Jiang, B. et al. (Eds.) (2025). Proceedings of the 33rd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

An Exploratory Study of a Generative AI Teacher Coach

Karishma SHANGHVI[a], Sridhar IYER[b]
[ab]*Center for Education Technology, Indian Institute of Technology, Bombay, India*
**24D1204@iit.ac.in*

Abstract: In low-resource educational environments, scalable teacher professional development remains a significant challenge. This exploratory study investigates the usability and perceived usefulness of a Generative AI (GenAI)–based teacher coaching platform designed to provide personalized, feedback-driven instructional support to teachers. The GenAI coach was built using established coaching and AI design models. The GenAI coach analysed audio from teacher talk during classroom teaching to provide personalised feedback and recommendations to teachers, including co-creating future goals and action plans. Teachers from a low-income K–12 school in Mumbai, India used the GenAI coach over multiple sessions. A mixed-methods evaluation included surveys and semi-structured interviews. Results indicated above-average usability and positive perceived value. Teachers highlighted strengths such as unbiased, personalised, evidence-based feedback, privacy and encouragement of self-reflection. Challenges included limited classroom capture, long processing times, and a need for lesson-specific feedback and pre-class planning support. The study offers early evidence that GenAI coaching tools can provide scalable and context-sensitive professional development in settings where access to human coaches is limited. Findings have implications for the future design of AI-enabled teacher development systems in India and other resource-constrained environments.

Keywords: Generative AI, Teacher Professional Development, Teacher Coaching, Scaling, Low-Resource Environments

1. Introduction

1.1 *Teacher Coaching as Professional Development*

Teacher coaching is a form of teacher professional development that enhances teachers' instructional practices and improves student learning outcomes. Extensive research has demonstrated that coaching, especially when individualised, sustained and context-dependent, supports teachers in reflecting on their practice, setting meaningful goals, and applying teaching strategies (Bush, 1984; Knight, 2007; Kraft et al., 2018; Reddy et al., 2021). However, in low-resource educational environments such as India, delivering high-quality instructional coaching at scale remains a substantial challenge (Knight, 2012). Constraints such as limited access to expert coaches, financial limitations, and logistical difficulties hinder the widespread adoption of traditional coaching models.

1.2 *Generative AI Based Teacher Coaching*

To address these challenges, this study explores the potential of leveraging generative artificial intelligence (GenAI) to provide personalised, one-on-one, scalable coaching to teachers. Specifically, a GenAI-based teacher coach was developed to offer teachers actionable feedback on their classroom teaching, recommend specific instructional practices and guide them to create future goals and action plans aimed at instructional improvement. The GenAI coach seeks to democratize access to coaching, particularly benefiting teachers

and schools in low-income contexts where conventional coaching is difficult to implement effectively.

To use the GenAI coach, the teacher registers on the mobile app and engages in an onboarding conversation. For every coaching cycle, the teacher records her voice while teaching in the classroom and uploads the recording to the app, along with a brief description of the class. The GenAI coach transcribes and analyses the recording based on a pre-specified rubric of teaching quality. Feedback is provided to the teacher with specific quotations from the transcript as evidence. The teacher and coach then engage in a collaborative conversation and co-create an action plan for improvement in the next class.

The current exploratory study was conducted to evaluate the usability and usefulness of the GenAI coach, by engaging three teachers from a private K-12 school in Mumbai that serves low-income students to gauge teacher perceptions regarding the value of the app as a professional development tool. The study aimed to uncover both strengths and limitations of the GenAI coach from the perspective of actual users within a real-world educational setting.

1.3 *Research Questions*

The study is guided by the following research questions:

1. How usable is the GenAI coach according to teacher users?
2. How useful do teachers perceive the GenAI coach for their professional development?
3. What improvements do teachers suggest to enhance the effectiveness of the GenAI coach?

2. Generative AI in Teacher Coaching

Teacher coaching is well documented as an impactful method for enhancing teacher practice and student achievement. Coaching models for teachers have evolved over the years, encompassing peer coaching (Joyce & Showers 1980), cognitive coaching (Costa & Garmston, 1994), instructional coaching (Knight 2007) and student centered coaching (Sweeney, 2011) amongst others. Human to human coaching generally involves cycles of human experts working with teachers, observing classrooms, providing feedback, and supporting reflective practice.

