



Volume 25 (2026), pp. 20-40  
*American Journal of STEM Education:  
Issues and Perspectives*  
eISSN 30.3-1190 | Print ISSN: 3069-0072  
<https://doi.org/10.32674/r3m20653>

## **Human vs. Robot: Exploring the Potentials and Pitfalls of AI-Led Mathematics Instructor Professional Development**

Alison S. Marzocchi

*California State University, Fullerton, United States*  
<https://orcid.org/0000-0002-2993-4646>

Nazgol Hadaegh

*California State University, Fullerton, United States*

---

### **ABSTRACT**

*Professional development (PD) is essential for the increasing number of mathematics instructors who aspire to improve their instruction. One PD format to support incremental instructional improvement is a reflection cycle. Using 31 authentic transcripts, we trained an artificial intelligence (AI)-powered chatbot to guide instructors in reflection cycles. We tested the chatbot by prompting it with reflection cycle segments and comparing its responses to those of student researchers. Data were qualitatively analyzed to find that student researcher and chatbot responses differed in three key ways: content, length, and tone. Findings have implications for widespread access to research-backed faculty PD. Overall, we explore the potential and pitfalls of AI-led instructor PD, with implications for AI's role in programs that support instructional improvement.*

**Keywords:** Chatbot, faculty development, faculty improvement, professional development, reflection cycles

© Marzocchi & Hadaegh, 2026. Published by Star Scholars Press.  
This article is distributed under the Creative Commons Attribution 4.0

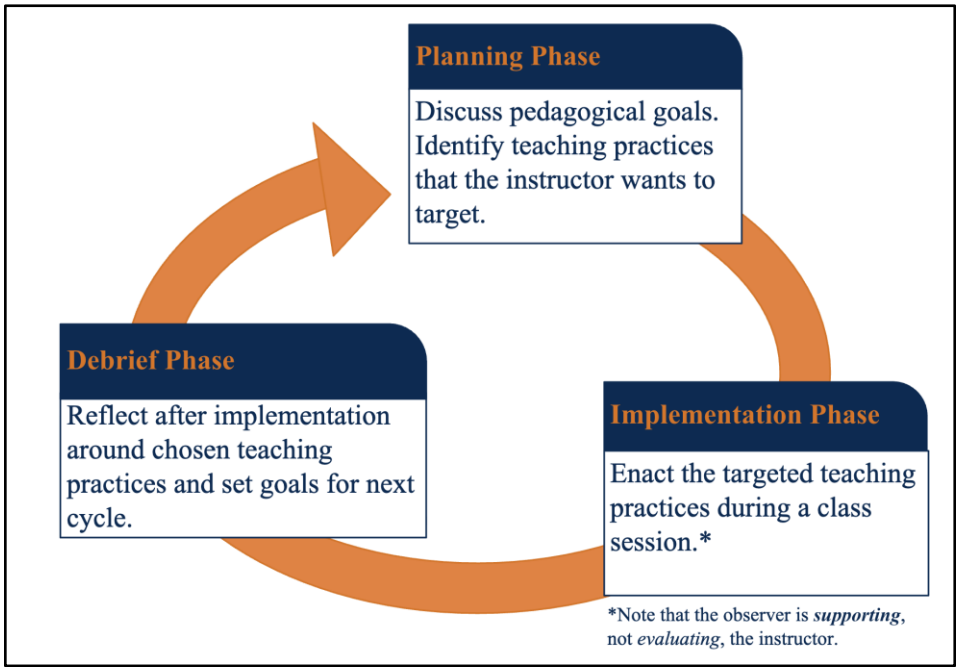
## INTRODUCTION

With research-backed guidance on effective teaching practices, many mathematics instructors desire to improve their instruction and need support to aid the transition (Bouwma-Gearhart, 2012). Effective professional development (PD) is a key mechanism for supporting instructional improvement (Brownell & Tanner, 2012). However, instructors have noted time as a barrier to participation in PD (Brownell & Tanner, 2012; Cook-Sather et al., 2024). Thus, our research turns to artificial intelligence (AI) tools to address the shortage of time for instructors. We already know that AI tools can provide instructors with customized support, assistance with assessment, and faster assessment (Ifenthaler et al., 2024). We seek to determine whether AI tools can also increase access to high-quality instructor PD while reducing the time constraints posed by traditional PD. We do so while carefully considering the impact of AI integration in educational settings, with an eye toward ethical implications, quality, and accuracy (Almasri, 2024).

Accordingly, we foreground the study's central purpose: to compare the quality of feedback generated by a trained chatbot with that of trained student researchers during instructor reflection cycles. To orient the reader to our analytic trajectory, we briefly note that our comparison centers on the three key dimensions of content, length, and tone. Together, these dimensions illuminate what may be gained or lost when reflection cycles shift from human to AI facilitation.

