

生成式人工智能辅助教案设计的主体性表达研究

宋沁怡

张全真 *

南京大学, 中国

摘要

生成式人工智能 (GAI) 在国际中文教育中日益普及, 但对使用者独立思考与个性化表达的影响尚缺实证检验。研究基于某高校“国际中文课堂教学法”课程, 对比 GAI 普及前后教案 (2022 年 33 份、2025 年 37 份), 结合三项量化指标与质化编码 (标准化表达 A、个性化表达 B) 考察其变化。量化结果显示, 普及后教案标准术语密度上升 63%, 词汇多样性下降 23%, 组内相似度上升 19%, 前两项达显著水平, 后者以效应量呈现 ($r=.32$)。质化编码显示, 个性化表达总量未见显著下降, 但标准化内容大量叠加使其比重明显压缩; B 类内部亦现结构性分化: 表面创意呈增长趋势, 个人声音显著减弱。研究为理解技术中介下准教师主体性表达的变迁提供依据。

关键词

生成式人工智能, 教师专业认同, 主体性表达, 趋同化, 国际中文教育

1 引言

在国际中文教育领域, 教案设计是课程体系中的核心训练环节, 承载着培养准教师教学理念与实践能力的双重功能。近年来, 多项研究发现生成式人工智能 (GAI) 辅助写作可能导致语言表达的同质化 (homogenization), 削弱使用者的个人声音 (Alvero et al., 2024), 创意空间的多样性会显著收窄, 思维逐渐趋同 (Anderson et al., 2024); 同时, 大脑活动也呈现“认知债务累积”的特征——对 AI 输出的被动接受逐渐替代了主动的认知加工 (Kosmyna et al., 2025)。张务农、汤洁 (2023) 从辩证唯物主义视角指出, 当人工智能替代人完成“非知到知”的认知建构过程时, 使用者的主体性反而被削弱。教案设计不仅是一项技术性操作, 更是准教师形成教学理念、发展专业身份的重要过程, 也关涉其专业认同、自我效能感与教学动机等心理层面的养成。如果教案设计这一关键训练环节因 GAI 的介入而趋于程式化, 准教师可能在获得“格式规范”的同时, 失去发展个人教学风格的重要锻炼机会。随着 GAI 普遍赋能国际中文教育的教案设计, 问题随之产生: 当教学者借助 GAI 设计教案时, 其独立思考与个性化表达是否也面临被“代劳”的风险?

当下研究多从“学习者”或“技术应用”角度出发, 关注 GAI 对学生学习效果的影响, 而从“准教师主体”角度考察 GAI 对教学设计过程影响的研究十分有限。已有讨论或集中于宏观层面的机遇与风险分析 (袁羲、吴应辉, 2023), 或聚焦于国际中文教师生成式人工智能素养评估

* 通讯作者。联系电邮: zhqzhen@nju.edu.cn

框架的构建 (范慧琴等, 2024), 但对 GAI 介入后准教师在具体教学设计文本中的主体性表达发生何种变化, 仍缺乏基于真实文本的实证考察。丁安琪、蒙小凤 (2025) 通过对 32 份 AI 生成教案案例的研究, 指出学生生产出的教案普遍存在“文化语境表层化、情境适应性不足及创新力抑制”等三重风险。

本研究将教案中体现设计者独立判断、个人风格与教学创见的内容概念化为“主体性表达”。选取 GAI 普及前 33 份教学设计(2022 年 5 月)与 GAI 普及后 37 份教学设计(2025 年 5 月)进行对比研究¹, 通过混合方法设计系统考察主体性表达的变化, 以期理解技术中介下准教师主体性表达的变迁提供实证依据, 并为教师教育中 GAI 的合理定位提供参考。

2 研究综述

教案设计在教师专业发展中的意义远超其技术功能 (Schön, 1983)。本研究将“主体性表达”界定为教案文本中体现设计者独立判断、个人风格与教学创见的内容, 即设计者基于个人专业知识、实践经验与教学判断所做出的独特设计选择与表达。殷杰 (2024) 和田勇、鲁璐 (2025) 分别从哲学层面与高等教育实践层面, 讨论了 GAI 介入对人类主体性的潜在影响, 本文将其归纳为认知依赖、表达趋同、反思弱化三个面向。这些理论关切为本研究提供了问题意识的源头: 如果 GAI 的介入导致上述三个面向的主体性表达减弱, 准教师专业成长中最需要主动建构的能力, 或许正在被技术悄悄代劳。“主体性”可以从几个相互关联的理论视角来理解, 但在实际教案中, 这些面向往往相互交织而非彼此孤立。

从学科教学知识 (Pedagogical Content Knowledge, PCK) 理论的视角来看 (Shulman, 1986), 教案设计首先涉及的是知识整合的个性化问题。Shulman 强调, 教师需要掌握的不仅是学科内容本身, 更是“使内容变成可教的知识”——即针对特定主题, 选取最有效的类比、例证和解释方式, 并理解不同背景学生可能带来的前概念与误解。这种知识本质上是个性化的: 面对同一教学内容, 不同教师会基于各自的学科理解和对学生的认知做出不同的设计选择。在教案中, 这一面向主要体现为对特定学生群体学习困难的针对性预判 (B3), 以及融入个人理解的创意活动设计 (B4)。需要指出的是, B4 在理论上属于知识整合的个性化体现, 但在 GAI 辅助语境下, 其设计主体究竟是学生本人还是技术工具, 需要结合具体文本特征加以甄别, 这也是本研究后续分析将重点关注的问题。