Since 2022, the emergence of generative AI (GenAI) offers new possibilities in coaching, with large language models (LLMs) capable of analyzing complex classroom data and generating nuanced feedback. Generative AI can result in more complex, open-ended coaching, providing personalised feedback and adaptive dialogue for teacher professional development.

3. Theoretical Basis for the Generative AI Coach Design

While designing GenAI coach application, the researchers identified models for crafting the AI training prompt (Kamphorst, 2017; Terblanche, 2020), the teacher coaching process (Knight, 2007) and the teacher competency rubric (Danielson, 2011). See Figure 1 alongside for theoretical model use across the GenAI coaching process.

3.1 *AI Training Prompt*

Kamphorst (2017) lists eight essential features of effective e-coaching systems, social ability, credibility, context awareness, personalisation and learning, data integration, proactivity, behaviour change modelling and planning support.

The Designing AI Coach (DAIC) framework (Terblanche, 2020) outlines a structured and ethical approach to developing AI coaching chatbots for organisational settings. The model integrates both coaching science and AI design principles. The model requires specific design considerations for the AI coach such as defining the organisational context, creating boundaries of ethical conduct, designing for the relationship aspects of the AI coach and

coachee, choosing the theoretical coaching model, and specifying the required output of the coaching exercise (Terblanche, 2024).

The prompt design of the GenAI coach platform used in this study was based on both the Kamphorst criteria and the DAIC framework.

3.2 *Teacher Competency Rubric*

Multiple teacher competency rubrics were explored to provide a rich rubric for the GenAI coach to analyse teacher talk in the classroom. The Framework for Teaching (FFT) model (Danielson, 2011) was chosen because of its wide implementation and research validity. Under the FFT, four domains of teacher competency are described with elements and rubrics, which are 1) planning and preparation, 2) the classroom environment, 3) instruction, and 4) professional responsibilities. For the purpose of the instructional coach design, since only classrooms are observed, Domains 2 and 3 were chosen as relevant.

3.3 *Teacher Coaching Process*

While there exist many teacher coaching processes (Costa & Garmston, 1994; Sweeney 2011), the theoretical coaching model chosen for the GenAI coach is popular and research validated Impact Cycle from the Instructional Coaching model (Knight, 2007). The cycle is a structured coaching model for K-12 teachers that guides instructional coaches and teachers through a three-phase process: Identify, Learn, and Improve. In the *Identify* phase, coaches observe teachers and identify strengths and areas for improvement. The *Learn* phase involves a collaborative conversation between the coach and teacher to discuss classroom observations, share feedback and explore teaching strategies. Finally, in the *Improve* phase, the coach and teacher collaboratively create an action plan and teacher applies the strategies in the classroom. The interaction process between the GenAI coach and teacher followed the Impact Cycle (see Figure 1).

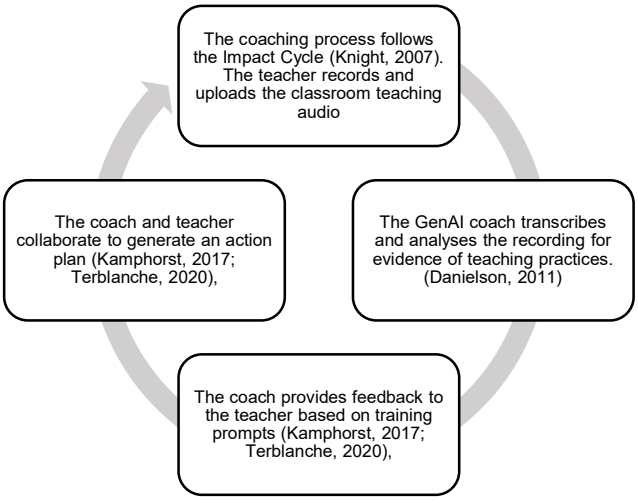

Figure 1. Theoretical models used in the GenAI coach design

4. Research Design

Three teachers from a K-12 school in Mumbai, teaching different grades and subjects, participated in the study. The teachers used the coach for three consecutive classes, after which, each teacher was interviewed by a school-based researcher. The interview protocol was designed to elicit detailed feedback on usability, usefulness and suggestions for improvement (see Table 1). Interviews were video-recorded, transcribed verbatim and analysed to inform the following three research questions.