Supported by seven years of NSF funding (META: Mathematics Equity through Teaching Actively DUE # 2142122 and Student Engagement in Mathematics through an Institutional Network for Active Learning (SEMINAL) DUE # 1624610), our project has been providing California State University, Fullerton (CSUF) mathematics faculty with a robust and dynamic PD program grounded in a community of practice framework (Marzocchi & Soto, 2023; Marzocchi et al., 2024; Soto & Marzocchi, 2021; Stone-Johnstone et al., 2024). All PD activities are guided by the TEAM Tool (Bustos et al., under review), a student-developed, research-backed tool to support instructor goal setting, reflection, and gradual improvement. Instructors are invited to participate in a wide variety of PD activities, ranging from low-stakes (e.g., a casual brown-bag lunch to discuss a particular teaching strategy) to high-stakes (e.g., a year-long instructor learning community to redesign course projects). One such activity is a reflection cycle.

Reflection cycles consist of three phases: planning, implementation, and debrief (see Figure 1). Reflection cycle participation is voluntary, and reflection cycles are always non-evaluative; they are meant to support self-reflective instructional improvement. Historically, two trained student researchers have guided an instructor through a reflection cycle. In the planning phase, the student researchers guide the instructor to set goals for an upcoming lesson based on recommended teaching practices from the TEAM Tool. In the implementation phase, the student researchers attend the instructor’s class and take notes on how the instructor enacts the selected teaching practices. In the debrief phase, the student researchers guide the instructor in reflecting on the implementation of their chosen teaching practices and setting goals for the future. A subsequent reflection cycle can be scheduled to continue working towards the same goals or set new goals.



**Figure 1.**  
*Depiction of the Three Phases of a Reflection Cycle.*

Feedback from participants in the reflection cycle has been favorable, with many instructors volunteering for multiple cycles (Marzocchi et al., 2026). However, logistical challenges are a barrier to broad participation in reflection cycles. Challenges include the need for trained student researchers to lead the reflection cycles and coordinating the schedules of three individuals for three

separate meetings. Challenges will be exacerbated with the loss of funding to continue our project.

Consideration of these challenges led us to explore our current project. With four years of reflection cycle data, all recorded and transcribed, and a mountain of literature to support instructional improvement and effective instructor PD, is it possible to train a custom AI bot to lead instructor reflection cycles? If successful, any instructor can participate in as many reflection cycles as desired at any time. The implication for access is monumental. Yet how does the quality of an AI-led reflection cycle compare to the quality of an expert-human-led one? What is gained and lost in the transition from human to robot? This led us to our research question: When a university mathematics instructor sets a goal or seeks feedback on their instruction, how does the quality of a response from a trained student researcher compare to that of a trained chatbot?

By integrating a socio-technical lens with reflective practice, we conceptualize instructor professional development as an interplay between human expertise, technological mediation, and instructor agency. Thus, comparing student researcher-led and chatbot-led responses provides insight not merely into the *quality of output*, but into how different socio-technical configurations shape reflective conversations. This framework guides our analysis of content, length, and tone in feedback delivered by humans versus an AI chatbot.

## LITERATURE REVIEW

Our PD project, META, was motivated by the goal of supporting mathematics instructors in addressing longstanding opportunity gaps in postsecondary mathematics while increasing awareness of and capacity to address the opportunity gaps that are magnified for historically marginalized students (Engle & Tinto, 2008; Rasmussen & Ellis, 2013). Many undergraduate science, technology, engineering, and mathematics (STEM) instructors have not received substantial pedagogical training; therefore, institutions must support instructors in building teaching expertise over time (Laursen, 2019). META sought to meet this need by offering mathematics instructors a variety of low-stakes and high-stakes PD activities (Marzocchi & Soto, 2023; Soto & Marzocchi, 2021; Stone-Johnstone et al., 2024) under a community of practice framework (Marzocchi et al., 2024; Oliver et al., 2024).

As mentioned, one such PD activity is a reflection cycle. According to Reinholz (2016), “reflection is the act of processing an experience, action, or practice” (p. 444). In this spirit, META reflection cycles support mathematics faculty in considering ways to engage students in equity-minded, active learning within their classroom environment. Consistent with the literature, we begin each cycle by preparing the instructor for prospective reflection, which sets the tone and expectations for the experience, and conclude with a retrospective reflection,

which self-assesses, reconciles, and strategizes for future implementation (Reinholz, 2016). Faculty are invited to participate in multiple reflection cycles, allowing for the interplay between prospective and retrospective reflections (Reinholz, 2016). Over time, faculty build their skills in self-assessment (Reinholz, 2016). Like Reinholz, Stone-Johnstone, and Shah (2020), our reflection cycles engage faculty in goal-setting, implementation, and reflection to develop reflective practitioners over time. This improves the sustainability of the learning experience, as instructors can continue to reflect and improve outside of the formal structure of a reflection cycle.