与此不同, Schön (1983) 所强调的则是经验反思如何在文本中被表达出来。Schön 区分了“行中反思”和“行后反思”两种机制, 强调专业实践者通过将默会知识转化为明确表达来形成专业判断力。Hatton & Smith (1995) 在此基础上进一步区分了不同层次的教师反思, 较高层次的反思涉及对自身教学信念和经验的深入审视, 而非仅停留于对教学过程的描述。在教案中, 这一面向主要体现为将真实经历或个人感悟融入教学设计 (B1)——也就是说, 如果同样的内容可以出现在任何一位学生的教案里, 那么它很可能不属于这一面向。

在更具体的语言层面, 主体性还体现在表达方式本身的个人化特征上。Wang et al. (2024) 的研究发现, 学生虽然认可 AI 生成文本的专业性, 但普遍反映其中缺乏个人声音的传达。Meyer (2023) 也观察到 AI 生成文本虽然形式规范, 却缺乏人类写作者特有的轶事、幽默和风格偏好。在教案文本中, 这一面向体现为使用非程式化的、带有个人腔调的表述方式 (B2), 包括口语化的表达、恰切的比喻, 以及那些读来会让人觉得“只有这个人才会这样说”的句子。

上述三个面向虽各有侧重, 却共同指向一个心理层面的事实: 教案并非中性的技术文档, 而是准教师专业身份的书写场域。教师专业认同并非静态属性, 而是在持续的意义建构中被塑

造与再塑造的过程 (Beijaard et al., 2004); 将自身经验、教学信念与决策理由外化为可被辨认的表达, 正是这一建构过程在教案写作中的具体形态; 当这一外化过程被工具部分接管, 影响的便不只是文本的个性, 还包括支撑专业成长的内在动机与自我效能感——李明超 (2023) 即指出, 人工智能的深度介入可能弱化教师的能动思考并引发自我怀疑。因此, 本研究所考察的主体性表达, 在文本层面表现为个性化内容的多寡, 在心理层面则关涉准教师职业认同与教学动机的建构空间; 这也是本研究将 B1 个人经验融入与 B2 口语化 / 比喻性语言这两类“个人声音”编码视为关键观测点的原因。

据此, 本文将上述面向操作化为如下编码体系 (见表 1)。

表 1. 主体性表达的理论面向与编码对应关系

理论面向	理论来源	在教案中的表现	对应编码
知识整合的个性化	Shulman (1986)	针对特定学生的教学预判; 融入个人理解的活动设计	B3、B4
经验反思的外化	Schön (1983); Hatton & Smith (1995)	个人经历融入; 教学信念的表达	B1
语言表达的个人化	Wang et al. (2024); Meyer (2023)	口语化表达; 比喻性语言; 个人风格	B2

标准术语密度、词汇多样性与组内文本相似度所测量的是文本的形式特征, 它们本身并不等同于主体性, 也无法直接反映设计者的独立判断与教学洞察。本研究之所以引入这三项指标, 是因为主体性表达的削弱在文本层面通常先以“趋同化”的形式显现: 当设计判断更多让渡于外部模板或工具输出时, 文本会趋向于更高的术语依赖、更窄的用词选择与更强的彼此相似。就此而言, 三项指标是趋同化这一前置条件的可观测代理, 而非主体性的直接测度。

具体到本研究的编码体系, 三项指标与质化类目之间存在可对应的方向性预期: 标准术语密度指向 A2 术语堆叠与 A3 公式化目标表述所刻画的程式化倾向; 词汇多样性的下降对应 B2 口语化 / 比喻性语言的收窄, 即 Meyer (2023) 所说的轶事、幽默与风格偏好的消失; 组内文本相似度的上升则指向“同样的内容可以出现在任何一位学生的教案里”这一去个人化状态, 与 B1 个人经验融入的流失相呼应。换言之, 量化部分预测的是质化编码应当呈现的方向; 两者构成方向上相互印证的证据链, 但因同源于教案文本, 尚不构成严格意义上的独立三角验证。

这一设计也有助于回应一种可能的替代解释: 术语密度的上升是否只是意味着专业成熟度的提高? 若确系专业成熟度提升所致, 则理应伴随 B3 针对性教学预判的同步增强——因为对特定学生学习困难的预判正是 PCK 发展的核心标志 (Shulman, 1986), 也是专业成长最难以速成的部分。同时, 专业成熟即便使教学判断在“做什么”上更趋共识, 通常也不会抹平“怎么说”层面的个人差异——成熟的专业写作恰以稳定的个人声音为标志 (Meyer, 2023); 而本研究测得的组内相似度基于 TF-IDF (词频 - 逆文档频率) 词面特征, 其上升指向的正是表达层的趋同, 难以用判断层的成熟加以解释。本研究的实测结果与这一解释的预期方向并不一致 (详见第 4、5 节), 这也构成了将量化趋同与主体性让渡相勾连的经验依据。

本研究主要聚焦两个问题:

- (1) GAI 普及前后, 模拟课堂教案设计有哪些不同表现特征? 在标准术语密度、词汇多样性和组内文本相似度上是否存在显著差异?
- (2) 在教案设计这一具体学科场域, 主体性的让渡究竟发生了吗, 又以何种方式发生? 两组教案在标准化表达与个性化表达的类型与分布上有何差异?

