写作课教师们肯定了 ChatGPT 在实际应用中的即时反馈、省时高效等方面表现优越, 认为极大地提高了写作评价效果。然而, 面对描述文写作(descriptive writing)、说明文

写作(expository writing)和议论文写作(argumentative writing)3 种不同体裁的写作任务, ChatGPT 并未能将不同体裁写作的要点进行提炼和突出, 一以概之地从内容连贯、语言词汇、组织结构、创意与想象、投入与影响 5 个维度进行反馈, 不利于学生对于课程重难点的掌握。同时, 关于学生英语写作中 AI 的参与度问题, 有教师表达了担忧。在教学过程中, 发现有学生直接使用 AI 生成作文, 将生成好的作文参与评价。这属于学术伦理问题, 需要教师在英语写作教学中积极开展课程思政教学, 对学生加强学术诚信教育, 引导学生正确使用 AI 工具, 避免学术不端行为。但是, 面对学生提交的作业, 如何能够甄别其真伪, 辨别究竟是出自学生之笔还是 AI 之手, 这对于教师而言是一个不小的挑战。

2) 学生问题调查与访谈。本研究小组于 2024 年 4 月对学习“基础英语写作”课程的 75 名大一英语专业的学生发放了调查问卷, 回收有效问卷 70 份, 有效回收率为 93.33%。调查问卷内容聚焦 ChatGPT 的使用, 涵盖所持态度、使用感受、问题及建议等。对于 AI 这一新事物, 大一新生表现出较快的适应能力。在经历了最初短暂的质疑后, 很快转而愿意在英语作文评改方面与新的 AI 技术进行交互。2024 年 6 月在访谈过程中, 90% 以上的学生都认为利用 AI 可以有效地修改自己的作文, 通过与 AI 的反复互动、基于 AI 的即时反馈, 他们可以多次修改直至最终形成自己满意的写作。相较于批改网, 100% 的学生认可 ChatGPT 的评价范围, 认为其更为具体、专业、智能化。但是, 超半数的同学对 ChatGPT 评价的认可程度低于教师评价。一方面, 当前 AI 无可避免的“套话”评价也让学生对于 AI 评价的可行性和有效性持一定质疑态度; 另一方面, 在学生眼中, 教师评价具备更高的权威性、更强的针对性和更为独特的人文关怀。当然, 在问卷中, 有 10% 的学生坦言自己曾直接利用 AI 生成作文, 以快捷地完成写作任务。总体而言, 英语写作评价方面, 学生的认可程度排序为: 教师评价 > ChatGPT 评价 > 批改网评价。ChatGPT 有潜力、有实力、有动力成为英语写作课堂的智能助手。当然, 在此之前还是需要加强对学生如何使用 AI 工具的有效指导, 引导学生合理用好、用对、用深这把“双刃剑”。

三、AIGC 应用于英语写作自动评价的建议

数字化时代, 教育的数字化转型是必然趋势。

引入 AI 自动评价模型是智能时代的趋势, 是国家教育政策的导向, 也是教育深化改革的重要角色。AIGC 引入英语写作自动评价能够有效弥补作文评分受到个人单方面评价偏好的影响, 其评价与最终分数更为可观, 同时也可以减少作文分数虚高的情况。在保证对单词、语法、句法等的评价标准的基础上, AIGC 能够强化对情感、态度、价值观等相对较为抽象的标准的评改, 辅助学生进一步提升作文水平。鉴于以 ChatGPT 为代表的 AI 大语言模型比批改网在英语写作自动评价上具有更为显著的优越性, 接受并应用 AI 大语言模型进入课堂, 用 ChatGPT 取代批改网, 采用“教师+批改网”的写作评价模式可能成为更多教师与学生的选择。然而, AIGC 在英语写作自动评价领域的进一步推广应用仍然需要时间。在这个过程中, 课堂上不仅要融入 AI, 更要着力于培养师生的数字素养, 落实走深数字课堂, 并推动多元教学。虽然是基于“基础英语写作”这门课程的教学实践, 但是鉴于高校英语写作教学的高度相似性和相关性, 本研究所提出的以下 3 条建议也同样适用于其他英语写作课程。