Although we have seen numerous benefits of reflection cycles (Marzocchi et al., 2026), we acknowledge that our efforts were supported by NSF funding, which provided compensation for student researchers and instructors. We also recognize time as a noted barrier to instructor participation in PD (Brownell & Tanner, 2012; Cook-Sather et al., 2024). Thus, our study investigates the potential of widely available AI tools, particularly large language models, as a mechanism to provide PD to instructors at no monetary cost and on a flexible timeline. Research has shown that AI presents numerous advantages in the educational context, including enhancing instructional methods, streamlining assessment processes, supporting students in acquiring new skills, facilitating learning, and delivering personalized instruction (Almasri, 2024; Chichekian & Benteux, 2022; Ifenthaler et al., 2024; Slimi, 2023). Currently, most research on the benefits of AI in education has focused on the student experience. For example, Gupta and Chen (2022) investigated the impact of an intelligent, inclusive AI-powered tutor that can support *student* learning at scale. Labrado-Antolín (2026) examined the emotional tone and cognitive strategies of undergraduate students who used a chatbot. We wondered whether a chatbot tool could be designed to support *instructor* learning at scale.

Some research is showing the potential of AI tools to support instructors (Emre, 2025; Maanu et al., 2025). Emre (2025) documented the use of AI tools in higher education and reported that instructors used chatbots for pre-class preparation, such as syllabus design and lesson planning. In Maanu et al.'s (2025) study, pre-service mathematics teachers utilized AI-assisted tools in collaborative learning environments for mathematics education. They reported positive perceptions of these tools for teaching and learning. Use of the tools enhanced collaboration and reflective engagement among mathematics teachers, showing promise for AI tools in instructor PD. Ruiz and Yambao (2026) found that participation in professional development positively influenced instructors' readiness to incorporate AI tools. Coen and Cuddapah (2026) found that pre-service teachers were willing to use AI tools to generate ideas and increase efficiency but were concerned about accuracy and ethics, as well as using the tools for assessment.

Framing instructor professional development as a *socio-technical* practice helps clarify how AI-mediated reflection interacts with human judgment and the tools that structure reflective dialogue. In this view, AI does not replace reflection; rather, it shapes what information is surfaced, how feedback is scaffolded, and how instructors make sense of practice. These interactions are evident in emerging work showing that AI-focused PD can increase educators' confidence and intentions to use AI while also revealing persistent concerns about time, institutional policy, and data privacy (Riggs, 2025). In mathematics education specifically, AI-oriented PD has been shown to influence teachers' acceptance and planned use of generative AI, while highlighting both instructional affordances and risks (Shrestha & Yi, 2026). Conceptually, this aligns with socio-technical models of teacher learning that position reflection as emerging from interactions among people, artifacts, and practices; within these models, AI tools can scaffold noticing and goal-setting, while humans contribute contextual knowledge, affective attunement, and situated interpretation.

While AI tools offer promising support for instructors (Ifenthaler et al., 2024), it remains essential to address ethical considerations and ensure the quality and accuracy of these technologies (Almasri, 2024; Ifenthaler et al., 2024). Thus, this research compares authentic student researcher responses during instructional PD with those of a trained chatbot. We seek to uncover the potential, while recognizing the pitfalls, of AI-facilitated instructor PD.

## RESEARCH METHOD

From 2022 to 2024, trained student researchers facilitated 31 reflection cycles with voluntary instructor participants. Planning and debrief meetings were recorded and transcribed. We utilized the transcriptions of the 31 planning meetings and 31 debrief meetings, along with student researcher training materials and the TEAM Tool, to create and train a custom chatbot on the SchoolAI platform (schoolai.com).

The dataset for this study consisted of eighteen reflection cycles from nine individual instructors. Among these instructors, four completed one reflection cycle each, three completed two cycles, and two completed four cycles. The participant group was diverse, comprising six women and three men across various faculty ranks: one part-time lecturer, three full-time lecturers, two assistant professors, two associate professors, and one professor.

The subset of eighteen reflection cycles was selected by identifying instances within transcripts where the instructor asked the student researcher for feedback or stated a goal. All feedback-seeking or goal-stating instructor prompts were compiled alongside the student researcher's responses to those prompts. Instructor prompts were copied and pasted verbatim into our chatbot, and the chatbot's responses were compiled.

Next, we conducted a side-by-side comparison of the student researcher's response and the chatbot's response for each feedback-seeking or goal-stating instructor prompt. The TEAM Tool was used as the evaluation framework, with coding of instances of TEAM Tool teaching practices in student researcher responses and chatbot responses. The data were then analyzed for the affordances and limitations of each response. Once compiled, affordances and limitations were analyzed for themes. Three key differences between student researcher responses and chatbot responses emerged from the data: the type of feedback, the quantity of feedback, and the tone of response.

## RESULTS

After comparing student researcher responses to chatbot responses, after each was provided with the same prompt by an instructor, the three key differences were:

- (1) Content: the student researchers provided more practice-based suggestions, whereas the chatbot provided more literature-based strategies;
- (2) Length: in terms of quantity, the chatbot provided a greater number of suggestions and feedback, prompted for more detail, and asked more follow-up questions; and
- (3) Tone: when an instructor reported a challenge, the student researchers were more likely to offer empathetic support, whereas the chatbot consistently provided lists of possible solutions.