3 研究设计

3.1 研究样本

本研究样本来自某高校“国际中文课堂教学法”课程,该课程面向汉语国际教育专业三年级学生,要求学生独立完成教案设计作为期末作业。两个年度的课程由第二作者主讲,其作业要求与评分标准的跨年度一致性由第二作者作为任课教师予以确认。研究团队收集了两个学期的教案:2022年春季学期33份(GAI普及前,编号T01—T33)和2025年春季学期37份(GAI普及后,编号G01—G37),共计70份。两组学生来自同一专业、修读同一课程、由同一教师授课、使用相同教材,主要的系统性变量为GAI工具的可及程度(其他可能的干扰因素详见第5节)。根据课程作业提交时的GAI使用声明,2025年全部37份教案均声明使用了GAI辅助;学期末的问卷调查亦显示,26位应答者(占该组样本的70.3%)全部报告在课程学习中使用过GAI,无人选择“从未使用”,两组之间的系统性差异有据可查。

3.2 研究方法

本研究采用混合方法设计,以量化分析和质化编码为主要手段,辅以问卷数据进行补充验证。

3.2.1 量化分析

对70份教案进行文本提取后,先剔除字数少于2000字的过短样本(2022年、2025年各4份),以避免不完整文本对指标计算的干扰;下述三项指标均基于筛选后的样本计算:

(1)标准术语密度:统计每千字中标准教学术语的出现频率。术语清单包含“教学目标”“操练”“导入”“交际能力”等54个词条,涵盖教学目标、教学过程、教学方法等10个类别,基于《国际中文教学通用课程大纲》和主流教材用语编制。该指标衡量教案对专业术语的依赖程度。

(2)词汇多样性(TTR):采用类符/形符比(Type-Token Ratio)衡量用词丰富程度,即文本中不重复词语的数量与总词语数量之比。TTR值越高,表明用词越多样。本研究使用jieba中文分词工具对教案文本进行分词处理。由于TTR易受文本长度影响,上述过短样本的剔除亦有助于降低极端长度对该指标的干扰。

(3)组内文本相似度:基于TF-IDF(词频-逆文档频率)向量化方法将每份教案转化为数值向量,再计算同组教案两两之间的余弦相似度均值。该指标衡量同一组内教案之间的文本趋同程度——值越高,说明该组教案的遣词造句彼此越相似。需要说明的是,该指标的组间检验以两两配对的相似度值为分析单位;由于同一份教案参与了多个配对,配对值之间存在非独立性,其显著性水平可能被高估,故本研究对该指标主要参考其效应量与变化方向,而非p值本身。

由于样本量较小($n < 40$)且数据不满足正态分布假设,上述三项指标均使用Mann-Whitney U检验进行组间差异的统计检验。该方法属于非参数检验,不要求数据服从特定分布,适用于本研究的样本特征。

3.2.2 质化编码

使用 MAXQDA 24 软件对 70 份教案进行编码分析。基于前期研究经验和本研究的理论框架（见第 2 节表 1），建立了包含两个一级编码和八个二级编码的体系。

（1）A. 标准化表达：A1 模板化框架，指教案整体结构套用固定模板；A2 术语堆叠，指并列罗列多种教学法术语而无具体阐释；A3 公式化目标表述，指使用套话式的目标描述，内容可替换至任何课文；A4 程序化教学步骤，指按固定流水线排列教学环节。

（2）B. 个性化表达：B1 个人经验融入，指将真实经历或个人感悟融入教学设计；B2 口语化/比喻性语言，指使用非学术化的生动表达；B3 针对性教学预判，指基于个人理解对特定学生群体的学习困难做出具体分析；B4 创意活动设计，指设计具有原创性的教学活动。

为确保编码的可信性，研究采取了以下程序性策略。编码由第一作者完成：第一作者在通读全部 70 份教案的基础上进行首轮编码。在编码手册的形成过程中，第一作者与第二作者就类目界定与边界案例进行了多次讨论，并依据讨论意见对部分编码作出修订与删减，逐步形成界定明确的编码手册。首轮完成后，第一作者再依据定稿的编码手册对全部 70 份教案逐份进行第二轮通检，重点复核前后期编码标准的一致性与语义单元的切分。全部编码在 MAXQDA 24 中完成并存档，每处编码均与对应文本段落逐一关联，编码判断可完整回溯。编码体系的理论依据（见表 1）与最终编码结果经第二作者审定。编码遵循“按独立语义单元编码”的原则，即同一段落包含多个不同判断时分别编码，同一类型的重复表述仅计一次。两组编码频次的差异检验以每份教案的各类编码次数为分析单位，采用 Mann-Whitney U 检验。

3.2.3 问卷数据与研究伦理

2025 年学生在学期末完成了一份关于 GAI 使用情况的问卷调查，面向该学期全体选课学生发放，回收有效问卷 26 份（回收率 70.3%），内容涉及 GAI 工具的使用频率、使用场景和主观体验等。问卷结果作为辅助证据，用于对文本分析的发现进行补充验证。

就使用粒度而言，问卷结果呈现如下分布。在使用频率上，7.69% 的学生报告每天使用 GAI，30.77% 每周使用 3—5 次，38.46% 每周使用 1—2 次，23.08% 偶尔使用，无人选择“从未使用”。在使用场景上，选择“教案设计灵感启发”的比例最高（73.08%），其次为“教学反馈与改进建议”（53.85%）与“教学理论知识梳理”（50%），而“课件制作”（23.08%）等外围用途占比相对较低。这表明 GAI 在本课程中的介入集中于教学设计的构思与修改环节，而非仅用于语言知识查询或格式排版，亦即其作用点恰好落在本研究所关注的主体性表达最可能发生让渡的环节上。

在具体工具的选择上，学生于提交作业时须声明是否使用 GAI 及所使用的工具，2025 年全部 37 份教案均声明使用了 GAI 辅助，所涉工具包括 ChatGPT、DeepSeek、文心大模型、豆包、Claude、Grok 等多个主流通用大模型，部分学生同时使用 2 至 3 种。可见学生所接触的并非单一模型，而是多个通用大模型的组合，这一分布也提示本研究观察到的趋同化并非某一特定工具的产物。需要说明的是，本研究未收集提示词策略与人机交互轮次等更细粒度的过程性信息，相关讨论见第 5 节。

