

生成式人工智能反馈在国际中文教学设计中的应用研究 ——基于多反馈模式的实证比较

孙冕

西京学院, 中国

李悦菡 *

西安培华学院, 中国

摘要

在数智化背景下, 生成式人工智能为国际中文教师的教学设计能力发展带来新机遇。现有研究多关注 AI 辅助二语习得, 而关于教师如何利用 AI 反馈促进专业发展的实证探讨仍显不足。本研究以国际中文职前教师为对象, 比较了控制组、同伴反馈组、AI 反馈组与 AI+ 同伴反馈组在教案设计任务中的表现。研究发现, 各组修改稿质量均显著提升, 其中 AI 反馈组效果最优, AI+ 同伴反馈组次之。AI 反馈通过步骤化的引导优化了教案的逻辑与结构; 同伴反馈在情境适配与情感支持方面更具优势; 人机协同在反思与创新上具有互补效应。研究表明, 未来国际中文教育与人工智能的深度融合应以人机协同而非技术替代为发展方向。本研究为教师专业发展与教学设计优化提供了实证依据与思路。

关键词

国际中文教育, 生成式人工智能, 反馈, 教学设计

1 引言

在全球化与数字化背景下, 国际中文教育正面临数智化转型, 人工智能为教学设计、课堂互动与资源生成带来全新机遇 (郑艳群, 2023)。人工智能在“教”与“学”两个方面均能提供有效支持 (胡壮麟, 2023; 蔡薇, 2023), 成为继教师与学生之外的“第三支点”, 促使教学关系和学习方式出现重构趋势 (史中琦、吕伯宁, 2025)。在汉语教学与学习中, 生成式人工智能 (Generative Artificial Intelligence, 在本文中简称 AI) 既能够在编写方式、教案内容、学习资源与课堂活动等方面为国际中文教学设计提供助力 (韩开、徐娟, 2023), 又能够通过实时交互为教师与学习者提供反馈与指导机会。在教师端可辅助教案编写、练习设计等; 在学生端可提供口语对话、语音纠错等功能 (储诚志, 2025)。就教学中的反馈而言, 由于人数、语言背景和熟练程度等客观条件的限制, 提供有效和个性化的反馈是一项费时费力的任务 (Wiboolyasarin et al., 2024)。与之相比, AI 提供即时、个性化反馈, 已成为重要的反馈手段。然而, 当前国际中文教师群体对 AI 反馈的使用尚处于探索阶段, 相关研究多聚焦于学生端的使用效果。关于国际中文教师使用 AI 反馈辅助教学设计的实证研究仍相对有限, 亟需探究其应用效果。

* 通讯作者。联系电邮: yuehan0024@163.com

基于以上问题,本文聚焦于 AI 反馈辅助国际中文教学设计,探讨其在教师教学赋能中的应用路径与实效,回应国际中文教育数智化发展的需求,以期为 AI 赋能国际中文教师的专业发展提供理论依据与实践参考。

2 文献综述

2.1 生成式人工智能在国际中文教育中的应用研究

以 ChatGPT 和 DeepSeek 为代表的生成式人工智能快速发展,不仅能自动处理语言教师的常规任务,还能实时生成个性化反馈,提供多样化建议(储诚志, 2025)。从理论层面看,研究者普遍认为 AI 不仅能深度参与教学活动、支持各环节实施,还能帮助国际中文教师更高效地讲解汉语规则与文化知识(郑艳群、许丽婧, 2023; 崔希亮, 2025)。从实证层面看,已在语法、词汇等方面的教学中尝试引入 AI 并评估其应用效果。语法教学上,大语言模型能够有效赋能中文语法教学资源的生成(宾帅等, 2025),并通过即时反馈与个性化互动促进学习者对语言规则的长期保持与内化(余晓菲等, 2025)。词汇教学上, AI 工具在快速搜索、汇总信息及宏观逻辑列举上具有一定的优势,能够为新手教师提供基础性的“脚手架”支撑(孟凯, 2024)。总的来说,在教案编写、活动设计、教学资源生成等方面, AI 能从教学目标、语言表达和文化适配等维度提供建议,减轻教师负担并提升教学效率。然而,当前关于 AI 赋能国际中文教育的研究仍以理论探讨为主,基于真实教学场景的实证研究尚不充分,本文拟从教师视角出发探讨 AI 在教学中的应用与实效。

2.2 AI 反馈在二语教学中的应用与研究

自动生成反馈与个性化辅导是人工智能在教育领域的重要应用方向之一(Zawacki-Richter et al., 2019)。在国际中文教育中, AI 反馈已应用于语言技能的纠偏(王治敏等, 2025),具有多元教学价值。现有研究表明, AI 反馈在写作等领域能提升学习者动机、水平与参与感(Meyer et al., 2024; Wang et al., 2024)。对比研究显示,大语言模型在解释、提问与针对性上优于新手反馈,部分内容甚至超过专家(Jacobsen & Weber, 2025);人机协作模式在主题思想、写作框架、文本呈现等方面的反馈质量优于单一来源反馈(Yang et al., 2025)。质性研究进一步揭示了师生态度:师生普遍认为 AI 反馈因其无偏见而更可信(Ruwe & Mayweg-Paus, 2023),但也对部分建议持批判态度(Tay, 2024)。学习者往往在情感上批评,认知上通过互动设定目标并调整策略,行为上倾向采纳局部反馈(Ma et al., 2025)。

尽管 AI 反馈在教学中具有潜力,但其局限性不容忽视:教师在使用 AI 时普遍面临采纳反馈策略不清与技术依赖问题(储诚志, 2025),原因在于 AI 对上下文理解不足以及产生幻觉的风险(Xu et al., 2024),整体实用性低于教师反馈。对此, UNESCO (2024) 提出《教师人工智能能力框架》,强调教师应掌握“人工智能教学法”与“人工智能赋能专业发展”等能力。只有实现教师认知转型与角色重塑, AI 才能真正赋能国际中文教育(赵杨, 2025)。综合来看,关于 AI 反馈的研究尚处于起步阶段,鲜有研究从教师视角探讨其在教学中的价值。现有成果多集中于英语二语教学, AI 反馈在国际中文教育中的应用效果尚待验证。因此,有必要深入探讨 AI 如何在教学各环节中支持国际中文教师专业发展,提升其教学设计与反思能力。

2.3 生成式人工智能反馈赋能教学设计

教学是教师在特定目标引导下,为促进学习而组织的系统规划 (Gagné et al., 2007; 丁安琪, 2014),是融合教学、学科知识与教学逻辑的高级认知活动。其核心包括教学需求与目标分析、教学内容与方法设计、教学活动与评估安排等环节 (崔永华, 2008)。在结果导向型教学体系中,教师逐渐由知识传授者转变为学习设计者 (Biggs et al., 2022);不仅需要构思教学活动,更需确保教学设计能有效促进二语习得 (廖建玲, 2013)。《国际中文教师专业能力标准》(T/ISCLT 001-2022)将教学设计界定为核心教学技能,强调目标导向性、结构性与适应性。然而,职前教师常出现目标模糊、环节不完整、跨文化意识不足等问题 (闻亭、刘晓海, 2023),并因缺乏资源与指导而难以深入理解设计内涵,甚至降低兴趣与信心 (吴澜等, 2024)。而反馈是提升教学设计的重要途径。研究表明,同伴反馈能增强教师参与度和技能 (Ma et al., 2018),自我反馈亦能促进专业成长 (Ellis, 2009)。然而,传统反馈存在负担重、延迟、质量不一以及依赖元认知等局限。生成式人工智能的普及为教学反馈带来新的契机,推动语言教育迈入新阶段 (Chapelle, 2025)。现有研究指出, AI 反馈提高了教学效率与精度 (Novita, 2023),能明确教学目标与评价标准 (Pishtari et al., 2023),还能提升教学设计的结构化与系统性,推动教师由“以教为中心”向“以学为中心”转变 (Zhang & Ilisko, 2025)。

综上, AI 在提升教学设计质量与效率方面潜力显著,但相关研究多停留于理论与观察,系统的实证探讨仍有限。尤其在国际中文教学设计中, AI 反馈与传统反馈的差异及人机协同效果尚缺乏验证。因此,本文围绕教学设计任务,比较控制组(自我修改)、同伴反馈组(同伴反馈,后文简称同伴组)、AI 反馈组(生成式 AI 反馈,后文简称 AI 组)以及 AI+ 同伴反馈组(人机协同反馈,后文简称 AI+ 同伴组)在辅助教学设计与促进专业发展上的差异。