For each key difference, we provide three examples of instructor prompts in the subsections below, followed by the student researcher's response and the chatbot's response to the given prompt.

### **Key Difference 1: Content**

When an instructor prompted a student researcher, the student researcher was more likely to respond with a practice-based suggestion. The practice-based suggestions were often grounded in what they had seen in classrooms or in leading other instructors through reflection cycles. On the other hand, when provided with the same prompt, the chatbot was more likely to respond with a list of literature-based strategies. Literature-based strategies were drawn from the TEAM Tool. Table 1 presents three illustrative examples that demonstrate how the student researcher and chatbot responded differently to the same instructor prompt.

**Table 1***Key Difference 1: Content*

Instructor prompt	Student researcher response	Chatbot response
Instructor #1: “So, about the whiteboards, I always request better classrooms that are full of whiteboards, but everybody wants them.”	“What professors do is that they give these little plastic folder whiteboards...We've noticed that when it's a smaller classroom where there's not a lot of whiteboard space, those could have been used like that because it's only the size of a regular piece of paper.”	“Using individual whiteboards can be a fantastic way to encourage participation. Bring Small Personal Whiteboards, Use Paper Alternatives, Utilize digital platforms like Jamboard or Padlet, Collaborative Work, Incorporate 'My Favorite Mistake.'”
Instructor #2: “You brought up [digital whiteboards], but I've never quite figured out how to implement it. There's no way I could do that in this classroom because it would take up the entire whiteboard.”	“I have seen it implemented but all of the classrooms were big. They have a screen, and they have the whiteboard. [Otherwise,] it is difficult. I think it's just a nice resource for students to have so that they could collaborate more with each other.”	“Implementing tools like [specific digital whiteboards] can indeed be tricky in a classroom with limited whiteboard space. Here are some practical ideas you might consider to use these tools while keeping the setup manageable: Limited Display Area, Use of Smaller Screens, Collaborative Groups, Anonymous Contribution, Time Limits, Feedback Loop.”
Instructor #2: “I do think that [we need to find] some kind of activity that is relevant to the content that will maybe help [the students] see their understanding in what these objects mean. I'm starting to formulate an idea of how I can relate to the two [topics] to	“Yeah, I think again like I mentioned in the planning phase this is really hard to implement. It needs planning...I actually did research with [a professor where we] asked the students to use calculus 2 to find the rate of poverty.”	“You're definitely on the right track in recognizing the importance of connecting content to students' lives. Start Small: Use Student Interests. Real-World Examples, Collaborative Projects. Integrating real-world connections into your curriculum is a

---

help it feel more relevant to them.”

---

journey...What are your thoughts on these steps?”

---

These examples show how human facilitators rely on contextualized, experience-driven feedback, while the chatbot emphasizes structured, literature-aligned strategies.

## Key Difference 2: Length

When an instructor prompted a student researcher, the student researcher often provided a focused and targeted response or asked a specific follow-up question. In contrast, chatbot responses were longer, including more suggestions and feedback, prompting for more detail, and asking follow-up questions. Table 2 presents three illustrative examples that demonstrate how the student researcher and chatbot responded differently to the same instructor prompt.

**Table 2**

*Key Difference 2: Length*

---

Instructor prompt	Student researcher response	Chatbot response
Instructor #3:  “I'm pretty sure I want to focus on [teaching practices] ...number five, creates a classroom environment where students feel comfortable contributing. Number eight, uses mindful language by considering how word choice can impact student learning. And then the last one, number 15, establishes an environment where mistakes are embraced as a natural part of learning.”	“Okay, thank you and is there a reason why you chose these specific three teaching practices?”	“Excellent choices! Let's reflect on each practice:  Creating a comfortable environment for contributions: How do you plan to create this environment? What challenges do you anticipate in making students feel comfortable to contribute?...  Using mindful language: Can you share how you currently incorporate mindful language in your classroom? Are there any specific situations where you feel you could improve in this area?...”

---

---

Embracing mistakes as part of learning: How do you currently address mistakes in your class? What strategies might you adopt to reinforce that mistakes are part of the learning process?...

Once we've gone through these, let me know if there's any extra practice you want to focus on and if there's anything specific about your classroom that we should know."

---

Instructor #4:

"I don't know in that format [where one student reporter shares out for the group] how I can get some feedback from [the students], so I'll try that thumbs up thumbs down [strategy] and you guys can observe and let me know if it was effective, or if you have other suggestions."

"I believe these would be good implementations of that practice, definitely. Especially with the group work, and having a reporter, which is a great practice. It's also important to make sure that the whole group understands what they're all presenting. So I think that's a good angle."

"Regarding teaching practice 16 (assessing understanding), using the thumbs-up/ thumbs-down method is a solid start. Have you thought about any other methods to gauge understanding during the class, perhaps through quick checks or peer discussions? You could try incorporating quick exit tickets or digital polls for anonymous feedback. This can help gather more insights on student understanding. Does that sound feasible? Do you want to focus on any additional teaching practices, and is there anything specific about your classroom dynamics that we should know?"

