International Society of the Learning Sciences

Learning by Teaching an LLM: A New Approach to Writing Instruction

Sydney Miller, University of Illinois Urbana–Champaign, srm16@illinois.edu
Nigel Bosch, University of Illinois Urbana–Champaign, pnb@illinois.edu

Abstract: This preliminary study investigates how different instructional modalities – i.e., video lecture versus learning by teaching (LbT) with a large language model (LLM) – influence students' understanding and application of rhetorical strategies in argumentative writing. Using a within-subjects design, 46 participants completed both a video condition and a LbT condition wherein they instructed an LLM-powered "student" on rhetorical strategies. Each session covered a distinct set of three strategies, with order and content counterbalanced. Learning was evaluated through pre-tests, immediate post-tests, delayed post-tests, and post-lesson essays. Participants used more rhetorical strategies in essays written after the LbT condition (2.70 vs. 2.43 on average) and made fewer errors on immediate (1 vs. 4) and delayed (7 vs. 15) post-tests. Participants who completed the video condition second showed a greater decline in strategy use (55% vs. 29%). These findings introduce a replicable experimental framework for studying LLMs as interactive tools for writing instruction.

Purpose

Large language models (LLMs) are increasingly used in educational contexts, raising concerns about how they shape student learning in writing. While some scholars argue that generative AI can support reflection and metacognition when used intentionally (Anders, 2023; De Matas, 2023; Graham, 2023; Luckin et al., 2016), others warn that these tools may encourage product-oriented writing and outsourcing of cognitive effort (Cotton et al., 2024; Vee, 2023). To date, most empirical work positions LLMs as feedback providers or drafting aids, emphasizing textual output rather than deeper cognitive engagement (Akiba & Garte, 2024; Hansen et al., 2025). Learning by teaching (LbT) offers an alternative framework, as explaining concepts and responding to questions can strengthen understanding and transfer (Duran, 2017; Fiorella & Mayer, 2013; Lachner et al., 2021; Roscoe & Chi, 2008), including in AI-based teachable agent systems such as *Betty's Brain* (Okita & Schwartz, 2013; Segedy et al., 2015). Recent advances in LLMs make it possible to extend LbT to open-ended domains like writing through conversational interaction (Malik et al., 2024; Salewski et al., 2023; White et al., 2023). This preliminary study compares a traditional lecture condition with an LbT condition in which students teach rhetorical strategies to an LLM-powered agent, examining post-test performance, strategy use, and essay outcomes.

Method

Forty-six undergraduate students completed two 30-minute learning sessions in a counterbalanced within-subject design: a video lecture condition and a learning-by-teaching (LbT) condition. After each session, students wrote an essay and completed a short multiple-choice quiz. An optional delayed post-test was administered at least two weeks later (n = 21).

Instruction focused on six rhetorical strategies central to persuasive writing instruction: ethos, pathos, logos, concession & rebuttal, analogies/metaphors, and rhetorical questions (Lunsford et al., 2018; Ramage & Bean, 2019). Three strategies were taught per condition, with strategy and topic assignments rotated using a Latin-square design to ensure full counterbalancing across conditions and to minimize potential differences in topic familiarity or strategy difficulty.

In the LbT condition, students taught rhetorical strategies to a simulated "student bot" powered by an LLM, which was prompt-engineered to behave as a novice learner (e.g., asking questions, making minor errors) to support explanation and correction (Debbané et al., 2023; Miller & Bosch, 2026). In the video condition, students watched pre-recorded lectures covering the same content.

Students produced two essays (one per condition; N = 92), which were manually coded for use of the strategies taught in the corresponding session. System logs (e.g., keystrokes, timestamps, typing speed, backspace frequency, time on task) and in-person administration were used to verify authentic engagement.

Results

We assessed learning through two modes: (1) post-lesson essays, in which students were prompted to apply recently learned rhetorical strategies, and (2) a series of knowledge tests: a pretest, two immediate posttests (after each learning condition), and a delayed posttest conducted at least two weeks after the initial session.

Students completed a pre-test assessing baseline knowledge of the six rhetorical strategies before exposure to any learning condition. Average accuracy was 84% (*range* = [50%, 100%]), indicating limited prior familiarity with the material. These scores provided a baseline for evaluating post-intervention learning and retention.

Following each learning condition, students completed a post-test on the three rhetorical strategies covered in that session. Immediate post-test accuracy was high across both groups – 97% following the video condition and 99% following the LbT condition – suggesting that both instructional modes successfully conveyed the target material in the short term (see Table 1).

Results revealed a clear divergence between conditions. Out of 22 total incorrect responses, 15 (68%) were from strategies learned in the video condition, whereas only 7 (32%) were from strategies learned in the LbT condition. This pattern indicates that information acquired through the teaching-based interaction with the LLM was retained more effectively over time.

For essays written after the video condition, students used an average of 2.43 relevant strategies, with 12 of 46 students (26%) successfully incorporating all three strategies they had just learned. For essays written after the LbT condition, students used an average of 2.70 relevant strategies, with 15 of 46 students (33%) demonstrating full transfer by incorporating all three.

Order effects

Students who began with the LbT condition followed by the video condition were more likely to show a decrease in strategy use between sessions, with 55% employing fewer rhetorical strategies in their second essay than in their first. In contrast, among those who completed the video condition first and the LbT condition second, only 29% demonstrated such a decline.

