

基于教学智能体 ArguePro 的英语议论文写作反馈：一项定性研究

Pedagogical Agent Feedback in EFL Argumentative Writing: A Qualitative Study of ArguePro

赵恒暄^{1*}, 王欣², 武永³

¹²³ 北京邮电大学人文学院

*hengxuan2002@163.com

【摘要】 本研究基于北京邮电大学 Trufar 平台开发了英语议论文写作教学智能体 ArguePro, 探讨学生对其反馈的评价及处理方式。通过对 6 名学生的刺激回忆与半结构化访谈, 发现学生普遍认可该智能体在审题、结构和论证方面的积极作用, 但部分学生仍指出其存在过于结构化、词汇超出学习者水平等局限。在反馈处理过程中, 学生表现出较高的反馈素养, 能够结合具体考试情境与个人表达需求对建议进行选择性和多轮追问。研究表明, 人工智能辅助写作应关注学习者差异与使用情景, 教师需引导学生有效使用智能体, 开发者亦应优化交互设计以提升人机协作效果。

【关键词】 教学智能体; 英语写作; 人机协作; 反馈素养

Abstract: The study investigates how students evaluate and process feedback provided by ArguePro, a pedagogical agent designed for English argumentative writing. Through stimulated recall and semi-structured interviews with six students, the findings show that learners generally recognize its positive role in topic analysis, structure, and argumentation, while also pointing out limitations such as over-structuring and vocabulary beyond the syllabus. Students demonstrate a high level of feedback literacy by selectively adopting suggestions and engaging in iterative questioning based on examination contexts and personal expression needs. The study suggests that AI-assisted writing should consider learner diversity and usage context. Teachers need to guide students in using the agent effectively, and developers should also optimize interaction design to enhance human-AI collaboration.

Keywords: pedagogical agent; EFL writing; feedback literacy; human-AI collaboration

1. 引言

随着大语言模型 (Large Language Models, LLMs) 的普及, 越来越多的学习者开始借助 AI 工具修改和润色英语作文。ChatGPT、Gemini、DeepSeek 等通用大语言模型, 虽能提供高效的反馈 (Guo & Wang, 2024; 任伟等人, 2024), 但在信任度和教学价值上, 学生认为其不如教师反馈 (Escalante et al., 2023)。这些通用大语言模型的反馈虽然对快速提升语言水平有帮助, 但在符合教学情境和学科要求方面依然有所欠缺 (Henderson et al., 2025), 在一定程度上限制了实际教学效果。

为回应上述问题, 本研究基于北京邮电大学 Trufar 平台设计并开发了一款面向大学英语六级 (CET-6) 议论文写作的教学智能体 ArguePro。尽管该教学智能体已完成开发, 但学习者如何理解并评价来自不同写作维度的反馈, 以及在修订过程中如何基于自身判断采纳或拒绝这些建议, 仍缺乏细致的实证描述。

基于上述研究背景，本研究旨在回答以下两个研究问题：

- (1) 学生如何评价写作教学智能体所提供的写作反馈？
- (2) 学生如何处理写作教学智能体所提供的写作反馈？

2.研究综述

2.1.教学智能体

教学智能体（Pedagogical Agent, PA）的现有理论与实证研究多建立在理工科学习任务之上，如数学解题（Kautzmann & Jaques, 2019）、科学实验模拟（Makransky et al., 2019）或游戏编程训练（Yalçin et al., 2022），但是针对英语议论文写作这类开放且强调逻辑的任务，相关实证研究仍存在缺口。尽管卢宇等（2024）验证了智能体能有效应用于语言类任务，但尚未深入探讨学习者在真实交互中的主观体验和使用感知，这为本研究提供了切入点。