---

Instructor #5:

"Yeah that's super funny. It's like even from a musical theater standpoint it's like you

"It's insightful that you're recognizing the impact of body language on student

---

“I feel like [I want to work on] body language. When you just basically turn your back on the class they’re like ‘ok I guess we’re doing something else here.’ Yeah that’s something I’ve picked up on, if you keep on staring, they’re gonna keep staring at me like ‘what’s next?’ I also learned a good tip from [a colleague]. I was observing him one time and he asked a question and he’s like, ‘what do we think about this?’ And instead of looking at the class, scanning the class waiting for a response, he looked at the board and [made a thinking face]. I was like, it’s such a small but clever idea...to take the eye pressure off of ‘anybody?’ and instead to be like ‘I’m going to look over here with you.’”

always think about [the audience] looking at the stage. Like, ‘okay what’s this character looking at, oh they’re looking at the board, okay what’s that all about?’”

engagement. Turning your back to the class can indeed send a signal that might disengage students.

The tip from [your colleague] about looking at the board while asking questions is a clever way to reduce pressure on students. This approach creates a more inclusive environment, allowing them to think without feeling spotlighted.

Here are some next steps you might consider:

**Set a Goal for Body Language:** Make a conscious effort to maintain eye contact with students and face them when asking questions. You could practice this in your next class and observe how it affects student interaction.

**Continue Using Think-Pair-Share:** As you mentioned, incorporating this strategy more regularly can help foster communication. Consider pairing this with your goal of enhancing body language by facing students when prompting discussion.

**Gather Student Feedback:** After implementing these strategies, ask for feedback from your students about

---

their comfort level in contributing. This can help you assess what's working and what might need further adjustment.

How do these ideas resonate with you? Are there any specific strategies you'd like to focus on or adjust moving forward?"

---

As shown, the chatbot expanded on prompts more extensively, whereas student researchers typically kept their feedback focused and succinct.

### **Key Difference 3: Tone**

When an instructor shared a challenge with a student researcher, the student researcher often responded with encouraging feedback, lived experience, or supportive affirmations. The response is often connected to human experiences. On the other hand, when presented with the same instructor report of a challenge, the chatbot typically responded with brief, positive encouragement, followed by several literature-based strategies. Overall, the student researcher responses, perhaps not surprisingly, felt more authentically human, whereas the chatbot responses felt more robotic. Table 3 presents three illustrative examples of how the student researcher and the chatbot responded differently to the same instructor prompt.

**Table 3**  
*Key Difference 3: Tone*

---

Instructor prompt	Student researcher response	Chatbot response
Instructor #5:  “My [class] is very quiet...so I want to find ways, even for the last four classes that I have and then thinking of next semester, I really want to start	“Sometimes it helps if the professor comes to me beforehand, before calling on me. This usually works after doing group work. If students are working in groups and then [instead of the professor] asking	“I understand that it can be challenging to encourage participation, especially with only a few classes left. For the practice of creating a comfortable environment, you might reflect on:

---

---

enacting these practices so that students do talk more.”

afterwards to share in class, if they ask someone beforehand, then it gives the student time to ask the professor for clarification if they're maybe struggling at that point or they can even ask their peers.”

What specific actions have you taken in the past to encourage student interaction?

Are there particular activities or formats that have worked better than others for engaging your students?

How do you currently respond to student contributions, and how might that impact their willingness to share?

Thinking about these questions can help you identify potential adjustments for your current class and plan for the next semester. What are your thoughts?”

---

Instructor #6:

“Some [student] groups kind of form naturally. So that's why I try to break them up. I'm like, okay, you go here you go here and try to get people to make a connection and [then they] can study outside of class time. You know, when you come in, you just have another face to say hello to and you just raise that comfort level. So the more I can mix [the groups] up. So anytime we do a little work like that, I'll say I'll purposely pair or trio intentionally so that

“So yeah, that's totally understandable. I definitely think that you did a wonderful job...not only did you meet the three teaching practices, but I thought I saw an overlap on the other teaching practices, like multiple approaches to support the success of every student.”

“It's great to hear that you're actively working to diversify student interactions by intentionally grouping them. This approach not only helps students get to know each other but also fosters a supportive learning environment.

Here are a few suggestions to balance structured grouping with the need for flexibility:

Mix and Match: Maintain a list of different group formations (e.g., pairs, trios, or larger groups) and

---

they get to meet different [peers]. But then there's other days where it's just like, okay, just work with the people around you for a few minutes. I can't spend too much time organizing groups in a short period. So I kind of need to fly by the seat of my pants depending on how things go.”

rotate students through them. This can help you quickly organize groups based on the day’s objectives.

**Quick Grouping Strategies:** Use random methods, like drawing names or numbers, to quickly form groups on the spot. This can save time while still ensuring a mix of students.

**Flexible Group Tasks:** Allow students to work with their immediate peers for simpler tasks, but encourage them to collaborate with different classmates on more complex assignments.”

---

Instructor #6:

“Some classes and some groups collaborate a little bit better and it's just sometimes something that happens at the beginning of the semester that makes things click...And then there's a group over here that didn't know each other before.”

