Bridging Explainable Modeling and Actionable Feedback: Multimodal AI-Augmented Pedagogical Interventions for Developing Active Listening Skills in Collaborative Problem Solving

Xiaomeng Huang(✉) and Xavier Ochoa

New York University, New York City, NY, USA
{xiaomeng.huang,xavier.ochoa}@nyu.edu

Abstract. Active listening—the ability to move beyond hearing to interpret diverse perspectives and build on others' contributions during collaboration—is a foundational skill in collaborative problem solving (CPS). Yet it is rarely explicitly taught because teachers lack scalable ways to observe, assess, and support its development during group work. This research addresses this longstanding challenge in CPS by leveraging multimodal AI to create new pedagogical interventions for active listening skill development. These interventions bridge explainable learner modeling and dialogic, actionable feedback. We introduce the Theory—Measurement—Evaluation (TME) framework, an explainable modeling architecture that operationalizes theory-informed behavioral indicators into interpretable multimodal learner models using uncertainty-aware probabilistic modeling. Rather than producing black-box evaluations, ***explainability*** is achieved by linking observable multimodal evidence directly to theoretically grounded constructs. These explainable evaluations are then translated into pedagogical feedback grounded in feedback theories, where ***actionability*** is achieved through structured dialogic guidance that promotes reflection and skill learning while preserving learner agency. This work contributes to AIED by advancing pedagogically grounded explainable AI and demonstrating how multimodal AI can augment pedagogical interventions to support the development of complex collaboration skills.

Keywords: Multimodal AI · Active Listening Skills · Collaborative Problem Solving

1 Introduction

Collaborative learning is central to education, not only as a way to build understanding across perspectives, but also as preparation for civic life, where collaboration is essential [4]. Yet participation in collaboration is rarely equal [2,12].

E. G. Blanchard et al. (Eds.): AIED 2026, CCIS 3033, pp. 254–260, 2026.
https://doi.org/10.1007/978-3-032-29794-5_39

Studies consistently show that some students dominate while others are sidelined, which limits both individual learning and group outcomes. Most interventions to date have focused on efforts to design equitable conditions—structuring tasks, scripting activities, or assigning group roles [2]. Far less attention, however, has been given to students' own capacity to notice and regulate unequal participation within their groups. A promising but underdeveloped pathway for building such capacity is active listening. The National Communication Association defines active listening as "the process of receiving, constructing meaning from, and responding to spoken and/or nonverbal messages" [13]. This definition highlights core skills such as hearing, understanding, remembering, interpreting, evaluating, and responding [1]. In collaborative learning settings, active listening requires interpreting and responding to peers' contributions in ways that make space for diverse voices and foster shared understanding. Despite its importance, active listening is rarely taught in schools. Teachers often lack the time and resources to observe subtle interactional dynamics and provide feedback, and even research has largely emphasized regulating speaking (e.g., turn-taking, talk time) [10] over listening [12]. As a result, one of the most critical levers for improving equitable collaboration—students' ability to listen actively—has remained underdeveloped.

Recent advances in multimodal learning analytics (MmLA) and multimodal AI models offer new opportunities to observe and model collaborative interaction through rich multimodal data [9]. However, our prior semi-systematic review of collaborative analytics systems [5] identified persistent limitations, including insufficient operationalization of theory-informed indicators, limited use of interpretable modeling approaches, and a lack of pedagogical grounding in analytics-supported feedback design. Many systems rely on black-box modeling techniques [11] or static dashboards that provide limited actionable guidance for learners [8]. To address these challenges, this research adopts an explainable modeling approach inspired by Evidence-Centered Design (ECD) and integrates dialogic feedback principles [14] to translate the explainable learner models into pedagogical feedback.

Two design principles guide this work: *explainability* and *actionability.* Explainability is conceptualized not merely as model transparency, but as a pedagogical bridge between automated assessment and instruction. Bayesian Skill Tracing is used to estimate theory-informed constructs from multimodal behavioral evidence while preserving uncertainty. Rather than assigning global labels (e.g., "good" or "poor" listener), the system models multiple sub-constructs (e.g., understanding checks) together with confidence estimates. These estimates are derived from aggregated multimodal evidence (e.g., clarifying questions or head nodding) detected using multimodal AI models. In this approach, AI supports evidence extraction, while evaluative interpretations remain theory-aligned and transparent. *Actionability* extends this bridge by structuring feedback across cognitive, structural, and social-affective dimensions [14], ensuring that learners not only receive guidance on what to do next but also understand the rationale, feel capable of acting, and retain agency in the learning process. For example, instead of presenting behavioral metrics (e.g., counts of clarifying questions) through dashboards, the system presents estimated levels of the "understanding checks" construct together with examplary moments (e.g., showing video clips when learners asked clarifying questions, or other understanding check behaviors). Learners are then invited to reflect on their intentions at these specific moments, their perspectives are acknowledged, and actionable guidance (e.g., strategies for inviting peers' contributions) is provided in context. Guided by these principles, this research addresses the following questions:

– ***RQ1***: How can theory-informed behavioral indicators of active listening be translated into explainable multimodal learner models, and to what extent do these models produce faithful and pedagogically useful assessments?
– ***RQ2***: Does dialogic and actionable feedback derived from explainable learner models support changes in learners' active listening skill development, including self-awareness, motivation, and observable listening behaviors, in collaborative problem solving?