2.4 理论框架

本研究借鉴 Ellis (2010) 的纠正性反馈框架,从情感、认知和行为等层面探讨职前教师如何看待、接受和应用不同来源的反馈。该框架从学习者视角出发考察反馈效应,涵盖个体因素、情境变量、参与度与学习成效四个维度。其中参与度为核心维度,用于分析学习者对反馈的反应方式,具体包括情感、认知和行为。该框架已广泛应用于二语习得研究,例如比较二语写作者在教师与同伴反馈下的参与差异 (Cheng et al., 2023)。随着人工智能的发展,有研究将该框架用于分析研究生在修改学术研究计划书时与 ChatGPT 的互动方式 (Koltovskaia et al., 2024)。该理论为本研究提供了分析路径,帮助解释学习者如何接收与利用反馈,并将其转化为具体的教学设计行动。鉴于职前教师在训练过程中同样处于学习者角色,因此,本研究以纠正性反馈框架为基础,从情感、认知和行为三个层面探讨职前教师对不同来源反馈的感知、接受与应用方式,并据此设计反馈体验问卷。

2.5 研究问题

本研究围绕以下三个问题展开:

1. 不同类型的反馈能否有效提升职前教师的教学设计质量?
2. 相较于传统反馈,生成式人工智能反馈在教学设计质量提升方面的效果是否更为显著?
3. 不同反馈类型下,职前教师在情感、认知与行为层面的反馈参与有何差异?

3 研究方法

3.1 被试

本研究对象为中国某高校汉语国际教育专业的 68 名本科三年级学生。该群体经过本科阶段系统培养，属于国际中文职前教师，已具备基本的教学设计能力，能够完成简单的教案写作任务。同时，由于被试的教学能力仍处于发展阶段，对不同反馈类型的干预效果更为敏感。被试采用等量随机分组，每组 17 人，分为控制组、同伴反馈组、AI 反馈组及 AI+ 同伴反馈组。四组在年龄分布上无显著差异 ($p > .05$)，具体情况见表 1。所有被试在参与前均签署知情同意书。研究数据仅用于科研与教学并严格保密，符合伦理规范。

表 1 被试基本概况

组别	人数	年龄	标准差	性别
控制组	17	21.12	1.05	3 男，14 女
同伴组	17	21.41	0.80	2 男，15 女
AI 组	17	21.06	0.75	3 男，14 女
AI+ 同伴组	17	21.06	0.97	2 男，15 女

3.2 实验变量与研究工具

本研究采用 4 (反馈方式：控制组、同伴组、AI 反馈组、AI+ 同伴组) \times 2 (写作时间：初稿与修改稿) 的混合设计，反馈类型为被试间自变量，写作时间为被试内自变量。因变量包括：(1) 教案写作质量得分 (初稿、修改稿及差值)；(2) 反馈参与度问卷结果，涵盖自我意识、评估素养、目标设定、动机与参与、反馈处理五个维度，反映被试的认知、情感与行为反应。为确保实验变量测量的可靠性，本研究结合定量与定性工具，对研究任务与测评体系进行了规范设计。研究工具包括教案写作任务、教案质量评价量表与反馈体验测量问卷三部分。

3.2.1 教案写作任务

研究任务基于《HSK 标准教程 3》第三课，适用于 HSK 中级水平，既能考察职前教师对核心知识点的把握，又能检验其综合教学设计能力。选用统一课文可有效控制教学内容变量，减少材料差异带来的干扰。

3.2.2 教案质量评价量表

量表基于《国际中文教师专业能力标准》(T/ISCLT 001-2022) 构建，涵盖六个维度：(1) 教学目标明确性与可操作性；(2) 教学内容科学性与贴近性；(3) 教学流程完整性与时间安排合理性；(4) 教学活动设计的任务性与互动性；(5) 跨文化情境创设与适配性；(6) 教案语言规范性与术语准确性。每个维度采用 7 点 Likert 量表进行评定，所有教案由两名具有三年以上国际中文教学经验的教师依据统一评分标准进行独立评分，取平均分作为总分。

3.2.3 测量问卷

本研究的问卷设计基于 Ellis (2010) 的纠正性反馈框架,从情感、认知与行为三个维度切入,融合语言分析、人工智能辅助反馈视角,归纳为五个核心维度:(1)自我意识,个体对自身教案写作质量的自我评价;(2)评估素养,对反馈内容及评价标准的理解与应用;(3)目标设定,教案修改的目标明确性与可操作性;(4)动机与参与,修改的内在动机与投入程度;(5)反馈处理,接收反馈后的理解、采纳与应用。针对四组不同任务情境对题项数量与表述进行了相应调整,确保问卷既能反映各组的真实体验,又能保持跨组的可比性。问卷采用 Likert 五级量表,每个维度 2 道题,共 10 题,并配有开放式主观题。问卷通过“问卷星”平台统一发放与回收。为验证问卷的内部一致性信度,本研究在数据分析阶段采用 Cronbach's α 系数进行分析。结果表明,问卷整体信度良好,各组均达到可接受标准(详见 4.2.1 节),可用于后续统计分析与结果比较。受样本规模限制(每组 17 人),本文未对问卷进行因素分析以检验其结构效度。问卷的维度设置与题项设计基于 Ellis (2010) 的纠正性反馈理论框架,具有明确的理论依据,主要作为研究工具用于比较不同反馈条件下被试的体验差异,其效度更多体现在理论一致性与内容合理性层面,而非结构建模层面的统计验证。

3.3 实验流程

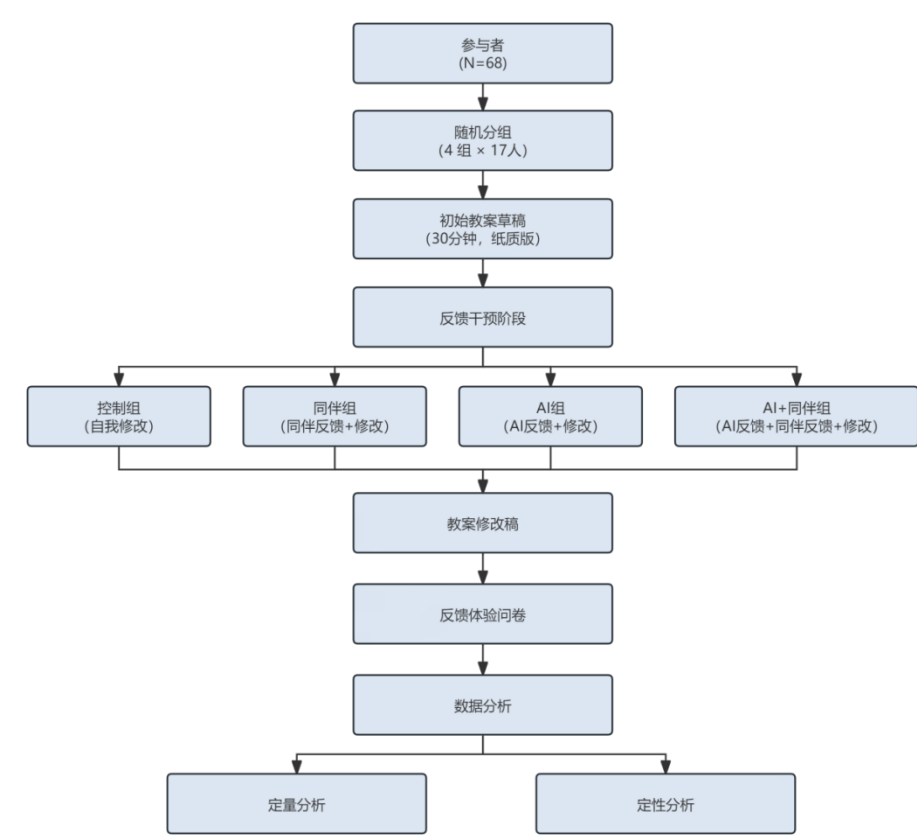
实验分两阶段、间隔一周进行,以减少短期记忆效应干扰并确保组间可比性(见图 1)。第一阶段为教案初稿写作,所有被试限时 30 分钟,纸笔独立完成。

第二阶段为教案修改,四组分别接受不同的反馈:控制组直接依据初稿进行自我修改;同伴组被试两两互评后修改;AI 组将教案初稿上传至 DeepSeek 平台,使用“深度思考 R1”模型生成反馈后修改;AI+ 同伴组先完成同伴互评,再结合 AI 反馈进行综合修改。AI 相关组的输入内容为被试撰写的教案初稿文本,不包含原课文材料,以确保反馈条件与同伴组一致并避免教学内容差异干扰。本次实验为 AI 提供统一指令语如下:“这是一份汉语作为二语教学的教案,请你以国际中文教师的视角进行评价,并提出 5 条具有针对性的修改建议。请围绕教学目标、教学内容、教学活动、教学流程及教学语言等方面提出改进意见。”AI 生成的反馈主要聚焦于教学目标的清晰度、教学活动设计的合理性、课堂流程的衔接性以及教学语言的规范性等方面。AI+ 同伴组在收到同伴反馈后,再以相同指令使用 AI 生成反馈,并整合两类意见进行修订。所有组别在完成修改稿后,均统一填写反馈体验问卷,以评估其在认知、情感与行为维度的反馈体验差异。