“So for the [group] work, are you taking participation points for that? Because I heard that you'll get extra credit or you'll get more points or something like that.”

“It’s wonderful to hear that your students collaborated effectively during the quiz and communicated their mathematical understanding. This is crucial in a math class, where discourse can often be challenging.

Your reflection on wanting students to explain their work more often is a great goal. Here are some thoughts to consider moving forward:

**Time Management:** Since time is a constraint, consider breaking down lessons into smaller segments that allow for

---

brief student explanations or discussions without feeling rushed. You could allocate specific time slots for students to share their work.

**Accountability:** Your practice of having students submit pictures of their work is excellent for holding them accountable while promoting collaboration. You might also consider peer reviews of each other's work to reinforce learning.

**Collaborative Formats:** Exploring more varied collaborative formats, like jigsaw activities where each student becomes an 'expert' on a specific problem type, could help maintain engagement while allowing for shared learning."

---

These examples illustrate a distinction between human relational presence and the chatbot's more detached, strategy-focused tone.

## **Summary**

Beyond describing differences in content, length, and tone, these findings highlight how human-led and AI-led reflection cycles provide different versions of professional support. The student researchers' practice-based suggestions, focused follow-ups, and relationally attuned affirmations contain core features of human mentoring. In contrast, the chatbot's literature-based strategies, expansive prompting, and consistently neutral tone reflect a more structured, tool-mediated mode of facilitation aligned with socio-technical perspectives. In this version of the PD experience, the guidance available to instructors is shaped by the AI system's design, training data, and parameters. Thus, differences in content, length,

and tone reveal how distinct human and technological mediators differently shape reflective dialogue and the kinds of learning instructors can engage in.

## DISCUSSION & CONCLUSIONS

Our comparison of student researcher responses to chatbot responses, both provided with the same instructor prompt in a reflection cycle format, revealed key differences in content, length, and tone. Perhaps not surprisingly, the chatbot-led reflection cycles had a more robotic, theory-driven tone, while the student-researcher-led reflection cycles had a more humanistic, practice-oriented tone. AI-mediated reflection cycles provide structure, prompts, and theoretically grounded feedback but may lack the nuanced affective and experiential knowledge that human mentors bring. Prior studies of AI-focused professional development report gains in teachers' confidence and intentions to use AI alongside persistent concerns regarding time, training, and data privacy; research in teacher education likewise highlights emerging benefits and challenges for AI-mediated teacher learning (Coen & Cuddapah, 2026; Meylani, 2024; Riggs, 2025).

The differences stated above between human-led and chatbot-led reflection cycles affect the experience and feedback for instructors seeking PD. While the chatbot offered accurate, literature-aligned strategies, its responses lacked the contextual sensitivity, emotional resonance, and adaptive attunement that human facilitators used to validate instructors' experiences or calibrate feedback to the moment. From a human-AI interaction perspective, this distinction reflects how AI systems, even when well-trained, mediate reflection through patterned structures rather than relational cues, which may shape how instructors interpret the feedback and the extent to which they see it as actionable. From a different perspective, Labrado-Antolin (2026) saw differences in the quality of chatbot interactions based on the emotional tone and cognitive strategies employed by the human user. These differences suggest that AI-mediated reflection cycles may be most powerful when complemented by human relationships or when intentionally designed to support, rather than replace, the relational dimensions of professional learning. Yet, when considering a goal of providing PD opportunities to instructors, where adequate time and funding could be challenging, well-trained and sufficiently researched AI tools could provide a benefit. Given Ruiz and Yambao's (2026) finding that professional development participation positively influenced instructors' readiness to incorporate AI tools, AI chatbots must be considered as a way to deliver professional development to increasing numbers of instructors at scale.

Overall, the chatbot shows promise in supporting faculty instructional improvement in the absence of PD opportunities. However, its training depended on three years of data collection from 31 authentic reflection cycles led by humans, well-developed training materials for student researchers, and a comprehensive

literature-informed tool to support instructional improvement. Thus, high-quality chatbots require labor to train. Additional research is required to determine whether a basic, untrained chatbot could provide valuable PD for faculty.

Although our findings highlight differences in content, length, and tone, the effectiveness of AI-led reflection cycles is likely shaped by contextual factors. Instructors who value relational mentorship or who are navigating complex, emotionally nuanced challenges may find human facilitation more supportive. On the other hand, those seeking quick, structured guidance for early-stage goal setting may find AI responses efficient and sufficient. Institutional norms, available mentoring resources, and the specific aims of a reflection cycle may further influence whether AI-generated feedback feels actionable or meaningful. Recognizing these boundary conditions clarifies that AI-led reflection may complement human-guided professional learning, without a full replacement.

