

基于多智能体辩论的多视角反思支架对职前教师教学设计能力的影响研究

Effects of Multi-Perspective Reflection Scaffolds Based on Multi-Agent Debate on Preservice Teachers' Instructional Design Competence

王依琳^{1,2}, 朱佳^{1,2*},

¹ 浙江全省智能教育技术与应用重点实验室

² 浙江师范大学教育学院

* jiazhu@zjnu.edu.cn

【摘要】 教学设计能力是教师专业发展的核心素养，而职前教师在成长过程中常面临理论与实践脱节的问题，且缺乏开展深度教学反思的有效支撑。针对这一现实困境，研究开发了基于多智能体辩论模式的多视角反思支架系统。该系统设置主持人智能体、学习者中心智能体和内容中心智能体，借助多理论视角的辩论式分析，助力职前教师审视并优化教学设计方案。研究拟采用三组准实验设计，从教学设计知识掌握、教学自我效能感、教案设计质量、反思行为深度等方面，评估该系统对职前教师的影响。目前，已完成系统的前期开发和小范围的测试，初步结果表明多视角辩论有助于教师加深对教学设计的认识和反思，提升其反思能力与实践水平。

【关键词】 多智能体；教学设计能力；反思支架；职前教师

Abstract: Instructional design competence is central to teacher professional development, yet pre-service teachers often face theory-practice gaps and insufficient reflection support. This study developed a multi-agent debate system comprising a moderator, learner-centered, and content-centered agent to help teachers examine instructional design from multiple theoretical perspectives. A three-group quasi-experimental design is planned to evaluate the system's effects on instructional design knowledge, self-efficacy, lesson plan quality and reflection depth. Initial system development and small-scale testing have been completed, with early results showing that multi-perspective debate effectively bridges theory and practice, deepening teachers' understanding and reflection on teaching design.

Keywords: multi-agent systems, instructional design competence, reflection scaffolds, pre-service teacher

基金项目：本文系浙江省尖兵领雁项目“跨领域多模态数据融合驱动下的多智能体智能决策服务研发及示范应用”（2026C02A1236）、国家自然科学基金面上项目“知识融合增强下基于多元教学事理图谱的生成式教学路径规划及应用研究”（62577050）的研究成果。

2. 前言

教学设计作为连接教学目标与教学实践的核心环节，直接决定教学质量与学生学习效果，其科学性与创新性是教师专业能力的集中体现（丁安琪等，2025）。然而，教学设计要求教师在复杂情境中权衡多种教学理念，平衡不同教学目标，做出专业的决策。研究表明，专家教师与新手教师的关键差异不仅在于教学技能的熟练度，更在于其能否深入理解不同教学选择背后的教育学原理，并根据具体情境做出适当的判断（Berliner, 2001）。这种专业判断能力需要教师持续进行批判性反思才能逐渐形成。然而，当前教师在教学设计中获得的支持显然不足，师范生在面对学校的具体状况时，往往产生理论与实践的隔阂（杨燕燕，2012）；在职教师则因繁重的教学压力和琐碎的管理工作而忽略了前提性反思（曹永国，2014），导致教学设计停留在表面或经验层面，难以实现真正的专业成长。

智能技术工具正逐步变革传统教学设计模式，这虽有效提升了教学效率，却也可能导致教师过度依赖智能生成内容，弱化其独立思考与批判性反思能力。为此，本研究开发了基于多智能体辩论模式的多视角反思支架，通过呈现内容结构与学习活动两个互补视角之间的对话与碰撞，激发教师的批判性思维，支持其深入开展前提性反思。目前，系统已完成构建与小范围试测，初步反馈显示其有效推动了理论与实践的结合，研究团队将据此完善系统设计并制定了完整的实证研究方案。

3. 相关工作

2.1. 教学设计中的人工智能

人工智能应作为教师的协作伙伴，帮助其在复杂情境中做出更科学的教学决策并推动自身的发展。然而，当前人工智能在教学设计中的应用大多聚焦于效率提升，如快速生成结构完整的教案（周洪宇等，2025）。若人工智能仅呈现输出结果，却未展现背后的推理过程，教师将难以理解其依托的教育学原理，甚至可能从主动的课程设计者沦为被动的系统使用者（Blömeke et al., 2022）。由此可见，人工智能在教学中的应用应更多支持教师的思考与判断，而非替代决策，从而帮助教师在实践中不断反思与成长。