2.2. 人机协同中的交互与决策

在人机协同写作中，学习者并非仅仅被动接受反馈，而是需要进行主动决策。Sutton (2012) 提出的“反馈素养”概念强调了学生理解、判断并利用信息改进学习的能力。在生成式人工智能（GenAI）应用于语言教学的情境下，反馈素养的含义得到进一步延伸：学生不仅需要理解 AI 提供的建议，还需具备对其准确性与适用性进行批判性判断（Cui et al., 2025）。因此，探究学生在面对智能体反馈时的判断与采纳策略，正是本研究试图填补的实证空白。

3.研究设计

3.1. 研究对象

本研究采用便利抽样法，选取了北京邮电大学 22 名选修思辨阅读与写作的大二学生参与实验，所有参与者均已通过大学英语四级考试，具备基础的英语写作水平。为了深入了解学习者的使用过程，在教学实验结束后邀请了其中 6 名学生（S1-S6）进行深度访谈。



图 1 ArguePro 使用界面

3.2. 研究工具

实验所用的教学智能体为 ArguePro，其核心设计理念是提供过程性引导，而非单一的纠错结果。ArguePro 在设计之初就融合教育学中的支架理论（Scaffolding Theory），分步骤提供审题（Topic Analysis）、统一性（Unity）、论证发展（Development）、连贯性（Coherence）以及

批判性思维（Critical Thinking）五个维度的反馈。智能体首先以设定的五个维度为学生提供反馈，每一维度完成后都可以选择进入自由对话模式，即在界面下方对话框进行追问（如图1所示）。

3.3. 教学实验设计

教学实验历时四周。实验开始前学生接受了关于议论文写作规范和 ArguePro 操作流程的培训，随后分别在第一周和第三周完成课堂限时作文，每次都需在课下使用 ArguePro 获取作文初稿的反馈，并根据建议进行修改，提交修改稿。实验过程中的所有人机互动记录都通过平台进行了完整保存。

3.4. 数据收集与分析

两轮写作修改完成后，研究者对6名访谈对象开展了访谈。首先，访谈使用学生第二次写作任务的初稿和修改稿进行刺激性回忆（Stimulated Recall），旨在还原学生在修改作文时真实的心理活动。接着研究者与访谈对象开展半结构化访谈，进一步了解其主观体验。所有访谈录音随后转录成文本，并采用主题分析法（Thematic Analysis）进行处理。

4. 研究发现

4.1. 学生对智能体反馈的评价

受访者对 ArguePro 的评价正反并存。在审题与结构方面上，学生普遍认可 ArguePro 提供的反馈，认为能有效梳理题目要求，提供新的思路（S3）和更完善的结构（S6）。但这种标准化的框架也引发了学生的质疑：S1 批评其“太结构化”，例如每个论点后都强制要求举例，这在限时和篇幅有限的考试中显得多余。

在论证内容方面，ArguePro 的拓展能力受到好评。S4 称赞其提供了正反视角的分析及新颖思路，S6 也认为补充的数据资料让论证更充分。不过，反馈的呈现方式有时显得较为割裂。S4 指出，智能体将论证、逻辑和连接词的问题分开修正，导致建议之间衔接不足，整体显得分散且臃肿，增加了阅读和修改负担。

在语言表达与交互体验上，学生的评价更为复杂。S4 认为智能体生成的内容通俗易懂，不像通用大模型那样堆砌生僻词。相比之下，S1 和 S3 则指出智能体有时会使用超出考生真实水平的词汇。此外，S2 和 S6 还提到了输入框显示不全、无法中途切换中英文以及上下文记忆丢失等问题，这些技术层面的问题影响了人机协作的流畅性。

4.2. 学生对智能体反馈的处理

在与智能体的交互过程中，学生并不会简单照搬修改建议，而是会结合实际情况进行取舍。比如在有字数限制的情况下，S2 在修改时会优先控制篇幅，把那些可能让文章变冗长的建议直接放弃，只在规定字数内对内容进行必要的调整。S3 也表现出类似的倾向，在模拟考场写作情境下，他更关注论证是否完整，对那些并非关键的论据补充往往选择性采纳。