However, if a high-quality chatbot can be trained to guide instructors through reflection cycles, the implications for access will be monumental. One of the most significant barriers to instructor participation in PD is time (Brownell & Tanner, 2012; Cook-Sather et al., 2024). A chatbot will enable instructors to participate in a reflection cycle at their own convenience. This creates new opportunities for instructional improvement at a faster pace and on a larger scale. Our research also opens the possibility of broader dissemination of high-quality PD. If a team develops a high-quality PD model, perhaps on a limited funding timeline, they can then use the knowledge they gained to train an AI tool to replicate the PD program at other institutions. In this way, knowledge can be shared and built upon at a capacity not previously known, thanks to widespread access to AI. Of course, a critical, research-informed perspective is needed to ensure that we are not compromising quality or amplifying existing challenges and inequities. Thus, our direct comparison of AI-led PD with researcher-led PD is essential.

Future research will use the findings of this initial investigation to improve the reflection cycle chatbot before engaging instructors in reflection cycles using the chatbot. By having instructors utilize the chatbot for their own reflection cycles, we will be able to research the potential and pitfalls of AI-led reflection cycles compared to those led by student researchers.

## **AI TOOLS STATEMENT**

The authors used Grammarly to assist with language editing and formatting during manuscript preparation, and both Grammarly and Microsoft Copilot for proofreading during the revision process. ChatGPT EDU was used to locate missing article links. Microsoft Copilot was used to provide suggestions for incorporating some reviewer comments. All ideas, interpretations, and findings are the original work of the authors.

## ACKNOWLEDGEMENTS

Thank you to those who funded this research, including the National Science Foundation and the California State University, Fullerton Undergraduate Research Opportunity Center Engaging Graduate Students in Research and Scholarly and Creative Activities program. This material is based on work supported by the National Science Foundation under Grant # 2142122. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## REFERENCES

- Almasri, F. (2024). Exploring the impact of artificial intelligence in teaching and learning of science: A systematic review of empirical research. *Research in Science Education, 54*(5), 977–997. <https://doi.org/10.1007/s11165-024-10176-3>
- Bouwma-Gearhart, J. (2012). Research university STEM faculty members' motivation to engage in teaching professional development: Building the choir through an appeal to extrinsic motivation and ego. *Journal of Science Education Technology, 21*, 558–570. <https://doi.org/10.1007/s10956-011-9346-8>
- Brownell, S. E., & Tanner, K. D. (2012). Barriers to faculty pedagogical change: Lack of training, time, incentives, and... tensions with professional identity? *CBE—Life Sciences Education, 11*(4), 339–346. <https://doi.org/10.1187/cbe.12-09-0163>
- Bustos, F., Ceja, V., Filice, M., Luna Flores, O., Hofstetter, C., Suastegui, J., Stone-Johnstone, A., Marzocchi, A. S., Kurianski, K., & Soto, R. C. (under review). The TEAM Tool: Supporting Instructors to teach equity-minded active mathematics.
- Chichekian, T., & Benteux, B. (2022). The potential of learning with (and not from) artificial intelligence in education. *Frontiers in Artificial Intelligence, 5*. <https://doi.org/10.3389/frai.2022.903051>
- Coen, A., & Cuddapah, J. L. (2026). Artificial intelligence in action: How preservice teachers embark on the fast-paced journey of mindful AI use. *American Journal of STEM Education, 20*, 211–232. <https://doi.org/10.32674/keqjtt69>
- Cook-Sather, A., Braud, A., Kane, B., & Suresh, A. (2024). How student-faculty pedagogical partnerships counter adultism in higher education. *Taboo: The Journal of Culture and Education, 23*(1), 47–58. <https://files.eric.ed.gov/fulltext/EJ1445394.pdf>

- Emre, İ. E. (2025). Integrating generative AI tools in higher education. *American Journal of STEM Education: Issues and Perspectives*, 11(1), 51–68. <https://doi.org/10.32674/2gxxk8g16>
- Engle, J., & Tinto, V. (2008). *Moving beyond access: College success for low-income, first-generation students*. Pell Institute for the Study of Opportunity in Higher Education, Washington, DC. <https://files.eric.ed.gov/fulltext/ED504448.pdf>
- Gupta, S., & Chen, Y. (2022). Supporting inclusive learning using chatbots? A chatbot-led interview study. *Journal of Information Systems Education*, 33(1), 98–108. <https://jise.org/Volume33/n1/JISE2022v33n1pp98-108.html>
- Ifenthaler, D., Majumdar, R., Gorissen, P., Judge, M., Mishra, S., Raffaghelli, J., & Shimada, A. (2024). Artificial intelligence in education: Implications for policymakers, researchers, and practitioners. *Technology, Knowledge and Learning*, 29(4), 1693–1710. <https://doi.org/10.1007/s10758-024-09747-0>
- Labrado, M. (2026). Talking to machines: How communication style shapes student engagement with AI tutors. *American Journal of STEM Education*, 19, 37–58. <https://doi.org/10.32674/t2qznc90>
- Laursen, S. (2019). *Levers for change: An assessment of progress on changing STEM instruction*. American Association for the Advancement of Science. [https://www.aaas.org/sites/default/files/2019-07/levers-for-change-WEB100\\_2019.pdf](https://www.aaas.org/sites/default/files/2019-07/levers-for-change-WEB100_2019.pdf)
- Maanu, V., Boateng, F. O., & Larbi, E. (2025). AI-assisted instructions in collaborative learning in mathematics education: A qualitative approach. *American Journal of STEM Education: Issues and Perspectives*, 7, 11–36. <https://doi.org/10.32674/68zzkz60>
- Marzocchi, A. S., Stone-Johnstone, A., Kurianski, K., & Soto, R. C. (2026). Student-led reflection cycles to support faculty instructional improvement. *International Journal for Students as Partners*, 10(1), 201–218. <https://doi.org/10.15173/ijpsap.v10i1.6827>
- Marzocchi, A. S., & Soto, R. C. (2023). From the front lines of active learning: Lessons learned from those who are trying. *International Journal of Research in Undergraduate Mathematics Education*, 9(2), 524–555. <https://doi.org/10.1007/s40753-022-00176-x>
- Marzocchi, A. S., Stone-Johnstone, A., Kurianski, K., & Soto, R. C. (2024). Supporting mathematics instructors' transition to equity-minded active instruction using a community of practice framework. *Education Sciences*, 14(9), 1–13. <https://doi.org/10.3390/educsci14091001>
- Meylani, R. (2024). Artificial intelligence in mathematics teacher education: A systematic review and qualitative synthesis of contemporary research literature. *International Journal of Technology in Education Science*, 1(1), 63–91. <https://ijoftes.net/index.php/ijoftes/article/view/258>