3.4 数据分析

本研究的数据分析分为定量与定性两部分,所有统计分析均在 SPSS 29.0 软件中完成,显著性水平设定为 $p < .05$ 。定量部分通过信度检验、方差分析及效应量计算等方法比较四组在教案修改中的差异,定性部分则对问卷与访谈资料进行主题归纳与词频分析,以揭示职前教师对不同反馈的参与情况。

图 1. 实验流程图



4 研究结果

4.1 定量数据研究结果

4.1.1 评分信度与评分者一致性结果

为确保评分结果的信度与稳定性，本研究对初稿与修改稿的评分数据进行了 Pearson 相关分析、内部一致性分析和 ICC 分析进行检验。结果显示，两阶段评分均具有较高一致性(见表 2)。

表 2. 初稿与修改稿的评分信度与评分者一致性结果

	初稿	修改稿
Pearson 相关系数 (r)	0.731	0.881
ICC(2,k)	0.804	0.931
Cronbach's α	0.804	0.931

4.1.2 单因素方差分析结果

为比较四组在教案初稿与修改稿阶段的差异，本文对两阶段评分进行了正态性检验、方差齐性检验及单因素方差分析（见表 3）。初稿阶段除同伴组外均满足参数检验前提条件；同伴组轻微偏离正态（ $p = .027$ ），但因样本量相近且偏态程度有限，整体数据满足参数检验前提。组

间差异不显著 ($p = .810$), 效应量 $\eta^2 = 0.015$, 表明各组初始水平均衡。修改稿阶段同样满足参数检验前提, 组间差异显著 ($p < .001$), 效应量 $\eta^2 = 0.461$, 表明反馈类型对学习产出具有较大解释力。

表 3. 正态性与方差齐性检验结果

阶段	正态性检验 (Shapiro-Wilk)	方差齐性检验 (Levene)	ANOVA F(df)	p 值	η^2
初稿	控制组 $p=.972$	$F=0.327, p=.806$	$0.321(3,64)$	$.810$	0.015
	AI 组 $p=.112$				
	同伴组 $p=.027^*$				
	AI+ 同伴组 $p=.700$				
修改稿	控制组 $p=.753$	$F=1.088, p=.360$	$18.195(3,64)$	$<.001$	0.461
	AI 组 $p=.744$				
	同伴组 $p=.118$				
	AI+ 同伴组 $p=.502$				

4.1.3 组内前后测差异分析

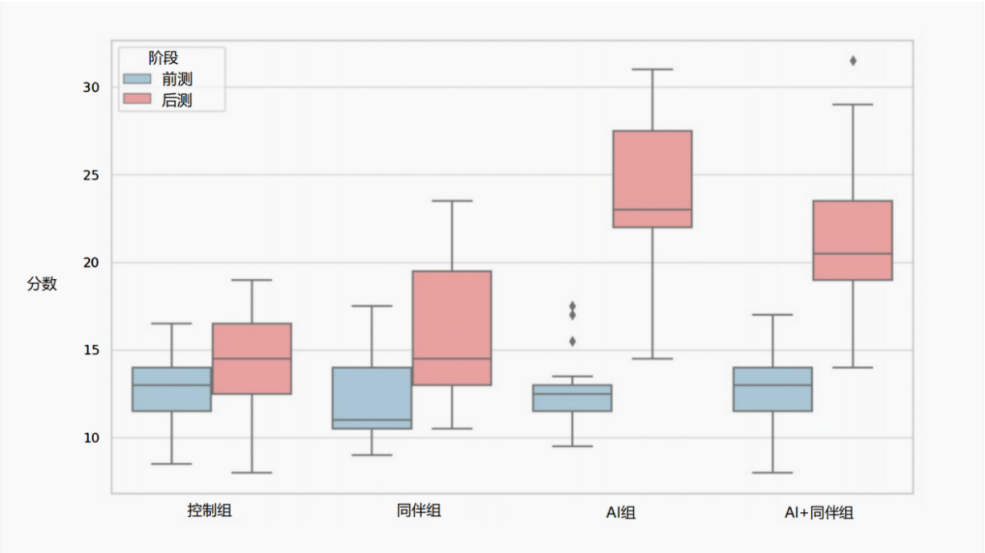
为进一步检验反馈在教案修改阶段的实际成效, 配对样本 t 检验结果 (见表 4) 显示, 四组被试在修改阶段得分均显著高于初稿 ($p < .05$), 效应量从中等到极大不等。其中, AI 组和 AI+ 同伴组的提升幅度最大 (Cohen's $d > 1.9$), 显著优于同伴组 ($d = 0.972$) 与控制组 ($d = 0.624$), 表明 AI 反馈尤其是与同伴反馈结合的形式, 在提升教案写作质量方面具有显著优势 (见图 2)。

表 4. 配对样本 t 检验结果

组别	初稿平均分	修改稿平均分	T 值 (df=16)	显著性 (p 值)	效应量 (Cohen'sd)
控制组	12.76	14.32	2.572	.020*	0.624
同伴组	12.03	16.15	4.007	.001**	0.972
AI 组	12.65	24.00	10.628	<.001***	2.578
AI+ 同伴组	12.56	21.38	8.085	<.001***	1.961

* $p < .05$, ** $p < .01$, *** $p < .001$

图 2. 四组被试教案初稿与修改稿得分的比较



4.1.4 交互作用与主效应分析

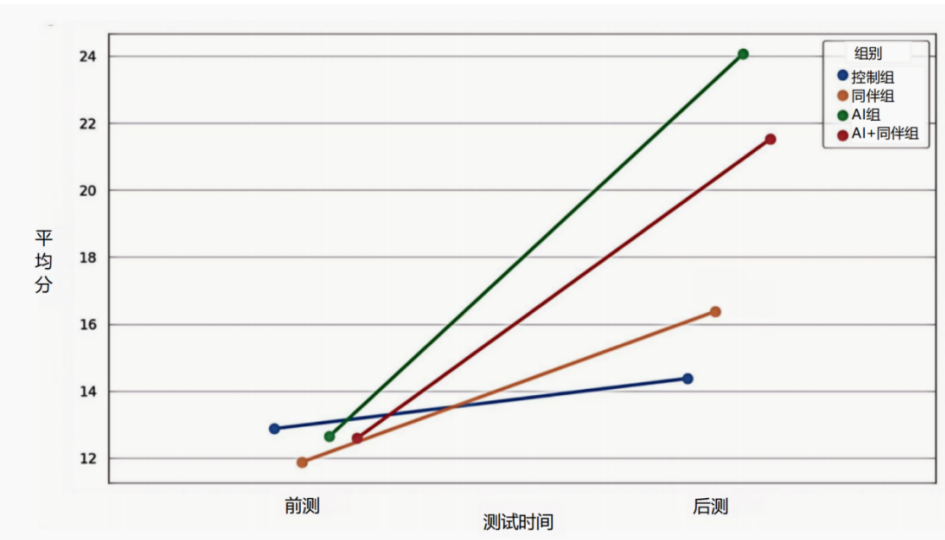
采用两因素重复测量方差分析,测试时间(初稿、修改稿)为被试内因素,组别(控制组、AI组、同伴组、AI+同伴组)为被试间因素。结果显示(表5),时间主效应显著, $F(1,64)=177.96, p<.001, \eta^2=0.735$, 修改稿得分高于初稿;组别主效应显著, $F(3,64)=13.31, p<.001, \eta^2=0.383$, 不同反馈效果存在差异;时间 \times 组别交互效应显著, $F(3,64)=20.96, p<.001, \eta^2=0.496$ 。图3显示,各组修改稿得分均提升,其中AI组和AI+同伴组提升幅度最大,显著优于同伴组与控制组,凸显AI反馈在提升教案质量方面的优势。

表 5. 教案写作表现的两因素重复测量方差分析结果

	F 值	df1	df2	P 值	偏 Eta 平方
测试时间	177.96	1	64	<.001***	0.735
组别	13.31	3	64	<.001***	0.383
测试时间 x 组别	20.96	3	64	<.001***	0.496