Table 1. *Surveys alignment with Research Questions*

Survey	Research Question addressed
System Usability Survey	RQ1
Usefulness Survey	RQ2
Suggestions for Improvement	RQ3

A 10-question standardised System Usability Survey (SUS) (Bangor, 2008) generated an average composite score out of 100 across the three teachers. These scores were interpreted using established SUS benchmarks to evaluate overall usability of the AI coach.

A 10-question usefulness survey (Table 2) was designed around social learning theory constructs (Bandura, 1997) and face validated with an external educational researcher. Questions and their corresponding constructs are mentioned in the Results section 5.2 below. The survey followed a Likert Scale of 1 = strongly disagree to 5= strongly agree.

At the end of the interview, teachers responded to open-ended questions on the strengths and weaknesses of the coaching experience. From the questions listed below positive feedback questions (1-3) and negative feedback questions (4-6) were asked.

1. What are you most excited about when using an AI teacher coach?
2. What was the most helpful feature of the AI teacher coach?
3. How did using the AI teacher coach influence your teaching, if at all?
4. What are your biggest concerns about using an AI teacher coach?
5. What was frustrating or difficult about using it?
6. What suggestions do you have to improve the AI teacher coach?

Answers were transcribed and categorized as positive and negative, and tallied for frequency. The five most frequently stated positive and negative aspects of the GenAI coach were listed.

5. Results

5.1 *Usability Survey*

The System Usability Survey (SUS) revealed an average usability score across the three teachers of 70.83 (see Table 2), which implies that the coach is generally considered to be acceptable and above average in terms of usability.

Table 2. *Average Scores on the Usability Survey*

#	Questions (Likert Scale: 1: Strongly Disagree, 5: Strongly Agree)	Average Score
1	I think that I would like to use the coach frequently.	4.00
2	I found the coach unnecessarily complex.	2.00
3	I thought the coach was easy to use.	3.67
4	I think I would need the support of a technical person to be able to use the coach.	2.67
5	I found the various functions in the coach were well integrated.	3.33
6	I thought there was too much inconsistency in the coach.	2.33
7	I would imagine that most people would learn to use the coach very quickly.	4.67
8	I found the coach very cumbersome to use.	2.67
9	I felt very confident using the coach.	5.00
10	I needed to learn a lot of things before I could get going with the coach.	2.67
	SUS SCORE (on 100)	70.83

5.2 *Usefulness Survey*

On a Likert Scale of 1 to 5, the average usefulness score was 3.98, implying significant usefulness for a teacher (See Table 3).

Table 3. *Average Scores on the Usefulness Survey*

#	Questions (Likert Scale: 1: Strongly Disagree, 5: Strongly Agree)	Construct	Average Score
1	The feedback I got was accurate.	Perceived Accuracy	3.33
2	The feedback I got was useful.	Perceived Usefulness	3.67
3	The coach supported my growth as a teacher.	Empowerment	4.00
4	I would recommend this tool to other teachers.	Adoption Intention	4.00
5	I see myself continuing to use this tool in the future.	Long-Term Use Intention	4.33
6	I found the AI's suggestions clear and understandable.	Clarity and Interpretability	4.33
7	I felt confident using the AI teacher coach.	Confidence	4.67
8	I was satisfied with my experience using the AI teacher coach.	Satisfaction	3.83
9	The coach's suggestions aligned well with my teaching goals and context.	Relevance and Goal Alignment	3.67
10	My views on AI in education have changed after using the AI teacher coach.	Belief Change	4.00
	Average Usefulness Score on 5		3.98

5.3 *Suggestions for Improvements*

Answers to qualitative questions were categorised into positive and negative aspects of the coach, and counted for frequency. Table 4 lists the five most common positives and negatives about the GenAI coach design and use.