2.2. 教育中的多智能体系统

已有研究表明，多个智能体围绕同一问题展开辩论，能显著提升推理质量和事实准确性（Du 等，2023）。这一辩论机制可推动智能体反复审视自身论点、识别逻辑漏洞，进而生成优于单一智能体的输出结果。然而，现有研究大多聚焦于提升智能系统自身能力，鲜少探讨将多智能体辩论机制应用于教师专业发展场景。从理论机制看，当教师同时接触来自内容结构与学习活动两个维度的分析时，两种视角间的差异形成认知冲突，打破原有认知平衡，驱动教师主动审视自身教学假设，进而激活元认知监控，推动其从描述性反思走向批判性反思。

4. 系统设计

本研究自主开发了多视角反思支架平台，包含三个智能体：主持人智能体负责协调对话全程，学习者中心智能体与内容中心智能体则分别代表课程论中两种经典立场，通过模拟多视角辩论引导教师深化对教学设计的专业理解。

3.1. 智能体角色设计

主持人智能体扮演协调者角色：分析教案并提炼核心议题，引导两个智能体展开平衡而有深度的对话，最终整合双方观点以启发教师反思。

学习者中心智能体围绕学生立场审视教案，关注学生已有经验与认知发展水平、自主探究空间的充分性以及学习情境的真实性，并据此提出促进学生深度参与的改进建议。

内容中心智能体忠实于学科知识体系的内在逻辑，审查教学内容的组织是否遵循学科顺序、知识结构是否严谨，并据此提出增强学科系统性的优化建议。

3.2. 辩论流程设计

系统采用多轮辩论机制，每轮辩论分为四个阶段：①问题聚焦：主持人提炼教案核心议题；②观点陈述：两个智能体分别从各自立场表态；③交互辩论：双方相互回应与反驳；④综合引导：主持人整合观点，促进教师全面反思。图1呈现了前两个辩论阶段的示例。

3.3. 提示词设计策略

系统采用结构化的提示词设计策略，分为五个层次：基础设定层定义智能体角色与风格、理论框架层嵌入核心理论原则、分析维度层明确关注的教学要素，对话策略层指导建设性辩论方式、情境适应层根据辩论阶段动态调整，确保各智能体在保持立场的同时生成有深度的对话内容。

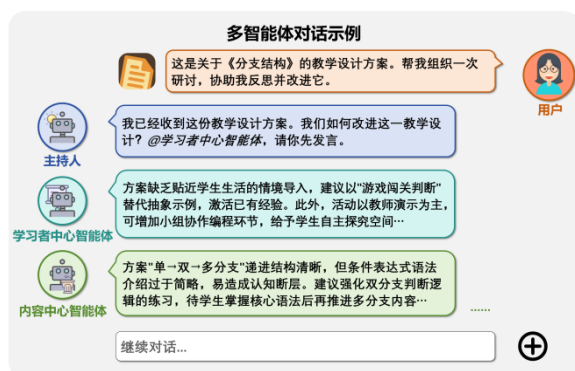


图 1 多智能体对话示例

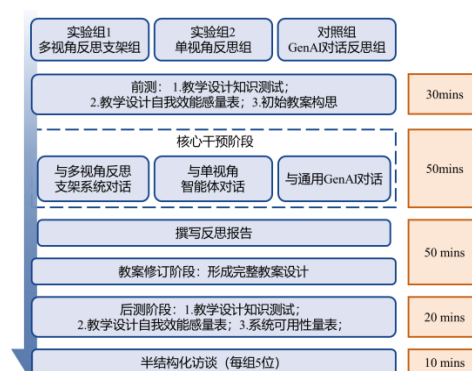


图 2 实验流程图

5. 研究方法与设计

基于前期系统开发与小规模试测（详见第5节），本研究计划开展准实验，系统评估多视角反思支架对职前教师教学设计能力的影响。研究核心假设为多视角辩论机制具有独立效应，优于单一视角对话。

4.1. 研究对象

本研究将以华东地区某所师范院校教育技术学专业和现代教育技术专业在读研究生作为研究对象，计划招募 75 名参与者，随机平均分配到多视角反思支架组、单视角反思组和 GenAI 对话反思组这三个组别，确保各组在性别和专业方向上的基本平衡。

4.2. 研究流程

研究将采用单次实验设计，在约 3 小时内完成前测、干预和后测的完整流程，如图 2 所示。前测阶段，所有参与者签署知情同意后，完成前测量表并就针对算法的基本概念与应用这一高中信息技术课程主题提交初始教案构思。前测结束后随机分组，各组接受为期 10 分钟的系统培训或任务说明。干预阶段，多视角反思支架组将教案提交至本研究开发的系统后，观看学习者中心智能体与内容中心智能体的 3-4 轮辩论；单视角反思组提交教案后，随机与仅持一种立场的智能体对话获取分析建议；通用 GenAI 组提交教案后与通用大语言模型对话获取改进建议。干预结束后，所有参与者有 50 分钟时间修订教案与撰写反思报告。后测阶段：所有参与者完成教学设计知识测试、教学设计自我效能感量表及系统可用性量表，并对随机抽取的部分参与者开展半结构化访谈。