* $p < .05$, ** $p < .01$, *** $p < .001$

图 3. 测试时间与组别对教案评分的交互效应



4.1.5 组间差异的事后比较结果

为检验修改稿得分的组间差异,采用 Tukey HSD 法进行事后两两比较(见表6)。结果显示,AI组得分显著高于同伴组($p < .001$)和控制组($p < .001$),与AI+同伴组差异不显著($p = .302$);AI+同伴组显著高于同伴组($p = .004$)和控制组($p < .001$),但与AI组差异不显著;同伴组与控制组差异不显著($p = .613$)。效应量分析表明,AI组相较控制组($d = 2.51$)和同伴组($d = 1.76$)为极大效应,AI+同伴组相较控制组($d = 1.67$)和同伴组($d = 1.10$)为中到大效应,AI组与AI+同伴组间为中等效应($d = 0.53$)。结果表明,AI主导的干预在统计和实践层面均具有显著优势。

表 6. 修改稿得分的组间差异 Tukey HSD 事后比较结果

组别	平均差 (I-J)	标准差	显著性 (p)	95% 置信区间下限	95% 置信区间上限
控制组 - 同伴组	-1.823	1.488	.613	-5.748	2.101
控制组 -AI 组	-9.676*	1.488	<.001***	-13.601	-5.752
控制组 -AI+ 同伴组	-7.059*	1.488	<.001***	-10.984	-3.135
同伴组 -AI 组	-7.852*	1.488	<.001***	-11.778	-3.927
同伴组 -AI+ 同伴组	-5.235*	1.488	.004**	-9.160	-1.300
AI 组 -AI+ 同伴组	2.617	1.488	.302	-1.307	6.543

*p < .05, **p < .01, ***p < .001

4.2 定性数据研究结果

4.2.1 问卷内部一致性

为检验问卷的内部一致性,分别计算了四组的 Cronbach’s α 系数。结果显示,控制组(0.811)与 AI 反馈组 (0.835) 均达良好水平,同伴反馈组 (0.725) 与 AI+ 同伴反馈组 (0.722) 处于可接受范围。总体而言,四组均达到研究可接受标准,可用于后续组间比较与统计分析。

4.2.2 描述性统计分析

为比较四组在五个核心维度的表现,计算了各题项的均值 (M) 与标准差 (SD) (见表 7)。结果显示,控制组得分较均衡,评估素养最高,反馈理解与运用最低。同伴组在动机与参与维度表现最优,显示其在激发学习投入方面的积极作用。AI 组在评估素养得分最高,但反馈理解与运用相对较低。AI+ 同伴组在动机与参与表现最佳,但反馈理解与运用最低。总体而言,不同反馈机制在评估素养、动机与参与等维度呈现差异化趋势,后续将进一步检验显著性差异。

表 7. 四组在五个核心维度上的描述性统计结果

组别	自我觉察与诊断 (SD)	评估素养 (SD)	目标设定与计划 (SD)	动机与参与 (SD)	反馈理解与运用 (SD)
控制组	3.79(0.59)	3.91(0.62)	3.82(0.83)	3.74(0.92)	3.62(0.72)
同伴组	3.74(0.69)	3.76(0.66)	3.94(0.58)	4.21(0.56)	3.82(0.77)
AI 组	3.88(0.72)	4.21(0.61)	3.91(0.87)	4.06(0.90)	3.68(0.66)
AI+ 同伴组	4.12(0.65)	3.76(0.53)	4.00(0.53)	4.26(0.66)	3.38(0.70)

4.2.3 数据分布特征与非参数检验结果

正态性检验 (Kolmogorov–Smirnov 与 Shapiro–Wilk) 显示部分变量显著偏离正态,因此采用非参数 Kruskal–Wallis H 检验分析。结果表明,“评估素养”维度在四组间差异显著 ($\chi^2 = 8.301$, $df = 3$, $p = .040$, $\varepsilon^2 = 0.124$, 中等效应)。表明不同反馈对该题项所反映的表现具有显著影响。Bonferroni 校正的两两比较显示,“同伴组 -AI 组” ($p = .019$) 与 “AI 组 -AI+ 同伴组” ($p = .033$) 在调整前接近显著,但调整后均不显著 ($p > .05$) (见表 8)。总体来看,不同干预方式在“评估素养”上的影响存在中等强度,但具体组间差异在多重比较下未达显著。

表 8. 四组在“评估素养”维度的非参数检验结果

组别	检验统计	标准差	标准检验统计	显著性	调整后显著性 *
控制组 - 同伴组	-11.588	6.183	-1.874	.061	.366
控制组 -AI 组	-2.971	6.183	-.480	.631	1.000
控制组 -AI+ 同伴组	-10.206	6.183	-1.651	.099	.593
同伴组 -AI 组	-14.559	6.183	-2.355	.019	.111
同伴组 -AI+ 同伴组	-1.382	6.183	-.224	.823	1.000
AI 组 -AI+ 同伴组	-13.176	6.183	-2.131	.033	.199

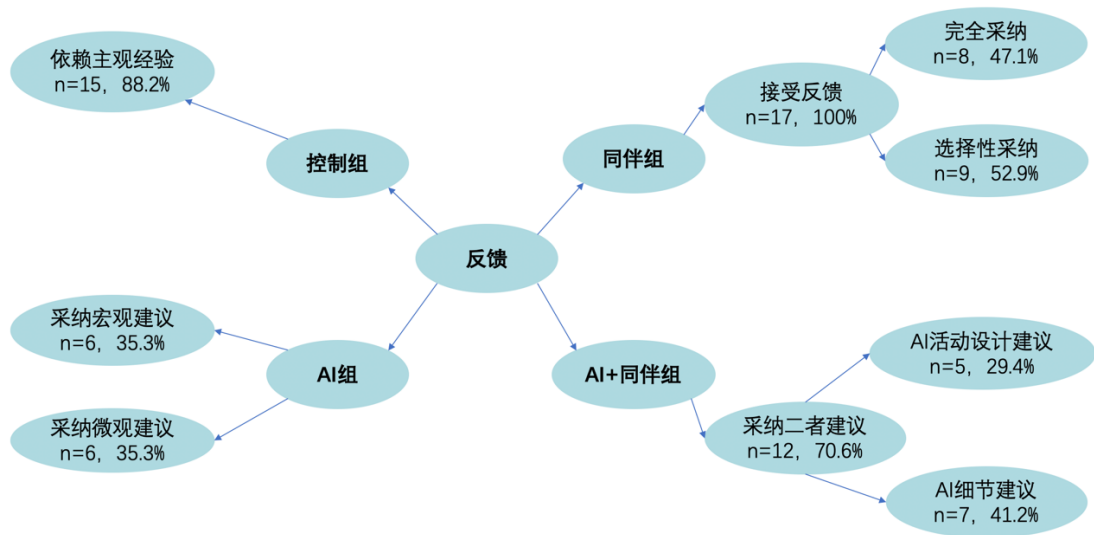
* 已针对多项检验通过 Bonferroni 校正法调整显著性

4.3 开放式问题结果分析

4.3.1 反馈类型与教案修改策略的关联分析

如图 4 所示，不同反馈类型在修改策略上差异明显。控制组大多数被试（88.2%，15 人，）依赖经验与直觉。同伴组全部被试均采纳反馈，其中 47.1%（8 人）完全采纳，52.9%（9 人）选择性采纳，主要根据个人目标筛选。AI 反馈组被试均能识别关键信息，35.3%（6 人）采纳宏观设计建议，35.3%（6 人）采纳细节补充。AI+ 同伴组发挥了人机协同的作用，超半数被试（70.6%，12 人）同时采纳了 AI 和同伴反馈的内容与结构建议。29.4%（5 人）和 41.2%（7 人）的被试接受了 AI 在活动设计与细节补充方面的建议。