- Oliver, J., Olkin, J., & Sisneros-Thiry, S. (2024). Leveraging a community of practice model to accelerate change. *PRIMUS*. Advance online publication. <https://doi.org/10.1080/10511970.2024.2354811>
- Rasmussen, C., & Ellis, J. (2013). Students who switch out of calculus and the reasons why they leave. In M. Martinez & A. Castro Superfine (Eds.), *Proceedings of the 35th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 457–464). <https://www.pmena.org/pmenaproceedings/PMENA%2035%202013%20Proceedings.pdf>
- Reinholz, D. L. (2016). Developing mathematical practices through reflection cycles. *Mathematics Education Research Journal*, 28, 441–455. <https://doi.org/10.1007/s13394-016-0175-1>
- Reinholz, D. L., Stone-Johnstone, A., & Shah, N. (2020). Walking the walk: Using classroom analytics to support instructors to address implicit bias in teaching. *International Journal for Academic Development*, 25(3), 259–272. <https://www.ingentaconnect.com/content/routledg/rija/2020/00000025/0000003/art00006>
- Riggs, V. (2025). Impact of AI literacy professional development on teaching: Evaluating changes in perceptions and classroom practices. *American Journal of STEM Education*, 7, 37–50. <https://doi.org/10.32674/rbajg012>
- Ruiz, C. C., & Yambao, F. M. P. (2026). AI literacy and professional development engagement as predictors of mathematics teachers' readiness for AI-enhanced instruction. *International Journal of Learning, Teaching and Educational Research*, 25(2), 655–675. <https://doi.org/10.26803/ijlter.25.2.29>
- Shrestha, S., & Yi, J. (2026). TPACK-based professional development for the AI era: fostering pre-service teachers' acceptance of generative AI in mathematics classrooms. *Journal of Mathematics Teacher Education*, 2026. <https://doi.org/10.1007/s10857-026-09747-0>
- Slimi, Z. (2023). The impact of artificial intelligence on higher education: An empirical study. *European Journal of Educational Sciences*, 10(1), 17–33. <https://files.eric.ed.gov/fulltext/EJ1384682.pdf>
- Soto, R. C. & Marzocchi, A. S. (2021). Learning about active learning while actively learning: Insights from faculty professional development. *PRIMUS (Problems, Resources, and Issues in Mathematics Undergraduate Studies)*, 31(3-5), 269–280. <https://doi.org/10.1080/10511970.2020.1746449>
- Stone-Johnstone, A., Kurianski, K., Soto, R. C., & Marzocchi, A. S. (2024). Mathematics equity through teaching actively: Centering equity in our active learning pursuits. *PRIMUS (Problems, Resources, and Issues in Mathematics Undergraduate Studies)*, 35(4-5), 479–492. <https://doi.org/10.1080/10511970.2024.2354813>

---

## Bios

**ALISON S. MARZOCCHI**, PhD, is a Professor of Mathematics at California State University, Fullerton, where she leads research and faculty development initiatives focused on equity-minded and active mathematics teaching. She is deeply committed to mentoring students and faculty, particularly those from historically excluded backgrounds. Email: [amarzocchi@fullerton.edu](mailto:amarzocchi@fullerton.edu)

**NAZGOL HADAEGH** is a graduate student in the Department of Mathematics at California State University, Fullerton, where she studies mathematics and conducts research focused on supporting instructors in improving their teaching. Her academic interests include mathematics education, the learning and teaching of mathematics, and motivating and supporting students through effective instruction. Email: [nhadaegh@csu.fullerton.edu](mailto:nhadaegh@csu.fullerton.edu)