另外，本研究使用的教案为课程作业提交稿，学生在课程开始时已知悉其作品可能用于教学研究。为保护隐私，论文中的教案均以编号指代，不呈现学生姓名。

4 研究结果

4.1 量化指标

三项量化指标在两组间均呈现方向一致的变化（见表 2）。按上述标准筛选过短样本后，最终纳入分析的有效样本为 2022 年 29 份、2025 年 33 份。

表 2. 两组教案量化指标对比

指标	2022 年 (n=29)	2025 年 (n=33)	变化幅度	p 值
标准术语密度 (次/千字)	7.19	11.74	+63%	p=.003
TTR 词汇多样性	0.4610	0.3539	-23%	p<.001
组内文本相似度	0.1486	0.1765	+19%	p<.001

注：组内文本相似度因配对值非独立，其 *p* 值可能被高估，解释以效应量为准（见 3.2.1）。

标准术语密度方面，2025 年教案每千字出现 11.74 个标准术语，较 2022 年的 7.19 个增加了 63% (*p*=.003)，表明 2025 年教案更频繁地使用规范化教学表述。词汇多样性方面，2025 年教案的 TTR 值 (0.3539) 较 2022 年 (0.4610) 下降了 23% (*p*<.001)，说明用词重复度更高、词汇选择范围更窄。组内文本相似度方面，2025 年教案的组内相似度 (0.1765) 较 2022 年 (0.1486) 上升了 19% (*p*<.001，效应量 *r*=.32)，表明同组教案之间的趋同程度明显增大。

需要说明的是，这些量化差异并不直接等同于表达质量的优劣，更多反映的是表达方式的变化。尽管如此，将三项指标结合起来看，仍可以观察到一个较为一致的趋势：GAI 普及后的教案在语言使用上表现出更强的标准化和趋同倾向。

4.2 质化编码

表 3. 两组教案编码频次对比（括号内为每份教案均值）

编码类别	2022 年 (n=33)	2025 年 (n=37)	变化 (按均值) 与 p 值
A 类 (标准化表达) 总计	31 (0.94)	111 (3.00)	+219% (<i>p</i> <.001)
A1 模板化框架	13 (0.39)	44 (1.19)	+202% (<i>p</i> <.001)
A2 术语堆叠	5 (0.15)	15 (0.41)	+168% (<i>p</i> =.012)
A3 公式化目标表述	8 (0.24)	31 (0.84)	+246% (<i>p</i> <.001)
A4 程序化教学步骤	5 (0.15)	21 (0.57)	+275% (<i>p</i> <.001)
B 类 (个性化表达) 总计	113 (3.42)	111 (3.00)	-12% (<i>p</i> =.621)
B1 个人经验融入	14 (0.42)	5 (0.14)	-68% (<i>p</i> =.020)
B2 口语化 / 比喻性语言	15 (0.45)	4 (0.11)	-76% (<i>p</i> =.011)
B3 针对性教学预判	47 (1.42)	38 (1.03)	-28% (<i>p</i> =.209)
B4 创意活动设计	37 (1.12)	64 (1.73)	+54% (<i>p</i> =.110)
B:A 比值	3.6:1	1:1	—

注：*p* 值为 Mann-Whitney U 检验（双侧）结果，分析单位为每份教案的各类编码次数；B:A 比值为组内比值，不作显著性检验。

对 70 份教案的 MAXQDA 编码结果显示, 两组在标准化表达与个性化表达的分布上呈现出系统性差异 (见表 3)。由于两组样本量不相等 ($n_1=33$, $n_2=37$), 直接比较绝对频次会受到组规模的干扰, 因此表 3 在呈现绝对频次的同时, 以括号标注每份教案的平均编码次数 (频次/样本量), “变化”一栏以及后文的所有百分比均基于均值计算。需要说明的是, B:A 比值为组内比值, 不受样本量差异影响, 故两种口径下结果一致。以每份教案的各类编码次数为分析单位进行 Mann-Whitney U 检验: A 类总计的组间差异高度显著 ($U=161$, $p<.001$, 效应量 $r=.64$), 四个子码亦均达显著 (A1、A3、A4 均 $p<.001$, A2 $p=.012$); B 类中, B1 ($p=.020$) 与 B2 ($p=.011$) 的下降达到显著, 而 B4 的上升 ($p=.110$)、B3 的回落 ($p=.209$) 及 B 类总计的变化 ($p=.621$) 均未达显著。

编码结果并非简单的“个性化减少”, 而是呈现出三个相互关联的变化。

首先, 从整体分布来看, 标准化表达出现了显著增长, 个性化表达的比重被压缩。B 类个性化表达的均值有所下降 ($3.42 \rightarrow 3.00$, -12% , 组间差异未达显著), 而 A 类标准化表达从每份 0.94 次增至 3.00 次, 增幅达 219% 且高度显著。B 与 A 的比值从 3.6:1 骤降至 1:1。换言之, GAI 并未消灭个性化表达, 而是在其之上叠加了大量标准化内容, 使得个性化表达在教案整体中的比重大幅下降。其次, B 类内部结构发生重组。B4 创意活动设计的均值反而上升了 54% ($1.12 \rightarrow 1.73$, 未达显著), 而 B2 口语化/比喻性语言下降了 76% ($0.45 \rightarrow 0.11$)、B1 个人经验融入下降了 68% ($0.42 \rightarrow 0.14$), 两者的降幅均达统计显著。这一分化说明, GAI 擅长生成形式多样的教学活动, 但那些真正属于设计者个人声音的内容 (如用自己的话表达、融入个人经历、使用个性化语言) 正在明显流失。进一步来看, A 类四个子码的增幅全面且均衡, A1 模板化框架 (+202%)、A2 术语堆叠 (+168%)、A3 公式化目标表述 (+246%) 和 A4 程序化教学步骤 (+275%) 均呈现大幅增长, 表明标准化的渗透是系统性的, 覆盖了教案的各个组成部分。