图 4. 不同反馈类型下职前教师教案修改策略的分布

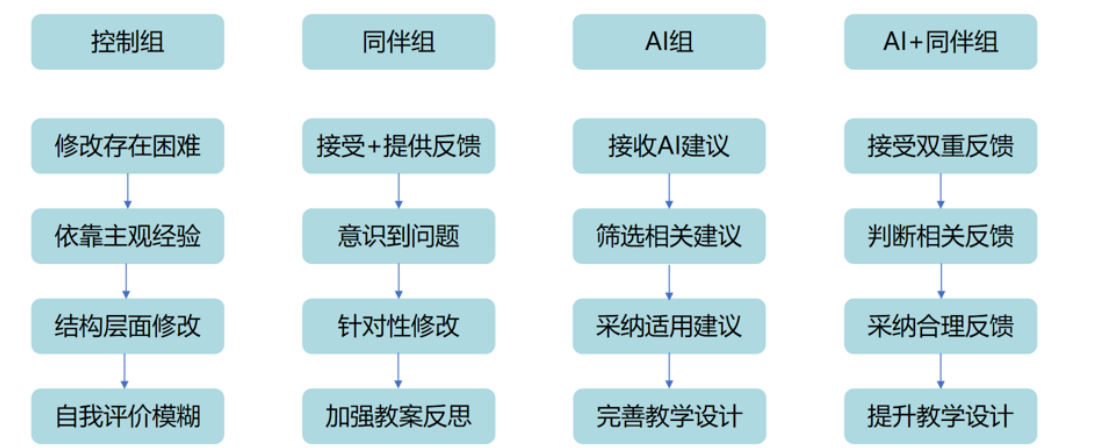


4.3.2 反馈采纳路径中的认知差异

如图 5 所示，控制组多数被试（58.8%，10 人）表示在结构与流程修改上存在困难，35.3%（6 人）直言“不知道如何修改”，整体以表层调整为主（结构 76.5%，13 人；内容 23.5%，4 人）。虽有 76.5%（13 人）对修改稿“有信心”，但因缺乏外部反馈，判断标准模糊。同伴组全员认可反馈，70.6%（12 人）在为他人提供建议后意识到自身问题并进行针对性修改。AI 组获得了

细节方面的指导，35.3%（6人）拓展教学思路，但23.5%（4人）认为其脱离教学实际。AI+同伴组中，76.5%（13人）倾向于结合实际情况综合判断，仅5.9%（1人）更偏向AI。对于AI反馈的评价，部分被试认为其全面（47.1%，8人）且高效（23.5%，4人），但也有人指出其存在文本识别障碍（29.4%，5人）、内容笼统缺乏针对性（47.1%，8人），甚至可能导致依赖（11.8%，2人）。

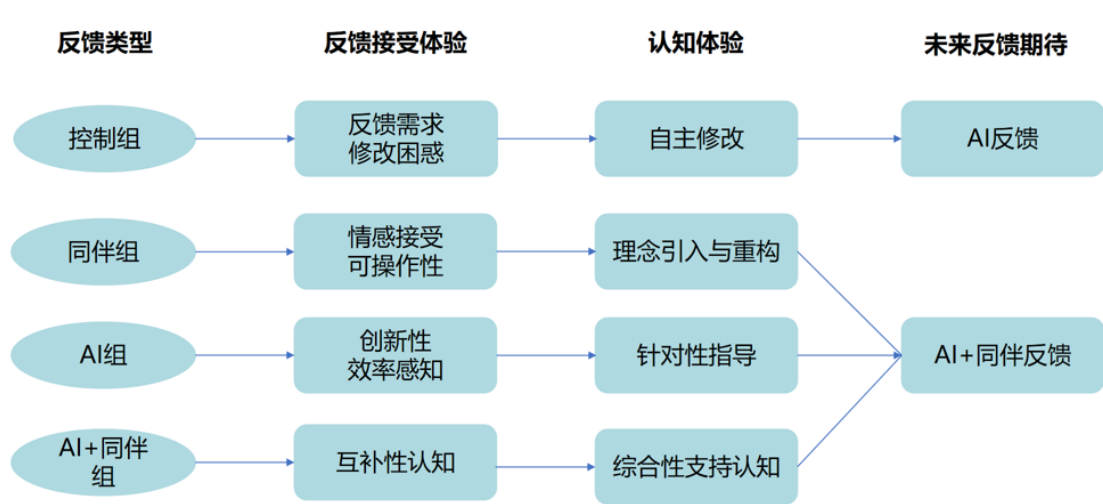
图 5. 四组的反馈采纳路径



4.3.3 反馈接受体验与未来偏好倾向

如图 6 所示，控制组全体被试希望获得更具针对性的反馈，重点集中在教学步骤（52.9%，9人）和内容（47.1%，8人）。同伴组 94.1%（16人）期待未来引入 AI 反馈，但也有少数人认为同伴反馈更贴近自身水平。在反馈接受体验方面，部分被试肯定同伴反馈的情感支持（35.3%，6人）和可操作性（29.4%，5人），但存在不够全面（52.9%，9人）、与个人想法冲突（17.6%，3人）等问题。AI 组认为其优势在于全面（47.1%，8人）与即时高效（29.4%，5人），不足在于内容笼统（41.2%，7人）、识别局限（35.3%，6人）及速度不稳定（11.8%，2人）。AI+同伴组从 AI 反馈中获取理论支持，从同伴反馈中获得实践启发。在未来期望中，76.5%（13人）倾向 AI+同伴结合反馈，仅 17.6%（3人）愿只接受 AI 反馈。

图 6. 四组职前教师的反馈接受体验与未来偏好



总体来看,与单纯依赖个人经验相比,获得反馈的被试能更系统地优化教学设计,并提升反思与自我监控等元认知能力。AI反馈与同伴反馈在功能侧重上存在差异,而AI+同伴反馈在教学设计中具有显著的协同优势。

5 讨论

5.1 反馈类型的整体效应与AI、双重反馈的比较优势

本研究发现,四组教案修改稿得分均显著高于初稿,表明反馈显著提升了职前教师教案设计质量(Lee, 2017; Kang & Han, 2015; Vasu et al., 2022)。其中, AI组提升最显著, AI+同伴组次之;同伴组与控制组的提升相对有限。这表明AI反馈能在短期内有效激发教学认知并优化教学设计。同时, AI具备自适应学习能力,可根据使用者水平与偏好调整提示(Jain & Jain, 2019),为职前教师提供个性化指导。这种特性突破了传统反馈评价标准单一的局限(Mureşan, 2023; Okunade, 2024),并兼具认知与元认知支持功能,帮助教师识别问题、形成方案并促进反思(Drake et al., 2014; Ge & Land, 2003; Winstone & Boud, 2022)。然而, AI反馈仍存在语义模板化、术语堆砌等问题,缺乏对具体教学情境的理解与适配,常需教师或同伴补充(Pishtari et al., 2023; Ndjama, 2025)。

AI+同伴组体现了双重反馈的互补效应。在反馈过程中, AI通常指出教案结构层级不清或教学目标表述模糊的问题,“AI指出我的教学目标写得不清楚,还帮我润色了语句,让目标更明确、表达更通顺”。与此同时,同伴反馈则更多关注教学情境的可行性与学生感受,“同伴的意见更贴近教学实际,比如提醒我活动设计要考虑学生兴趣和语言水平”,提出了更贴近课堂实际的建议。AI的标准化指导与同伴的经验补充形成互补,前者优化了教案的逻辑结构与语言规范,后者增强了教学设计的情境契合度与课堂可行性,从而促进教师在反思与自我监控中不断完善教学设计与专业思维。尽管AI组与AI+同伴组之间的差异不显著,但并不意味着双重反馈无效。相反,这表明AI能成为弥补经验不足的“认知脚手架”(Singh & Mueller, 2021)。双重反馈还能将抽象术语转化为易懂的教学语言,增强实用性与可操作性(Pishtari et al., 2023),但也可能因意见冲突和信息筛选带来认知负荷。此外,在AI+同伴组中,本研究采用了“先同伴反馈、后生成式AI反馈”的固定顺序安排。该顺序主要基于教学实施的可行性与认知加工的考虑:同伴反馈有助于激活学习者的问题意识,而生成式AI在结构整合与系统性建议方面具有优势,置于后位更有利于对前序反馈进行比对与整合。然而,这一顺序安排也可能引入近因效应,即学习者在修改过程中更倾向于采纳最后接收到的反馈建议,从而在一定程度上放大AI反馈的影响。因此,关于AI与双重反馈效果的比较结果仍需谨慎解读。未来研究可通过操控反馈顺序或采用平衡设计,进一步考察不同顺序对反馈采纳行为与学习成效的影响。从研究结果看,同伴组虽有提升,但反馈主要依赖个人经验,缺乏明确的修改步骤和逻辑,成效有限。控制组提升幅度最小,表明在缺乏外部反馈的情况下,职前教师主要依赖自我调控与经验,修改结果缺乏方向与深度。Tukey HSD事后比较进一步验证了仅依赖同伴互评成效有限,而AI或人机协同能提供更清晰和专业的反馈。

综上所述,本研究支持“AI应作为人类反馈的补充而非替代”的观点(Zhang & Ilisko, 2025)。在教学实践中, AI反馈适合满足效率和成果的即时需求,而人机协同反馈更能促进教师专业素养的持续发展,因此应根据具体目标灵活选择。

