Teach2Learn: Assessing the Learning among CS1 Students on How Well They Can Teach LLM-Simulated Students

Aizen Baidya and Ayush Pandey
University of California, Merced

RESEARCH QUESTIONS

RQ1) Does learning-by-teaching lead to higher performance in CS1 courses?

RQ2) Do students feel more confident learning computer programming when they can teach a simulated student?

RQ3) Do LLM simulated students create realistic learning-by-teaching scenarios for novice programmers?

CONTEXT

- 66% of students at UC use AI tools [1]
- AI usage increases with student level: 69% of seniors reported usage versus 58% of first years [1]
- Students routinely using generative AI showed significantly lower cognitive engagement [5]

DESIGN GOALS

- Develop a new learning approach
- Works in harmony with generative AI
- Inspired by learning-by-teaching [2, 3, 4]

Key Design Principles
1. Facilitate knowledge articulation
2. Safe and judgment-free learning environment
3. Enable an active and scaffolded discussion
4. Ease of use of the interface

Table 1: Key design principles motivating Teach2Learn's design

Traditional	Learning by Teaching
1. "Banking model"	1. Teaching one's peers
2. Students → containers	2. Students → teachers
3. Teachers → depositors	3. Peers → LLM-simulated
4. Non-participatory	4. Depth of understanding
5. Non-dialogical	5. Resistance to offloading

Table 2: Comparing traditional vs. the learning-by-teaching model

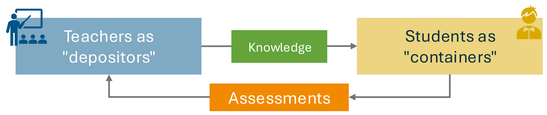

Figure 1: Flow of the traditional learning model

TEACH2LEARN DESIGN

Teach2Learn is a web-based platform. Sessions are open-ended and self-directed. Students are given full control over their teaching pace.

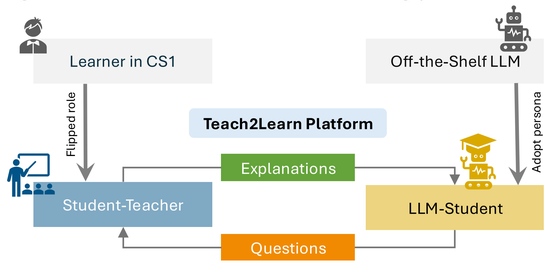

Figure 2: Flow of learning by teaching with Teach2Learn

Student-Teacher: Outlines topics of concern → Explains concepts step-by-step → Clarifies and addresses misconceptions.

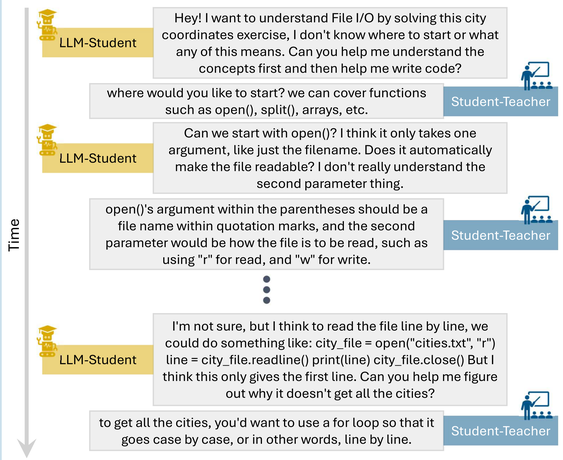

Figure 3: Sample conversation between student-teacher and LLM-student

PROCEDURE

Summer and Fall procedures were identical. File processing was the topic of learning.

Prior Exp	Summer	Fall
None	**66%**	27%
Self-Taught	6%	18%
High School	22%	**44%**
College	6%	11%

Table 3: Prior experience of the Summer (N = 32) and Fall (N = 55) cohorts.

Figure 4: Activities students engaged in before and following Teach2Learn.