然而, 以上趋势并非绝对。2025 年也有少数教案保持了鲜明的个人风格。如 G01 (B=9) 的教学反思深入具体, 作者反省自身教学习惯时写道:

讲解语法时, 我习惯性将知识点全部补充并给学生讲解 (如“在”的三种用法), 对于初级学生来说部分用法还没出现, 过早地教学反而会加重学生课堂负担。

这段反思同时体现了“经验反思的外化”(B1) 与“针对特定学生的教学预判”(B3)——既有基于自身教学经验的反省判断, 又包含对学习者的认知负荷的具体预估, 恰恰是 GAI 难以代劳的部分。G14 (B=8) 的活动设计则展现出基于具体文化比较的巧思: 教师要求学生“找到一种他们国家有而中国找不到的蔬菜或水果, 下一节课来讨论”, 并在课堂上引导学生比较“中国买菜使用‘斤’来算, 而大多数国家都是用‘公斤’来算”。该活动虽属 B4 创意活动设计, 但其设计思路嵌入了具体的跨文化对比视角, 需要学生主动从本国经验中寻找对比素材, 与 GAI 倾向生成的通用化活动框架形成明显对比。反之, 2022 年也有个别教案呈现明显的模板化特征。如 T02 (A=1, B=0) 的教学目标全部使用“通过 X 让学生掌握 Y”的固定句式——“通过课堂练习让掌握语言点和语法点”“通过讲解、问答、讨论和利用视频及图片让学生掌握文化点”“通过讨论提高学生的文化知识”——内容可替换至任何课文, 与本课具体议题几乎没有关联, 是 A3 “公式化目标表述”的典型形态。这说明趋同化是总体趋势而非个体必然。此外, 一份 2025 年教案 (G17) 的末尾保留了 AI 工具的对话结尾语——“希望能给你的教学带来启发! 如果需要具体某篇课文的详细教案模板, 也可以继续告诉我”——这一细节直接印证了 GAI 在教案设计中的深度介入。

问卷数据从学生视角为上述发现提供了补充佐证。在 2025 年学生的问卷中, 57.69% 的应答者反映 AI “缺乏个性化指导”, 这表明连使用者本人也意识到, GAI 提供的是通用化、模板化的辅助, 而非针对其个人教学理念的个性化引导。

5 讨论

从整体结果来看, GAI 对教案主体性表达的影响, 在总量层面更接近 “叠加” 而非 “削减”: B 类编码总量的组间差异未达统计显著 ($p=.621$), 即未能检出个人表达总量的显著下降; 考虑到本研究样本量有限、统计检验力不足, 这一结果宜理解为 “未观察到显著变化”, 而非 “确认不存在变化”。与之形成对照的是, A 类编码大幅且高度显著地增长 ($0.94 \rightarrow 3.00$, $+219\%$, $r=.64$), 使个性化内容在教案整体中的比重显著下降, B:A 比值由 3.6:1 降至 1:1。进一步分解可见, B 类占比由 78.5% 降至 50.0%, 其中约九成降幅可归因于 A 类的扩张——即便 B 类均值维持 2022 年水平, 其占比也会降至 53.3%。还需指出, B 类总量的稳定在构成上依赖 B4 的名义增长, 而 B1、B2 均显著下降; 鉴于 B4 的设计主体归属尚需甄别 (见第 2 节), 对 “个性化表达未受损” 的解读应保持克制。这一发现与 Kosmyna et al. (2025) 提出的 “认知债务” 概念相呼应: 学生并非不再思考, 而是其思考被嵌入了 GAI 预设的框架之中, 独立建构的空间在不知不觉中被压缩。孙冕、李悦菡 (2026) 在国际中文教育语境下也发现, AI 反馈虽能有效优化教案的逻辑与结构, 但在情境适配方面不及人际反馈, 与本研究揭示的 “GAI 擅长框架构建、弱于个性化适配” 的规律相互印证。这对准教师的专业发展可能产生深远影响——如果教案设计的练习功能被 GAI 的 “脚手架” 部分替代, 准教师形成个人教学风格和 PCK 的过程是否也会受到延缓, 值得持续关注。这一结果也回应了第 2 节提出的 “专业成熟度” 替代解释: 若 2025 年组的专业水平整体更高, 理应观察到 B3 针对性教学预判的增强与学情分析的进一步分化; 实测中 B3 不仅未见上升, 均值反而回落 28% (该回落未达显著, $p=.209$), 组内相似度则明显上升, 两项证据均与该解释的预期方向相悖。

B 类内部结构的重组是另一个值得深思的现象。B4 创意活动设计的名义增长 ($+54\%$) 与 B1 个人经验融入 (-68% , $p=.020$)、B2 口语化 / 比喻性语言 (-76% , $p=.011$) 的显著下降形成鲜明对比, 这种不对称性反映了 GAI 在教案设计中的一种能力边界: 它可以生成形式上多样的教学活动, 拓展选项空间, 却难以帮助教师表达教学决策背后的个人理由, 也无法参与教师身份认同的建构。Anderson et al. (2024) 关于大语言模型 (LLM) 使个体产出增多、群体层面却更趋同的发现, 以及 Meyer (2023) 对 AI 生成文本 “形式规范却缺乏个人经验注入” 的观察, 都与这一模式吻合。在教案设计中, B4 的增长正对应于形式层面的丰富, 而 B1 和 B2 的下降则是个人声音流失的直接体现。进一步看, B1 与反思性表达的收缩并不只是文本层面个人声音的流失, 也触及教师专业发展的心理维度。对准教师而言, 将个人经验与教学反思写入设计的过程, 是其形成教学信念、确立专业身份认同的重要环节; 已有研究指出, 人工智能的深度介入可能弱化教师的能动思考、引发自我怀疑等主体性困境 (李明超, 2023)。当个性化表达被标准化内容覆盖、教案形式日趋规整时, 准教师借由独立设计来建构职业认同、维系内在教学动机的空间也可能随之收窄——这也提示, 在国际中文教师培养中, 除数字素养外, 仍需关注其职业认同的养成与心理情感支持 (范慧琴等, 2024)。