5.2 AI 反馈促进职前教师的教学设计能力提升

问卷数据显示, AI 组在“评估素养”维度得分最高, 表明 AI 反馈对职前教师理解并内化教学要素具有显著促进作用。AI 具备即时交互能力, 能够提供条理清晰、针对性强的反馈 (Steiss et al., 2024), 引导职前教师掌握教学逻辑 (Rahman et al., 2022), 还能够根据学习者的认知水平动态调整反馈内容 (Su & Yang, 2023)。相比之下, 同伴反馈虽然在互评过程中促进了教学反思, 但缺乏明确标准, 部分被试仍凭个人经验进行判断, 导致评估效度受限。控制组对教学构成要素掌握不足, 部分被试甚至表示“不了解优秀教案需要具备哪些要素”, 说明在缺乏外部反馈的条件下, 自我认知发展明显受限。尽管组间总体差异显著, 但 Bonferroni 校正后的两两比较未达显著水平, 这可能与职前教师对反馈的理解、吸收与转化能力存在较大的个体差异有关。这一结果表明, 反馈并非对所有学习者产生同质化效果, 在教学实践中有必要引导职前教师发挥主体作用, 通过人机协同机制, 将反馈转化为可操作的教学实施方案, 而非停留于表层理解。

为进一步阐释人机协同具体运作方式, 本文引入乔哈里视窗模型作为解释隐喻, 将反馈过程中的信息划分为公开区、隐藏区、盲区与未知区 (董艳等, 2023)。公开区为双方共知的信息; 隐藏区是职前教师所掌握而 AI 无法获取的学生背景与期望等; 盲区代表 AI 所知但职前教师未察觉的知识盲点, 如教学设计环节与教学逻辑的深层链接; 而未知区则指向职前教师与 AI 在反馈互动中拓展出的新的教学理解与教学策略。有效的人机协同通过多轮交流揭示隐藏区、缩小盲区、扩展公开区并探索未知区。在这一过程中, 教师应发挥主导作用, 通过向 AI 明确教学目标、学情与需求, 提高反馈的针对性与适切性; AI 则以结构化反馈和知识储备支持职前教师的反思与决策。由此可见, 人机协同并非取代教师, 而是通过内容反馈与激活元认知, 帮助职前教师搭建专业发展的认知基础, 从而提升教学能力。

5.3 职前教师的反馈体验呈多维特征

本研究从“修改策略”“反馈采纳”和“体验偏好”三个维度对反馈路径进行分析。

修改策略方面, 控制组高度依赖经验与直觉, 仅做表层调整, 存在认知负荷大、策略单一等问题。同伴组在互评中有选择性地采纳建议, 但受个人经验与知识储备限制, 难以形成专业指导 (Drake et al., 2014; Esterhazy et al., 2021)。AI 组能识别并采纳结构与细节层面的优化提示, 同时持有批判性态度, 选择性接受建议 (Bai & Hu, 2017)。AI+ 同伴组整合能力更强, 修改结果兼具可行性与创新性, 并且能够调动元认知能力。这体现出双重反馈在促进教师反思和元认知发展上的优势。

采纳策略方面, 控制组多集中于教学流程调整, 缺乏教学理念、目标与学情的深度分析。同伴组通过互评形成了初步的教学反思, 但评价标准不明确。AI 组采纳结构与细节方面的优化建议, 拓展了教学思路。AI+ 同伴组能够在冲突意见中权衡择优, 体现出一定的批判意识 (Tay, 2024), 逐步加深对教学要素的理解。

反馈体验方面, 控制组普遍渴望收到系统的、具体的反馈, 尤其偏好专业性与操作性强的指导。同伴组尽管获得了一定的反馈支持, 但普遍表现出对 AI 反馈引入的期待。部分被试指出, 同伴建议在情境契合度和可理解度方面更具优势, 但也存在“反馈内容零散”“缺乏条理性”等问题。这表明, 被试已经意识到多元反馈的互补价值, 并形成了对人机协同模式的高度认可。AI 组因反馈响应迅速、结构清晰且覆盖内容全面, 普遍受到被试的认可。然而 DeepSeek 等大语言模型在情感识别、意图判断与常识推理等方面仍存在显著短板 (Atlas, 2023; Hong, 2023)。同时, AI 的过度介入可能削弱教师与学习者之间的互动联结, 尤其在课程规划与教学

评估等环节,若教师被系统功能替代,可能引发教学个性化的缺失与教师专业能力的弱化(Felix & Webb, 2024)。AI+ 同伴组普遍表现出对复合反馈模式的高度认同,这一偏好不仅体现了对认知支持与情感支持整合需求的反映,也为今后教学实践中构建多元混合反馈提供了现实依据。由此可见,教师在 AI 反馈系统的设计与使用中并非技术附庸,而应理解并转换相关建议,从而推动教学设计的高质量发展(Ruiz-Martín & Bybee, 2022)。这一发现表明,有效利用 AI 反馈的关键在于兼顾内容与互动过程,从而推动使用者从被动的信息接收者转变为具备批判性筛选与判断能力的主动学习者。

综合来看, AI 与同伴反馈各具优势: AI 在结构优化与逻辑梳理方面具有优势,但在情境理解与情感支持上仍有不足;同伴反馈虽能契合教学情境,但缺乏专业性。人机协同反馈能通过优势互补实现认知与情感的平衡。正如丁安琪、蒙小凤(2025)所说, AI 工具能够为新手教师的教学设计提供“合规性参照”,但要实现深度赋能还需要构建“AI 生成基础框架+教师主导本土化适配”人机协同模型。因此,未来反馈应以人机协作而非技术替代为发展方向:教师发挥专业判断与情感支持作用, AI 提供即时性与专业化指导(Lee & Zhai, 2024),不仅助力教学设计,也为教师专业成长提供了新契机。

6 结语

本研究以国际中文职前教师为研究对象,基于四组对照实验,系统比较了 AI 与传统反馈在教案写作质量提升及反馈体验上的差异,验证了生成式人工智能赋能国际中文教学设计的潜在价值。研究发现,生成式人工智能在优化教学逻辑与结构方面具有显著优势,能促进教师对专业知识的理解与应用;人机协同反馈具有互补效应。同时,职前教师对 AI 反馈普遍持积极态度,认可其在效率和逻辑上的优势,但也意识到 AI 在情境适配上的局限。因此, AI 反馈不应局限于技术层面的工具操作,而应作为支持教师专业成长的协作机制。有必要将 AI 教学法与专业发展能力纳入职前教师培养体系,帮助其掌握 AI 融合教学策略的设计与实施能力。本研究初步考察了 AI 反馈在提升职前教师教学设计质量方面的作用,但本研究样本量相对有限(每组 17 人),尽管统计分析满足基本要求,但在结果解释与推广时仍需保持谨慎,未来研究有必要在更大样本或多情境条件下进一步验证研究结论。在人机协同反馈条件中,本研究采用固定的反馈顺序(同伴反馈在前、AI 反馈在后),未对不同反馈顺序进行比较,该设计可能引入顺序效应,对相关结果的解释构成一定限制。此外, AI 生成与经反馈修改的教案在真实课堂中能否有效促进学生的语言技能与知识习得,仍有待进一步实证检验。未来研究可从动态视角探讨人工智能在国际中文教师专业发展中的作用,并基于生成内容构建共享教学资源平台,推动国际中文教育的持续创新与发展。

参考文献

- Atlas, S. (2023). ChatGPT for higher education and professional development: A guide to conversational AI. University of Rhode Island, DigitalCommons@URI.
- Bai, L., & Hu, G. (2017). In the face of fallible AWE feedback: How do students respond? *Educational Psychology*, 37(1), 67–81. <https://doi.org/10.1080/01443410.2016.1223275>
- Biggs, J., Tang, C., & Kennedy, G. (2022). *Teaching for quality learning at university* (5th ed.). McGraw-Hill Education (UK).