1) “教育数字化离不开人的作用, 尤其是教师的作用”^[7], 需要“具备数字素养, 具有人机互动、人机协同能力的教师”^[7]。这对教师的数字素养提出了新要求, 使用什么样的人工智能, 怎么更好地使用人工智能, 如何将课上教学、课下练习进一步与 AI 融合, 这都是对高校一线教师提出的问题与挑战。教师唯有主动了解当前趋势, 敢于跳出当前教学舒适圈, 积极尝试与 AIGC 交互, 大胆在教学改革中使用新的数字教学方式, 才能真正推进数字化赋能高等教育的进程。

2) 随着 AI 在应用层面的深度学习, 数字课堂的改革迎来新的机遇。①利用 AIGC、大数据分析赋能, 对全网英语写作课程资源进行整合, 精选优秀教学课程资源, 形成更为优质的精华教学资源集合。②数字多媒体与 AI 的交互, 能丰富教学内容, 提升教学课件的质量, 推动高水平课堂建设。③AIGC 投入到英语写作自动评价中, 可以提升英语作文评改质量, 促进英语写作自动评价向个性化发展。

3) AI 工具的一大核心就是交互, 体现在英语写作课堂中即为多元教学。课堂中教师与学生的交互作为传统教学模式一直持续, 但课下教师与学生的交流频次明显下降, 学生课下学习遇到难题无

法得到及时的反馈和帮助。以 ChatGPT 为代表的人工智能工具的强交互功能为该问题的解决提供思路。在当前 AI 发展的今天, 多元教学将不再仅局限于课上教师讲课、学生听课, 课下教师布置练习、学生完成练习。教师可以通过与 AIGC 互动, 提高备课效率, 丰富教学环节, 总结教学过程; 学生也可以通过与 AIGC 互动, 提前预习与复习教学重点, 解决遗留问题, 在交互中提升对学习的认知。

面对当前 AI 与高校教学深度融合的趋势, 最终是否使用才是关键。从学校到教师、学生层面, 必须更新原有的认知观念, 精确把握国家战略导向, 充分认识到 AI 在高校教育教学应用中的必然性与必要性, 主动适应在教学过程中 AI 的应用, 积极将学校办学、教师教学、学生自学与 AI 的发展与应用充分结合, 最大程度发挥数字素养的积极作用。

资助项目: 西北农林科技大学大学生创新训练国家级项目——数字化赋能高校英语写作“智能+”教学模式构建研究, 项目编号 202410712193; 教育部产学研合作协同育人项目“AI 赋能医学英语词汇教学实践基地”, 项目编号 240905468041047。

参考文献

- [1] 新华社. 中共中央关于坚持和完善中国特色社会主义制度 推进国家治理体系和治理能力现代化若干问题的决定 [EB/OL]. (2019-11-05) [2024-11-30]. https://www.gov.cn/zhengce/2019-11/05/content_5449023.htm.
- [2] 习近平. 高举中国特色社会主义伟大旗帜为全面建设社会主义现代化国家而团结奋斗: 在中国共产党第二十次全国代表大会上的报告 [EB/OL]. (2022-10-17) [202-11-30]. http://www.xinhuanet.com/mrdx/2022-10/17/c_1310670149.htm.
- [3] 新华社. 习近平主持中央政治局第五次集体学习强调 加快建设教育强国 为中华民族伟大复兴提供有力支撑 [EB/OL]. (2023-05-29) [2024-11-30]. http://www.xinhuanet.com/politics/leaders/2023-05/29/c_1129654921.htm.
- [4] 李奕华. 基于动态评估理论的英语写作反馈方式比较研究 [J]. 外语界, 2015(3): 59-67.
- [5] 陈慰. 基于自动评分系统的大学英语作文反馈研究: 以“句酷批改网”为例 [D]. 重庆: 重庆大学, 2013: 17-45.
- [6] 鲍昕慧. 作文自动评价系统应用于高中英语写作教学的实证研究 [D]. 南京: 南京师范大学, 2018: 22-40.
- [7] 袁振国. 教育数字化转型: 转什么, 怎么转 [J]. 华东师范大学学报 (教育科学版), 2023, 41(3): 1-11.

(责任编辑 孔 艳)