当然, 标准化程度的提升并非全然消极。2025 年教案在结构完整性、术语使用的准确性方面确实优于 2022 年——后者有部分教案过于简略 (如 T11, 仅 1117 字) 或结构散乱。GAI 可能帮助了基础较弱的学生达到基本合格的门槛。但这样的 “基准提升”, 是否以 “上限压缩” 为代价? 2022 年的优秀教案 (如 T18, B=13; T19, B=9; T15, B=8) 展现了丰富的个人教学智慧——融入高考经历的导入设计、带有个人幽默感的活动命名、基于教学直觉的详细偏误

预判——这种水平的主体性表达在 2025 年的教案中虽然并未消失（如 G01, B=9），整体上却变得更为稀少。总之，标准化可以缩小差距，也可能降低上限，这一矛盾值得教师教育群体正视。这一现象也提示，仅依靠提升教师的 GAI 使用能力可能并不足够——正如范慧琴等（2024）所提出的国际中文教师生成式人工智能素养评估框架所揭示的，“工具选择与运用”等技能维度需要与“评价与反思”等认知维度协同发展。Lingard（2023）曾建议学术写作者区分写作中可以交由 AI 处理的劳动性环节与需要自己完成的思考性环节，这一区分对教案设计同样具有启发意义。

上述标准化趋势可置于更宏观的理论脉络中加以理解。Sourati et al.（2026）综述了大语言模型对人类表达与思想的同质化效应，将这一现象上溯至 Ritzer（1983）提出的“麦当劳化”（McDonaldization）理论——一种追求效率、可预测性与控制而抑制情境丰富性的进程。本研究观察到的教案标准化趋势可视为这一宏观进程在国际中文教育场域的微观映现：GAI 帮助实现的“效率提升”与“结构规范”恰对应“麦当劳化”框架中的“效率”与“可预测性”维度，而 B1、B2 等个人声音的流失则呼应“情境丰富性”的抑制。值得注意的是，这一问题在中文学界亦有学者警示：张明慧等（2023）指出，生成式人工智能遵循“现有表达—重组表达—表现”的路径，与人类“思想情感—表达—表现”的创作路径存在根本差异；以重组既有表达为基础的生成方式，天然倾向于产出趋同化的文本——这一机制差异恰为本研究观察到的 A 类标准化表达全面增长（A1+202%、A2+168%、A3+246%、A4+275%）提供了技术机理层面的解释。

本研究采用的是准实验设计而非严格的随机实验，两组样本分属不同学期的不同学生，因此在因果推断上需要保持审慎。下面就几个主要的替代解释逐一讨论。

首先是教师因素。两组教案的授课教师为同一人，据其确认，两个学期使用的课程大纲和教材相同，作业要求一致，均未对教案格式提出特定的模板规范，也未提供示范教案供学生参照。但研究团队无法完全排除教师在课堂讲授中对规范性的隐性强调可能随时间有所变化。其次是学生群体差异。两组学生均为汉语国际教育专业三年级学生，学科背景和培养阶段一致，但个体层面的学业水平差异无法通过现有数据加以控制，未来研究可考虑纳入前置课程成绩等基线指标以增强可比性。

另一个不能忽视的是时代效应：2022 年至 2025 年间，国际中文教育领域的教学话语本身可能趋于更加规范化——学生接触到的期刊论文、教学视频和网络资源都可能比三年前更为术语密集。这一宏观趋势可能在一定程度上解释标准术语密度的上升，但难以单独解释 B 类内部结构的显著重组（B4 增加而 B1、B2 下降），后者更可能与 GAI 工具的具体使用方式有关。这一宏观趋势与第 2 节讨论的“专业成熟度”解释在群体层面同源，上文 B3 与组内相似度的证据对二者构成共同约束。

还需指出的是，本研究对 GAI 使用模式的刻画仍限于使用频率、使用场景与工具类型层面。学生采用何种提示策略、经历多少轮人机交互、以及在多大程度上对 GAI 的输出进行改写，这些变量很可能对应着不同的文本建构过程：一次性采纳完整输出与经多轮追问后择要改写并自行重述，即便使用同一工具，其教案中主体性表达的留存程度也可能存在实质差异。本研究的问卷未收集上述过程性信息，因而无法在 2025 年组内部进一步区分使用模式与主体性表达之间的关系，这也意味着本文报告的组间差异是不同使用模式叠加后的平均效应。未来研究可通过对话日志留存、屏幕录制或使用后即时回溯访谈等方式采集人机交互过程数据，以检验不同提示策略与交互深度对主体性表达的差异化影响。

综上，本研究的核心发现——标准化表达的系统性叠加以及 B 类内部“表面创意增加、深层个性减少”的结构性分化——不能被严格归因于 GAI 这一单一变量，但 GAI 的普及仍是目前最具解释力的系统性差异。未来研究可在以下方向加以深化：引入纵向追踪设计，考察同一批学生在使用 GAI 前后教案风格的变化；结合课堂实录等多模态数据，考察教案中的主体性