Table 4. *Top Five Positive and Negative Aspects of the GenAI Coach*

Positive Aspects	Negative Aspects
High potential for self-improvement in teaching practices	Long onboarding process
Confidentiality and privacy of feedback	No record of student talk or classroom video
Personalised, class specific feedback with evidence	Analysis takes too long (20 minutes on average per class)
Feedback is unbiased	Feedback should align to specific lesson goals, not general pedagogy
Feedback is offered in real time	Questionable use in student-centred classroom pedagogy

6. Discussion

The study highlights the promise of generative AI to design a teacher coaching tool. Usability scores (70.83) exceeded the SUS benchmark of 68 (Bangor et al., 2008), and usefulness ratings (3.98/5) suggest teachers found value in the experience. Teachers particularly appreciated unbiased, non-judgmental feedback. The private, self-directed format also offered a safe space for reflection.

Some challenges with GenAI coach use were captured. Teachers noted slow analysis. The AI captured only teacher talk, overlooking classroom dynamics such as non-verbal cues and student interactions, highly relevant in student-centered classrooms where student voice is important. Participants also requested better alignment between lesson goals and feedback,

suggesting a need for customizable, content-specific objectives beyond the Danielson Framework (Danielson, 2011). Extending the coach to pre-class planning support could strengthen the cycle of planning, implementation, and reflection (Knight, 2007).

Limitations include a small sample (three teachers from one school), short study duration, and reliance on self-reported data. Future research should involve larger, more diverse samples, integrate usage logs and classroom observations, and assess impacts on long-term practice and student outcomes, as well as differences across teacher proficiency and subjects. Further research has been planned to gauge accuracy and precision of the GenAI coach feedback in comparison to an expert human coach.

7. Conclusion

This study explored the potential of a GenAI teacher coach for scalable, personalized professional development. The findings suggest that the current application achieves moderate to good usability and is perceived as useful by teachers. The findings suggest that generative AI-based coaching applications have significant potential to supplement traditional professional development approaches, particularly in low-resource contexts.

References

Bandura, A. (1977). *Social learning theory*. Prentice Hall.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, *24*(6), 574-594. https://doi.org/10.1080/10447310802205776

Bush, R. N. (1984). Effective staff development in making schools more effective: Proceedings of three state conferences. Far West Laboratory.

Costa, A. L., & Garmston, R. J. (2002). *Cognitive coaching: A foundation for renaissance schools* (2nd ed.). Christopher-Gordon Publishers.

Danielson, C. (2011). *Enhancing professional practice: A framework for teaching* (2nd ed.). ASCD.

Joyce, B., & Showers, B. (1980). Improving in-service training: The messages of research. *Educational Leadership*, *37*(5), 379-385.

Kamphorst, B. A. (2017). E-coaching systems: What they are, and what they aren't. *Personal and Ubiquitous Computing*, *21*(4), 625-632.

Knight, J. (2007). *Instructional coaching: A partnership approach to improving instruction*. Corwin Press.

Knight, D. (2012). Assessing the Cost of Instructional Coaching. Journal of Education Finance. 38. 52-80. 10.1353/jef.2012.0010.

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, *88*(4), 547-588. https://doi.org/10.3102/0034654318759268

Reddy LA, Shernoff E, Lekwa A. A randomized controlled trial of instructional coaching in high-poverty urban schools: Examining teacher practices and student outcomes. J Sch Psychol. 2021 Jun; 86:151-168. https://doi: 10.1016/j.jsp.2021.04.001.

Sweeney, D. (2011). *Student-centered coaching: A guide for K–8 coaches and principals*. Corwin Press.

Terblanche, N. H. D. (2020). A design framework to create artificial intelligence coaches. *International Journal of Evidence Based Coaching and Mentoring*, *18*(2), 152-165. https://doi.org/10.24384/b7gs-3h05

Terblanche, N. H. D., & Tau, T. (2024). Exploring the use of a goal-attainment, artificial intelligence (AI) chatbot coach to support first-time graduate employees. *Career Development International*, *39*(3). https://doi.org/10.1177/09504222241287090