Overall, students demonstrated higher post-test accuracy, greater strategy use, and fewer delayed post-test errors after the LbT condition than after the video condition, with strategy use declining more often when the video condition followed the LbT condition.

Table 1
Summary of Learning Outcomes Across Learning Conditions

Measure	LbT	Video
Immediate Post-Test Accuracy	1 incorrect response	4 incorrect responses
Delayed Post-Test Errors	7 errors (32% of total errors)	15 errors (68% of total errors)
Average Strategies Used per Essay	2.72 per essay	2.47 per essay
Strategy Use Decline (%)	29% (video first → teaching second)	55% (teaching first → video second)

Discussion

Our findings suggest that teaching rhetorical strategies to an LLM can enhance both immediate learning and retention compared to video instruction. Students used more strategies in essays following the LbT condition (2.72) than the video condition (2.47), indicating stronger short-term transfer. Retention showed a similar pattern: 32% of delayed post-test errors were associated with LbT-taught strategies, compared to 68% from the video condition. Additionally, students who completed the video condition second were more likely to use fewer strategies in their second essay (55%) than those who completed the LbT condition second (29%), suggesting that LbT may better sustain engagement across sessions. Although preliminary, with additional data needed for statistical verification, these results align with prior LbT research and extend its application to AI-mediated writing instruction. Positioning students as instructors of an LLM highlights the potential of interactive, explanation-based learning and offers a replicable framework for integrating generative AI into writing pedagogy.

References

Akiba, D., & Garte, R. (2024). Leveraging AI tools in university writing instruction: Enhancing student success while upholding academic integrity. *Journal of Interactive Learning Research*, 467–480.

Anders, B. A. (2023). Is using ChatGPT cheating, plagiarism, both, neither, or forward thinking? *Patterns*, *4*(3).

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61*(2), 228–239.

De Matas, J. (2023). ChatGPT and the future of writing about writing. *Double Helix: A Journal of Critical Thinking and Writing*, *11*(1), 1–7. https://doi.org/10.37514/DBH-J.2023.11.1.09

Debbané, A., Lee, K. J., Tse, J., & Law, E. (2023). Learning by teaching: Key challenges and design implications. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW1), 1–34. https://doi.org/10.1145/3579501

Duran, D. (2017). Learning-by-teaching. Evidence and implications as a pedagogical mechanism. *Innovations in Education and Teaching International*, *54*(5), 476–484. https://doi.org/10.1080/14703297.2016.1156011

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, *38*(4), 281–288. https://doi.org/10.1016/j.cedpsych.2013.06.001

Graham, S. S. (2023). Post-process but not post-writing: Large language models and a future for composition pedagogy. *Composition Studies*, *v51*, 162–168.

Hansen, R., Prilop, C. N., Alsted Nielsen, T., Møller, K. L., Frøhlich Hougaard, R., & Büchert Lindberg, A. (2025). The effects of an AI feedback coach on students' peer feedback quality, composition, and feedback experience. *Tidsskriftet Læring Og Medier (LOM)*, *17*(31). https://doi.org/10.7146/lom.v17i31.148831

Lachner, A., Jacob, L., & Hoogerheide, V. (2021). Learning by writing explanations: Is explaining to a fictitious student more effective than self-explaining? *Learning and Instruction*, *74*, 101438. https://doi.org/10.1016/j.learninstruc.2020.101438

Luckin, R., Holmes, W., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

Lunsford, A. A., Ruszkiewicz, J. L., & Walters, K. (2018). *Everything's An Argument with Readings* (8th ed.). Bedford/St. Martin's.

Malik, A., Mayhew, S., Piech, C., & Bicknell, K. (2024). *From Tarzan to Tolkien: Controlling the Language Proficiency Level of LLMs for Content Generation* (arXiv:2406.03030). arXiv. https://doi.org/10.48550/arXiv.2406.03030

Miller, S., & Bosch, N. (2026, April). *Prompting for teachability: Designing novice personas in LLMs for learning by teaching contexts*. International Learning Analytics and Knowledge Conference (LAK).

Okita, S. Y., & Schwartz, D. L. (2013). Learning by teaching human pupils and teachable agents: The importance of recursive feedback. *Journal of the Learning Sciences*, *22*(3), 375–412. https://doi.org/10.1080/10508406.2013.807263

Ramage, J. D., & Bean, J. C. (2019). *Writing Arguments: A Rhetoric with Readings* (11th ed.). Pearson.

Roscoe, R. D., & Chi, M. T. H. (2008). Tutor learning: The role of explaining and responding to questions. *Instructional Science*, *36*(4), 321–350.

Salewski, L., Alaniz, S., Rio-Torto, I., Schulz, E., & Akata, Z. (2023). In-Context Impersonation Reveals Large Language Models' Strengths and Biases. *Advances in Neural Information Processing Systems*, 72044–72057.

Segedy, J. R., Kinnebrew, J. S., & Biswas, G. (2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, *2*(1), Article 1. https://doi.org/10.18608/jla.2015.21.3

Vee, A. (2023). Large language models write answers. *Composition Studies*.

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). *A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT* (arXiv:2302.11382). arXiv. https://doi.org/10.48550/arXiv.2302.11382