- Bin, Shuai (宾帅), Wang, Yamin (王亚敏), & Hu, Renfen (胡韧奋). (2025). 基于大语言模型的中文语法教学资源生成测评研究 [An evaluation study on LLM-based resource generation for Chinese grammar teaching]. *世界汉语教学 [Chinese Teaching in the World]*, 39(3), 405–419. <https://doi.org/10.13724/j.cnki.ctiw.2025.03.011>
- Cai, Wei (蔡薇). (2023). ChatGPT 环境下的汉语学习与教学 [Learning and teaching Chinese in the ChatGPT context]. *语言教学与研究 [Language Teaching and Linguistic Studies]*, 4, 13–23.
- Chapelle, C. A. (2025). Generative AI as game changer: Implications for language education. *System*, 132, 103672. <https://doi.org/10.1016/j.system.2025.103672>
- Cheng, X., Liu, Y., & Wang, C. (2023). Understanding student engagement with teacher and peer feedback in L2 writing. *System*, 119, 103176. <https://doi.org/10.1016/j.system.2023.103176>
- Chu, Chengzhi (储诚志). (2025). 人工智能时代国际中文教育体系重构的几个基底性问题 [Several Foundational Issues in Reconstructing the System of International Chinese Education in the Age of Artificial Intelligence]. *世界汉语教学 [Chinese Teaching in the World]*, 39(3), 292–294. <https://doi.org/10.13724/j.cnki.ctiw.2025.03.010>
- Cui, Xiliang (崔希亮). (2025). DeepSeek 如何为国际中文教育赋能 [How DeepSeek Empowers International Chinese Education]. *世界汉语教学 [Chinese Teaching in the World]*, 39(3), 294–295. <https://doi.org/10.13724/j.cnki.ctiw.2025.03.010>
- Cui, Yonghua (崔永华). (2008). 对外汉语教学设计导论 [An introduction to teaching Chinese as a foreign language]. 北京语言大学出版社 [Beijing Language and Culture University Press].
- Ding, Anqi (丁安琪), & Meng, Xiaofeng (蒙小凤). (2025). 从模仿到重构: AI 生成教案对国际中文教育新手教师教学设计能力的影响 [From imitation to reconstruction: The impact of AI-generated lesson plans on the instructional design competence of novice teachers]. *贵州师范大学学报 (社会科学版) [Journal of Guizhou Normal University (Social Science Edition)]*, 5, 42–52. <https://doi.org/10.16614/j.gznuj.skb.2025.05.005>
- Ding, Anqi (丁安琪). (2014). 美国“语言教师效能框架”探析 [An analysis of the “Language Teacher Effectiveness Framework” in the United States]. *中国教育学刊 [Chinese Journal of Education]*, 9, 94–98.
- Dong, Yan (董艳), Xia, Liangliang (夏亮亮), Li, Xinyi (李心怡), & Hou, Yanhua (侯彦华). (2023). ChatGPT 赋能学生学习的途径探析 [Analysis of the path to empowering student learning with ChatGPT]. *电化教育研究 [E-Educational Research]*, 44(12), 14–20. <https://doi.org/10.13811/j.cnki.eer.2023.12.002>
- Drake, C., Land, T. J., & Tyminski, A. M. (2014). Using educative curriculum materials to support the development of prospective teachers’ knowledge. *Educational Researcher*, 43(3), 154–162. <https://doi.org/10.3102/0013189X14528039>
- Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63(2), 97–107. <https://doi.org/10.1093/elt/ccn023>
- Ellis, R. (2010). Epilogue: A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, 32(2), 335–349. <https://doi.org/10.1017/S0272263109990544>
- Esterhazy, R., de Lange, T., Bastiansen, S., & Wittek, A. L. (2021). Moving beyond peer review of teaching: A conceptual framework for collegial faculty development. *Review of Educational Research*, 91(2), 237–271. <https://doi.org/10.3102/0034654321990721>
- Felix, J., & Webb, L. (2024). *Use of artificial intelligence in education delivery and assessment*. Parliamentary Office of Science and Technology.
- Gagné, R. M., Briggs, L. J., & Wager, W. W. (加涅 R. M., 布里格斯 L. J., 瓦格 W. W.). (1992/2007). 教学设计原理 (王小明等译, 第五版) [Principles of instructional design (5th ed.)]. 华东师范大学出版社 [East China Normal University Press].

- Ge, X., & Land, S. (2003). Scaffolding students' problem-solving processes in an ill-structured task using question prompts and peer interactions. *Educational Technology Research and Development*, 51(1), 21–38. <https://doi.org/10.1007/BF02504515>
- Han, Kai (韩开), & Xu, Juan (徐娟). (2023). 基于 AIGC 的国际中文教育教学设计创新 [Innovative practices in international Chinese education teaching design based on AIGC] [A]. 中文教学现代化研究——2023 年专题文集 [Studies on the Modernization of Chinese Language Teaching—Selected Papers 2023] [C]. 中文教学现代化学会 [Society for Modernization of Chinese Language Teaching], 13. <https://doi.org/10.26914/c.cnkihy.2023.124946>
- Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, 5(1). <https://doi.org/10.61414/jeti.v5i1.103>
- Hu, Zhuanglin (胡壮麟). (2023). ChatGPT 谈外语教学 [ChatGPT's talk about foreign language education]. 中国外语 [Chinese Foreign Languages], 20(3), 12–15. <https://doi.org/10.13564/j.cnki.issn.1672-9382.2023.03.003>
- Jacobsen, L. J., & Weber, K. E. (2025). The promises and pitfalls of large language models as feedback providers: A study of prompt engineering and the quality of AI-driven feedback. *AI*, 6(2), 35. <https://doi.org/10.3390/ai6020035>
- Jain, S., & Jain, R. (2019). Role of artificial intelligence in higher education: An empirical investigation. *IJRAR-International Journal of Research and Analytical Reviews*, 6(2), 144–150.
- Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal*, 99(1), 1–18. <https://doi.org/10.1111/modl.12189>
- Koltovskaia, S., Rahmati, P., & Saeli, H. (2024). Graduate students' use of ChatGPT for academic text revision: Behavioral, cognitive, and affective engagement. *Journal of Second Language Writing*, 65, 101130. <https://doi.org/10.1016/j.jslw.2024.101130>
- Lee, G. G., & Zhai, X. (2024). Using ChatGPT for science learning: A study on pre-service teachers' lesson planning. *IEEE Transactions on Learning Technologies*, 17, 1643–1660. <https://doi.org/10.1109/TLT.2024.3401457>
- Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Springer Singapore.
- Liao, Jianling (廖建玲). (2013). 国际汉语教学设计 [Instructional design in international Chinese language teaching]. 高等教育出版社 [Higher Education Press].
- Ma, N., Xin, S., & Du, J. Y. (2018). A peer coaching-based professional development approach to improving the learning participation and learning design skills of in-service teachers. *Educational Technology & Society*, 21(2), 291–304.
- Ma, X., Zhang, C., & Lee, I. (2025). *Learner engagement with AI-generated feedback among Chinese EFL students on second language writing* [Unpublished manuscript]. SSRN. <http://dx.doi.org/10.2139/ssrn.5197066>
- Meng, Kai (孟凯). (2024). ChatGPT 在国际中文词汇教学中的表现与应用策略 [The performance and application strategies of ChatGPT in CSL vocabulary teaching]. 国际汉语教学学报 [International Journal of Chinese Language Teaching], 5(1). <https://doi.org/10.46451/ijclt.20240102>
- Meyer, J., Jansen, T., Schiller, R., Liebenow, L. W., Steinbach, M., Horbach, A., & Fleckenstein, J. (2024). Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions. *Computers and Education: Artificial Intelligence*, 6, 100199. <https://doi.org/10.1016/j.caeai.2023.100199>
- Mureşan, M. (2023, August). Impact of artificial intelligence on education. In *Proceedings of the 32nd International RAIS Conference on Social Sciences and Humanities* (pp. 81–85). Scientia Moralitas Research Institute.