4.3. 研究工具

本研究综合运用定量测量工具和定性数据收集方法。教学设计知识测试改编自 Koh 等（2013）的量表，包含 15 道选择题，涵盖学习理论基础、教学策略知识、教学设计原则和学习者分析四个维度。教学设计自我效能感量表改编自 Holden 等（2011）的工具，包含 12 个项目，采用 7 点 Likert 量表。教案质量评分采用研究团队开发的评分量表，从教学目标清晰度、教学活动适切性、教学策略多样性、学生中心性和评估设计五个维度评分，每个维度 0-4 分，总分 20 分，由两名评分者独立评分并计算 ICC。反思深度编码框架 (RDCS) 改编自 Hatton 等（1995）的模型，将反思分为描述性反思（1 分）、比较性反思（2 分）和批判性反思（3 分）三个层次。系统可用性量表 (SUS) 采用 Brooke（1996）的标准工具评估系统接受度。半结构化访谈探索参与者的反思体验和教学设计思维变化。

4.4. 数据分析计划

定量分析采用单因素协方差分析，以前测得分为协变量，比较三组在教学设计知识增益、自我效能感增益、教案质量和反思深度等因变量上的差异。事后多重比较采用 Bonferroni 校正，并计算效应量（偏 η^2 或 Cohen's d）。质性分析方面，反思日志使用 RDCS 框架进行演绎

编码,访谈数据采用主题分析法。最后采用三角互证策略整合定量和定性发现,增强研究结论的可信度。

6. 初步试测与下一步研究计划

本研究于2025年12月邀请3位职前教师进行试用,参与者均为教育技术学专业研二学生,已修读教学设计相关课程但无实际教学经验。试用反馈显示:(1)系统通过多视角辩论帮助参与者意识到教学设计的多种可能性,有效促进了理论与实践的连接;(2)部分内容较为学术化,辩论轮次较多时存在认知负荷;(3)建议增加更具操作性的改进方案。

后续研究将围绕以下三方面重点推进:首先,结合前期试用反馈优化系统设计,对智能体的语言表达进行调整,完善辩论流程设计,并为教师补充更具实操性的教学改进建议。其次,遵循本研究第四章的研究设计,招募75名职前教师分为三组开展准实验研究,评估系统对教师教学设计能力与教学反思水平的实际影响。最后,在验证系统有效性的基础上,进一步探索多智能体辩论机制在其他教师专业发展场景中的应用潜力,如课堂管理、学生评价等,进一步探讨如何通过认知冲突、观点融合等机制,推动教师开展深度专业反思,助力其在更广域的教学实践中实现专业能力的持续成长与提升。

参考文献

- 丁安琪,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.贵州师范大学学报(社会科学版),(05),42-52.
- 杨燕燕.(2012).论教师职前实践教学的取向转换.教育研究,33(05),84-89.
- 周洪宇,余江涛.(2025).迈向通用人工智能时代的教师工作:挑战与出路.电化教育研究,46(07),5-12.
- 曹永国.(2014).在做什么,抑或知道在做什么——教师的前提性反思的危机与重建.华东师范大学学报(教育科学版),32(01),41-49.
- Berliner, D. C. (2001). Learning about and learning from expert teachers. *International journal of educational research*, 35(5), 463-482.
- Blömeke, S., Jentsch, A., Ross, N., Kaiser, G., & König, J. (2022). Opening up the black box: Teacher competence, instructional quality, and students' learning progress. *Learning and Instruction*, 79, 101600.
- Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189(194), 4-7.
- Du, Y., Li, S., Torralba, A., Tenenbaum, J. B., & Mordatch, I. (2023). Improving factuality and reasoning in language models through multiagent debate. In *Forty-first International Conference on Machine Learning*.
- Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, 11(1), 33-49.
- Holden, H., & Rada, R. (2011). Understanding the influence of perceived usability and technology self-efficacy on teachers' technology acceptance. *Journal of research on technology in education*, 43(4), 343-367.
- Koh, J. H. L., Chai, C. S., & Tsai, C. C. (2013). Examining practicing teachers' perceptions of technological pedagogical content knowledge (TPACK) pathways: A structural equation modeling approach. *Instructional Science*, 41(4), 793-809.