学生对反馈的取舍还体现在对个人写作风格的坚持上。S5 在接受部分结构优化建议的同时，明确表示希望保留原有的句式和段落安排，拒绝把主体论证段拆分成多个段落。S1 则对词汇使用非常谨慎，她会主动删除 ArguePro 推荐的生僻词汇，并要求“提供更贴近我水平”的替代表达，以避免作文显得过于刻意修饰。

当智能体给出的反馈较为笼统时，学生通常会通过进一步提问来寻求更具体的修改思路。S2 在面对“需要多角度论证”这一建议时，会继续追问应从哪些角度展开，以及如何落笔。S4 在缺乏合适论据的情况下，也会主动向 ArguePro 求助，并在多轮互动中逐渐学会用具体事实来支撑观点。因此，学生会通过不断追问和互动，来弥补智能体反馈不够具体的地方。

5. 讨论与启示

研究发现, ArguePro 能快速地对学生的作文给出反馈, 在审题构思、谋篇布局、提供论据方面发挥了积极作用, 但其固定的输出方式在一定程度上也限制了学生的自主调控。这种标准化结构的评价之所以出现分歧, 可能与学生的写作习惯和实际需求有关: 部分学生乐于接纳其提供的支架, 而也有学生认为预设模板(如每个论点都要求举例)束缚了自己的表达。同时, 工具出现的上下文遗忘问题, 一方面源于大模型本身的记忆限制, 另一方面也受到平台界面和交互设计不够完善的影响。

在与智能体互动的过程中, 学生展现出了较高的反馈素养。他们不会简单接受系统建议, 而是依据具体考场情境与个人表达习惯进行筛选、调整, 并通过多轮追问不断细化修改思路。需要注意的是, 部分学生之所以排斥超出自身水平的词汇, 很大程度上是因为学生在使用前没有设置好与自身水平相匹配的参数, 导致大模型输出与实际能力不符。这表明, 人机协作的障碍有时并不完全来源于技术本身, 也与使用者对工具的熟悉程度有关。

未来英语教学中, 人工智能辅助写作应关注学习者差异与使用情景, 教师有必要从写作反馈提供者转变为人机协作的引导者。一方面, 应帮助学生判断和取舍教学智能体提供的建议, 使其更好的服务于具体写作任务; 另一方面, 也需要加强对工具使用方式的指导, 例如如何进行个性化设置以及如何提出更有效的问题。在技术层面, 开发者应进一步优化智能体的上下文处理能力和交互过程, 使反馈更加连贯、易于使用, 从而提升人机协作的整体体验。

参考文献

- 卢宇, 余京蕾 & 陈鹏鹤. (2024). 基于大模型的教学智能体构建与应用研究. 中国电化教育, (07), 99-108.
- 任伟, 刘远博 & 解月. (2024). 英语写作教学中 ChatGPT 与教师反馈的对比研究. 外语教学理论与实践, (04), 30-38+60.
- 徐振国, 刘志, 党同桐 & 孔玺. (2021). 教育智能体的发展历程、应用现状与未来展望. 电化教育研究, 42(11), 20-26+33.
- Cui, Y., Meng, Y., Qian, X., & Tang, L. (2025). Developing a college student feedback literacy scale for the GenAI context. *Assessment & Evaluation in Higher Education*, 1-14.
- Kautzmann, T. R., & Jaques, P. A. (2019). Effects of adaptive training on metacognitive knowledge monitoring ability in computer-based learning. *Computers & Education*, 129, 92-105.
- Makransky, G., Wismer, P., & Mayer, R. E. (2019). A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. *Journal of Computer Assisted Learning*, 35(3), 349-358.
- Sutton, P. (2012). Conceptualizing feedback literacy: Knowing, being, and acting. *Innovations in Education and Teaching International*, 49(1), 31-40.
- Yalçın, Ö. N., Lalle, S., & Conati, C. (2022, March). An intelligent pedagogical agent to foster computational thinking in open-ended game design activities. In *Proceedings of the 27th International Conference on Intelligent User Interfaces* (pp. 633-645).