- Ndjama, J. D. N. (2025). The use of artificial intelligence in lesson delivery and evaluation in large-scale teaching environments. *Interdisciplinary Journal of Education Research*, 7(s1), 1-21. <https://doi.org/10.38140/ijer-2025.vol7.s1.08>
- Novita, R. (2023). AI in lesson planning: Improving teacher efficiency and instructional design. *Education and Information Technologies*, 28(5), 5673–5692. <https://doi.org/10.55606/jurripen.v4i2.5560>
- Okunade, A. I. (2024). The role of artificial intelligence in teaching of science education in secondary schools in Nigeria. *European Journal of Computer Science and Information Technology*, 12(1), 57–67. <https://doi.org/10.37745/ejcsit.2013/vol12n15767>
- Pishtari, G., Sarmiento-Márquez, E. M., Rodríguez-Triana, M. J., Wagner, M., & Ley, T. (2023, August). Evaluating the impact and usability of an AI-Driven feedback system for learning design [Paper presentation]. In *European Conference on Technology Enhanced Learning* (pp. 324-338). Cham: Springer Nature Switzerland.
- Rahman, N., Zulkornain, H., & Hamzah, H. (2022). Exploring artificial intelligence using automated writing evaluation for writing skills. *Environment-Behaviour Proceedings Journal*, 7, 547–553. <https://doi.org/10.21834/ebpj.v7iS19.4304>
- Ruiz-Martín, H., & Bybee, R. W. (2022). The cognitive principles of learning underlying the 5E model of instruction. *International Journal of STEM Education*, 9(1), 21.
- Ruwe, T., & Mayweg-Paus, E. (2023). “Your argumentation is good”, says the AI vs humans—The role of feedback providers and personalised language for feedback effectiveness. *Computers and Education: Artificial Intelligence*, 5, 100189. <https://doi.org/10.1016/j.caeai.2023.100189>
- Shi, Zhongqi (史中琦), & Lü, Boning (吕伯宁). (2025). AI 在中文教学中的实证研究: 探索与展望——《AI 在中文教学中的应用》特刊序言 [Empirical research on AI in Chinese language teaching: Exploration and prospects—Preface to the special issue “AI in Chinese Language Teaching”]. *国际汉语教学学报* [International Journal of Chinese Language Teaching], 6(3), 1–5. <https://doi.org/10.46451/ijclt.250607>
- Singh, K., & Mueller, J. (2021). Taking a nuanced view of the role of teacher feedback in the elementary classroom. *Teachers and Teaching*, 27(1–4), 95–115. <https://doi.org/10.1080/13540602.2021.1933415>
- Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students’ writing. *Learning and Instruction*, 91, 101894. <https://doi.org/10.1016/j.learninstruc.2024.101894>
- Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education*, 6(3), 355–366. <https://doi.org/10.1177/209653112311684>
- Tay, H. Y. (2024). AI feedback: Moving beyond the hype to integrating it into a coherent feedback pedagogy. *Routledge Open Research*, 3, 26.
- The International Society for Chinese Language Teaching (世界汉语教学学会). (2022). 国际中文教师专业能力标准 [International Chinese Language Teachers’ Professional Competency Standards] (T/ISCLT 001-2022). Retrieved May 5, 2025, from <http://www.shihan.org.cn/index/build/detail.html?id=239>
- UNESCO. (2024, August 8). *AI competency framework for teachers*. UNESCO. <https://www.unesco.org/en/articles/ai-competency-framework-teachers>
- Vasu, K. A., Mei Fung, Y., Nimehchisalem, V., & Md Rashid, S. (2022). Self-regulated learning development in undergraduate ESL writing classrooms: Teacher feedback versus self-assessment. *RELC Journal*, 53(3), 612–626. <https://doi.org/10.1177/0033688220957782>

- Wang, L., Chen, X., Wang, C., Xu, L., Shadiey, R., & Li, Y. (2024). ChatGPT's capabilities in providing feedback on undergraduate students' argumentation: A case study. *Thinking Skills and Creativity*, 51, 101440. <https://doi.org/10.1016/j.tsc.2023.101440>
- Wang, Zhimin (王治敏), Wang, Yifan (王一帆), & Xu, Yue (徐悦). (2025). 国际中文教育智能技术应用及趋势研究 [Research on the application and trends of intelligent technologies in international Chinese education]. 华文教学与研究 [Teaching and Research of Chinese as a Foreign Language], (1), 9–21. <https://doi.org/10.16131/j.cnki.cn44-1669/g4.2025.01.006>
- Wen, Ting (闻亭), & Liu, Xiaohai (刘晓海). (2023). 国际中文智慧教育视域下的教学设计 [Teaching design of international Chinese language wisdom education]. 语言教学与研究 [Language Teaching and Linguistic Studies], (4), 24–33.
- Wiboolyasar, W., Wiboolyasar, K., Suwanwihok, K., Jinowat, N., & Muenjanchoey, R. (2024). Synergizing collaborative writing and AI feedback: An investigation into enhancing L2 writing proficiency in wiki-based environments. *Computers and Education: Artificial Intelligence*, 6, 100228. <https://doi.org/10.1016/j.caeai.2024.100228>
- Winstone, N. E., & Boud, D. (2022). The need to disentangle assessment and feedback in higher education. *Studies in Higher Education*, 47(3), 656–667. <https://doi.org/10.1080/03075079.2020.1779687>
- Wu, Lan (吴澜), Wang, Axi (王阿习), & Dong, Yan (董艳). (2024). 职前教师人机协同教学设计能力培养实证研究——基于自我生成教学理论视角 [An empirical research on the development of pre-service teachers' human-machine collaborative instructional design abilities—from the perspective of self-generated instruction theory]. 电化教育研究 [E-Educational Research], 12. <https://doi.org/10.13811/j.cnki.eer.2024.12.014>
- Xu, Z., Jain, S., & Kankanhalli, M. (2024). Hallucination is inevitable: An innate limitation of large language models [Unpublished manuscript]. arXiv.
- Yang, H., Zhang, Y., & Guo, J. (2025). Exploring the effectiveness of cooperative pre-service teacher and generative AI writing feedback on Chinese writing. *Behavioral Sciences*, 15(4), 518. <https://doi.org/10.3390/bs15040518>
- Yu, Xiaofei (余晓菲), Wang, Xiyue (王曦悦), & Qi, Feng (祁峰). (2025). 基于 ChatGPT 的国际中文句式语法教学实证探析 [An Empirical Analysis of International Chinese Sentence Grammar Teaching Based on ChatGPT]. 国际汉语教学学报 [International Journal of Chinese Language Teaching], 6(3), 6–24. <https://doi.org/10.46451/ijclt.250502>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 1–27.
- Zhang, W., & Ilisko, D. (2025). AI for enhancing English lesson design and pedagogy in Chinese middle schools [Paper presentation]. In ENVIRONMENT. TECHNOLOGY. RESOURCES. *Proceedings of the International Scientific and Practical Conference* (Vol. 3, pp. 397–403). <https://doi.org/10.17770/etr2025vol3.8562>
- Zhao, Yang (赵杨). (2025). 数智时代国际中文教师的认知与定位 [Cognition and Positioning of International Chinese Teachers in the Digital-Intelligence Era]. 世界汉语教学 [Chinese Teaching in the World], 39(3), 313–314. <https://doi.org/10.13724/j.cnki.ctiw.2025.03.010>
- Zheng, Yanqun (郑艳群), & Xu, Liqiang (许丽婧). (2023). 论数字化转型时期国际中文教师知识体系的重构 [On the reconstruction of the international Chinese-language teachers' knowledge system in the digital transformation period]. 云南师范大学学报 (对外汉语教学与研究版) [Journal of Yunnan Normal University (Teaching and Research on Chinese as a Foreign Language)], 21(3), 22–30. <https://doi.org/10.16802/j.cnki.ynsddw.2023.03.005>

Zheng, Yanqun (郑艳群). (2023). 在教育变革和技术变革中思考国际中文教育的前景 [Comments on the prospect of international Chinese language education in the context of educational and technological change]. 天津师范大学学报 (社会科学版) [Journal of Tianjin Normal University (Social Science Edition)], 2.

投稿: 2025 年 8 月 29 日; 接受: 2025 年 12 月 30 日; 出版: 2026 年 1 月 26 日

作者简介

孙冕, 西京学院外国语学院助教, 中国人民大学硕士, 研究领域为国际中文教育。

李悦菡, 西安培华学院文学院助教, 中国人民大学硕士, 研究领域为国际中文教育。

An Empirical Study on the Application of Generative Artificial Intelligence Feedback in Instructional Design for International Chinese Education: A Comparative Analysis of Multiple Feedback Modes

Mian Sun

Xijing University, China

Yuehan Li

Xi'an Peihua University, China

Abstract

In the context of digital intelligence, generative artificial intelligence has created new opportunities for the development of instructional design competence among teachers of International Chinese. While existing studies primarily focus on AI-assisted second language acquisition, empirical research on how teachers utilize AI feedback to enhance professional development remains limited. This study investigated pre-service International Chinese teachers by comparing the performance of a control group, a peer feedback group, an AI feedback group, and an AI+peer feedback group in lesson plan design tasks. The findings revealed that all groups showed significant improvement in the quality of their revised lesson plans, with the AI feedback group achieving the most substantial gains, followed by the AI+peer feedback group. AI feedback optimized lesson plan logic and structure through stepwise guidance, whereas peer feedback provided advantages in contextual adaptation and affective support. Human-AI collaboration demonstrated complementary effects in fostering reflection and innovation. The study suggests that the integration of AI in International Chinese education should emphasize human-AI collaboration rather than technological substitution. These findings offer empirical evidence and insights for enhancing teacher professional development and optimizing instructional design.

Keywords

International Chinese Education, generative artificial intelligence, feedback, instructional design

Mian Sun, Assistant Lecturer at the School of Foreign Languages, Xijing University. Research field: International Chinese Education.

Yuehan Li, Assistant Lecturer at the School of Humanities, Xi'an Peihua University. Research field: International Chinese Education.