表达是否转化为实际教学行为；对保持高度个性化的“例外”教案进行深度访谈，探究其抵抗趋同化的策略与动机。

6 结论

本研究通过混合方法设计，对比分析了 GAI 普及前后共 70 份教案中主体性表达的变化，主要发现如下。量化分析显示，2025 年教案的标准术语密度上升 63%，词汇多样性下降 23%，组内文本相似度上升 19%，其中前两项差异达到统计显著水平，组内文本相似度则以效应量呈现 ($r=.32$)，三者共同印证了趋同化趋势的存在。质化编码显示，个性化表达的总量未见显著下降 (B 类总量组间差异 $p=.621$)，但标准化内容的大量叠加使其在教案整体中的比重明显压缩；与此同时，B 类内部发生结构性重组：个人声音类编码 (B1、B2) 显著减少，创意活动类编码 (B4) 呈上升趋势但未达显著。问卷数据亦从学生视角为上述发现提供了佐证。

基于以上发现，本研究的实践启示主要指向三个层面。在教学设计的评价上，仅关注结构完整性和术语规范性是不够的，还应有意识地关注设计者的个性风格——个人经验的融入、针对性的预判、课后反思的深度，这些恰是技术最难替代的部分，也恰恰是最值得在课程中保护的内容。在 GAI 的使用方式上，应引导学生将其定位为搭建框架的“脚手架”而非思考的替代者——正如 Lingard (2023) 所建议的，区分写作中的劳动性环节与思考性环节，将前者交给 AI 处理，将后者留给自己，并始终将 AI 的输出视为需要注入个人声音的初稿。然而，使用方式的引导若缺乏制度层面的支撑，往往难以持续落实。制度干预的着力点应在评价标准本身——在评分维度中明确纳入主体性表达指标，并设计要求学生就具体教学决策阐明个人判断依据的反思任务，从而使“个人判断不可被技术替代”从倡导性原则转化为课程考核的实质要求。还需指出的是，保护教案中的主体性表达，保护的不只是文本的个性，也是职前教师在专业成长中确立教学认同、维系内在教学动机的空间；这与技术增强环境下对国际中文教师职业福祉的关注是一致的。

保护和增强人类能动性应是设计和采用任何 AI 技术的核心原则 (UNESCO, 2023)。

在 GAI 深度融入教育实践的今天，如何在技术赋能与主体性保障之间取得平衡，是国际中文教育乃至整个教师教育领域亟需面对的共同课题。

7 研究局限

本研究在分析方法层面尚存两点局限。其一，质化编码由第一作者一人完成，虽有编码手册约束与全样本二次通检，编码本身仍带有一定的诠释性，也无法报告编码者间信度系数，未来研究应引入独立编码者，结合量化信度指标进行交叉验证。其二，词汇多样性采用全文 TTR 测量，该指标随文本长度增加而系统性下降；两组教案的平均篇幅相近 (2022 年约 8700 字，2025 年约 8500 字)，长度效应引致组间系统性偏差的可能性较低，但全文 TTR 仍非长度稳健指标，未来宜采用 MATTR (Moving-Average Type-Token Ratio)、MTLD (Measure of Textual Lexical Diversity) 等指标复核。

致谢

本文系世界汉语教学学会全球中文教育主题学术活动计划项目“国际中文模拟课堂教案设计中 GAI 辅助的表现特征、检测对比与评价研究” (项目批准号: SH25Y26) 的阶段性研究成果。

注释

- 1 需要说明的是,本研究将两组样本分别称为“GAI 普及前”和“GAI 普及后”,而非以“未使用 GAI”和“使用 GAI”进行区分。2025 年学生使用 GAI 的方式和程度存在个体差异,具体的使用频率、使用场景与工具分布见第 3.2.3 节。

参考文献

- Alvero, A. J., Lee, J., Regla-Vargas, A., Kizilcec, R. F., Joachims, T., & Antonio, A. L. (2024). Large language models, social demography, and hegemony: Comparing authorship in human and synthetic text. *Journal of Big Data*, 11(1), 138. <https://doi.org/10.1186/s40537-024-00986-7>
- Anderson, B. R., Shah, J. H., & Kreminski, M. (2024). Homogenization effects of large language models on human creative ideation. In *Proceedings of the 16th Conference on Creativity & Cognition* (pp. 413–425). ACM. <https://doi.org/10.1145/3635636.3656204>
- Beijaard, D., Meijer, P. C., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education*, 20(2), 107–128. <https://doi.org/10.1016/j.tate.2003.07.001>
- Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, 11(1), 33–49. [https://doi.org/10.1016/0742-051X\(94\)00012-U](https://doi.org/10.1016/0742-051X(94)00012-U)
- Kosmyrna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. *arXiv*. <https://doi.org/10.48550/arXiv.2506.08872>
- Lingard, L. (2023). Writing with ChatGPT: An illustration of its capacity, limitations & implications for academic writers. *Perspectives on Medical Education*, 12(1), 261–270. <https://doi.org/10.5334/pme.1072>
- Meyer, C. A. (2023). Generative AI: The voice of the other. *Double Helix*, 11. <https://doi.org/10.37514/DBH-J.2023.11.1.05>
- Ritzer, G. (1983). The “McDonaldization” of society. *Journal of American Culture*, 6(1), 100–107. https://doi.org/10.1111/j.1542-734X.1983.0601_100.x
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- Sourati, Z., Ziabari, A. S., & Dehghani, M. (2026). The homogenizing effect of large language models on human expression and thought. *Trends in Cognitive Sciences*. Advance online publication. <https://doi.org/10.1016/j.tics.2026.01.003>
- UNESCO. (2023). *Guidance for generative AI in education and research*.
- Wang, C., Aguilar, S. J., Bankard, J. S., Bui, E., & Nye, B. (2024). Writing with AI: What college students learned from utilizing ChatGPT for a writing assignment. *Education Sciences*, 14(9), 976. <https://doi.org/10.3390/educsci14090976>
- 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>