RESULTS

Confidence and Realism

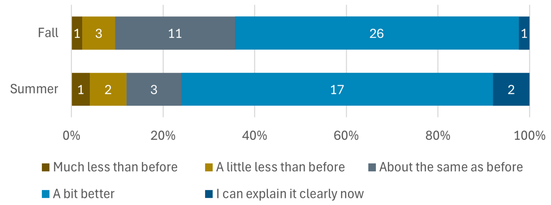

Figure 5: "After this activity, I understand file handling and processing..."

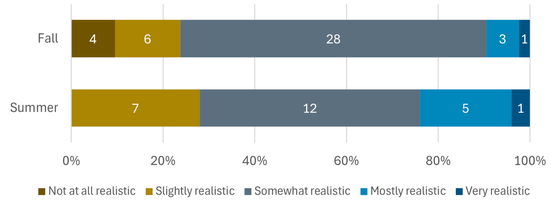

Figure 6: "How realistic did the simulated student feel?"

Student Performance and Learning

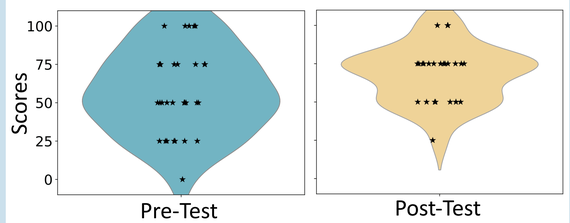

Figure 7: Scores for the Summer cohort.

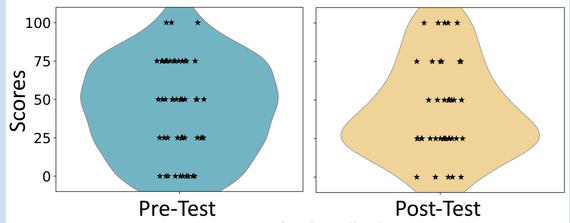

Figure 8: Scores for the Fall cohort.

We were unable to conclude that the activity caused a statistically significant difference in learning for either cohort.

Challenge to Thinking

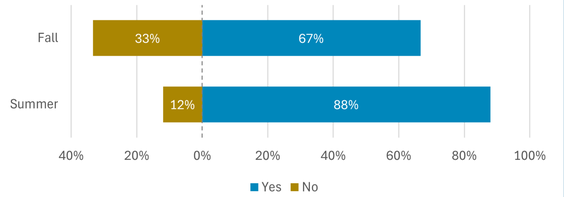

Figure 9: "Did the LLM-Student challenge your thinking?"

STUDENT FEEDBACK

+ "I liked how it was a simulated student since I have a hard time explaining things to people. So if I practice on this first and feel better, I can go to a real person after."
+ "Figuring out how to explain concepts helped me build my knowledge."
- "I didn't know whether or not the simulated student understood what I was trying to explain."
- "I wish I could see examples of what the student is doing, rather than a text box of what they're doing."

CONCLUSIONS

- 76% and 64% in the Summer and Fall cohorts reported programming confidence gains
- 88% and 67% respectively felt that the activity challenged their thinking
- [S] $\bar{x}$ = 57.03 → 68.00 | σ = 27.12 → 18.43
- [F] $\bar{x}$ = 44.09 → 43.45 | σ = 29.64 → 30.27
- Mixed impact on conceptual understanding and learning gains

FUTURE WORK

- Fine-tuning - Performing RLHF on annotated data from the pilots
- Student engagement - introduce activity as a regular, graded assignment
- Providing in-activity reward to students
- Generalizing to other topics and modules

REFERENCES

1. Institutional Research and Academic Planning, University of California. AI Use at UC. Oct. 2024.
2. Jin, Hyoungwook, et al. "Teach AI How to Code: Using Large Language Models as Teachable Agents for Programming Education". Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems.
3. Markel, Julia M., et al. "GPTeach: Interactive TA Training with GPT-based Students". Proceedings of the Tenth ACM Conference on Learning @ Scale.
4. Rogers, Kantwon, et al. "Playing Dumb to Get Smart: Creating and Evaluating an LLM-based Teachable Agent within University Computer Science Classes". Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems.
5. Rudrajit Choudhuri, Christopher Sanchez, Margaret Burnett, & Anita Sarma. (2026). GenAI's Impacts on Cognitive Engagement.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge support from CITRIS and the Banatao Institute at UC Merced.