- Fan, Huiqin (范慧琴), Wang, Lulu (王璐璐), Cao, Fang (曹芳), & Zhao, Runze (赵润泽). (2024). 生成式人工智能赋能国际中文教师数字素养提升研究: 评估框架、现状及建议 [On the empowering of generative artificial intelligence to enhance the digital competencies of international Chinese language teachers: Assessment framework, current status, and recommendations]. 天津师范大学学报 (社会科学版) [Journal of Tianjin Normal University (Social Sciences)], (3), 72–83.
- Li, Mingchao (李明超). (2023). 人工智能时代教师的技术隐忧与角色纾解——基于“主体性”教育哲学的视角 [Technology concerns and role relief for teachers in the age of artificial intelligence: From the perspective of educational philosophy of “subjectivity”]. 当代教育论坛 [Forum on Contemporary Education], (6), 105–112. <https://doi.org/10.13694/j.cnki.ddjylt.20230904.001>
- Sun, Mian (孙冕), & Li, Yuehan (李悦菡). (2026). 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 [A study on the application of generative AI feedback in international Chinese instructional design: An empirical comparison of multiple feedback modes]. 国际汉语教学学报 [International Journal of Chinese Language Teaching], 7(2), 12–30. <https://doi.org/10.46451/ijclt.260109>
- Tian, Yong (田勇), & Lu, Lu (鲁璐). (2025). 人工智能赋能高等教育: 主体性困境与路径重构 [Artificial intelligence empowering higher education: The dilemma of subjectivity and the reconstruction of pathways]. 成都理工大学学报 (社会科学版) [Journal of Chengdu University of Technology (Social Sciences)], 33(5), 96–111. <https://doi.org/10.3969/j.issn.1672-0539.2025.05.009>
- Yin, Jie (殷杰). (2024). 生成式人工智能的主体性问题 [Reconceptualizing agency with generative artificial intelligence]. 中国社会科学 [Social Sciences in China], (8), 124–145.
- Yuan, Xi (袁羲), & Wu, Yinghui (吴应辉). (2023). ChatGPT Plus 给国际中文教育带来的机遇、风险及应对策略 [The opportunities, risks, and coping strategies of ChatGPT Plus for international Chinese-language education]. 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), 53–62. <https://doi.org/10.16802/j.cnki.ynsddw.2023.03.011>
- Zhang, Minghui (张明慧), Yu, Lu (玉璐), & Yang, Jun (杨军). (2023). 禁用还是拥抱: 生成式人工智能与中国社科期刊原创把关 [Ban or embrace: Generative artificial intelligence and originality gatekeeping in Chinese social science journals]. 出版发行研究 [Publishing Research], (8), 49–55. <https://doi.org/10.19393/j.cnki.cn11-1537/g2.2023.08.031>
- Zhang, Wunong (张务农), & Tang, Jie (汤洁). (2023). 知与非知——再论人工智能应用对教学主体的影响 [Knowing and not-knowing: On the influence of artificial intelligence application on teaching subject]. 电化教育研究 [e-Education Research], 44(3), 36–43. <https://doi.org/10.13811/j.cnki.eer.2023.03.005>

投稿: 2026 年 7 月 23 日, 接受: 2026 年 9 月 15 日, 出版: 2026 年 10 月 16 日

作者简介

宋沁怡, 南京大学海外教育学院硕士研究生, 主要研究方向为国际中文教师专业认同与主体性、教师教育中的质性研究。

张全真, 南京大学海外教育学院副教授, 主要研究方向为汉语国际教育、汉语作为第二语言教学。

The Expression of Subjectivity in GAI-Assisted Lesson Plan Design

Qinyi Song

Quanzhen Zhang

Nanjing University, China

Abstract

Generative artificial intelligence (GAI) has been increasingly adopted in international Chinese language education, yet its impact on users' independent thinking and personalized expression has rarely been empirically examined. Drawing on lesson plans from a "Methodology of International Chinese Classroom Teaching" course at a Chinese university, this study compares two cohorts: 33 plans produced before the widespread availability of GAI (2022) and 37 produced after (2025). Combining quantitative indicators—standardized terminology density, lexical diversity, and intra-group textual similarity—with qualitative coding (Category A: standardized expression; Category B: personalized expression), the study examines changes in the expression of subjectivity in lesson plan design. Quantitatively, the post-GAI cohort shows a 63% increase in standardized terminology density, a 23% decrease in lexical diversity (TTR), and a 19% rise in intra-group similarity; the first two differences are statistically significant, while the change in intra-group similarity is reported as an effect size ($r = .32$). Qualitative coding further reveals that personalized expression has not been eliminated; rather, its proportional weight has been compressed, primarily by a substantial layer of superimposed standardized content. Within Category B, a structural realignment is observed: surface-level creativity shows an upward trend, while distinctive personal voice has markedly weakened. The study provides empirical evidence for understanding shifts in pre-service teachers' expression of subjectivity under technological mediation.

Keywords

Generative artificial intelligence, teacher professional identity, expression of subjectivity, homogenization, international Chinese language education

Qinyi Song, M.A. candidate at the Institute for International Students, Nanjing University. Her research interests include the professional identity and agency of Chinese language teachers, and qualitative research in teacher education.

Quanzhen Zhang, Associate Professor at the Institute for International Students, Nanjing University. Her research interests include international Chinese language education and teaching Chinese as a second language.