2 Theoretical Frameworks

This research integrates two complementary theoretical framework: the HURIER active listening model [1] to operationalize active listening as measurable constructs, and the PEARLS framework [3] to structure dialogic, pedagogically grounded feedback.

The HURIER model conceptualizes listening as a set of interrelated processes (e.g., understanding, interpreting, evaluating, and responding) that can be manifested through observable verbal and nonverbal behaviors. While originally developed for dyadic communication and self-report assessment, this research adapts HURIER for small-group collaborative problem solving by translating listening processes into multimodal behavioral indicators suitable for computational modeling. These theory-informed constructs provide the foundation for explainable learner modeling of active listening.

To turn assessment into pedagogically meaningful intervention, this work draws on the PEARLS framework, a structured dialogic debriefing model widely used in simulation-based learning. PEARLS organizes feedback around guided reflection, facilitated discussion, and targeted instructional support. Adapted for AI-mediated feedback, it provides a pedagogical structure for transforming explainable assessments into dialogic guidance that supports reflection, learner agency, and skill development.

3 Proposed Solution and Current Progress

Figure 1 illustrates the overall approach. An initial empirical study has been completed to develop and validate the explainable modeling approach, and a second study is planned to evaluate the effectiveness of the dialogic feedback intervention within the same experimental context: small-group, in-person collaborative problem-solving tasks. In these settings, groups of 3–4 college students engage in information-distribution challenges that require integrating unique pieces of knowledge held by each participant, creating conditions that necessitate active listening and shared understanding. Group interactions are captured through multimodal audio and video recordings. Interaction data are processed during collaboration, while feedback is delivered after the session to avoid disrupting natural group dynamics. The system generates explainable evaluations from multimodal data and pedagogical feedback aligned with the modeling and pedagogical principles described earlier.

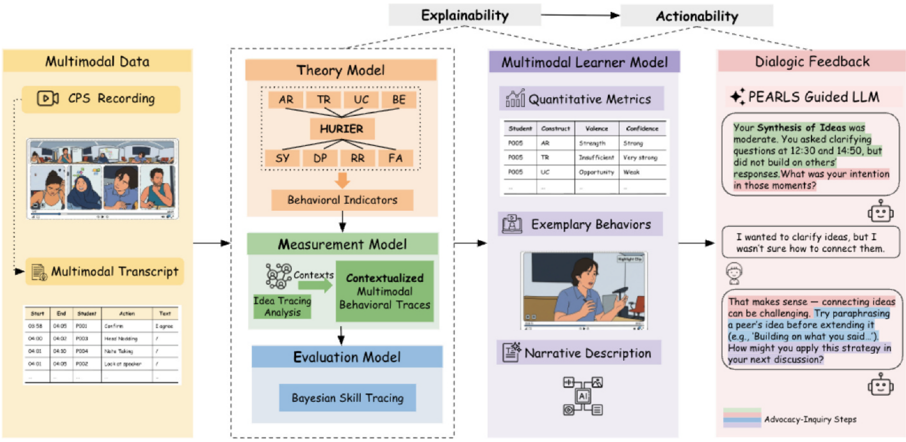

Fig. 1. System Overview: Bridging Explainable Modeling and Actionable Feedback

3.1 Explainable Modeling

To address RQ1, this work introduces the Theory-Measurement-Evaluation (TME) framework [7], an explainable modeling architecture (Fig. 1). TME transforms multimodal interaction data into interpretable learner models by explicitly linking theoretical constructs, observable behavioral evidence, and uncertainty-aware probabilistic evaluation.

TME consists of three stages. The **Theory Model** operationalizes HURIER-informed listening processes into theory-aligned behavioral indicators, including 19 verbal and 3 nonverbal behaviors (e.g., paraphrasing, clarifying questions, integrating peers' ideas, head nodding, note-taking) organized across eight subconstructs. The **Measurement Model** extracts and contextualizes multimodal behavioral traces, incorporating collaboration context through Idea Tracing Analysis (ITA) [6]. Multimodal AI techniques (e.g., language and vision models) are used to detect behavioral indicators aligned with theory-informed constructs.

The **Evaluation Model** applies Bayesian Skill Tracing, inspired by Bayesian Knowledge Tracing but adapted for collaborative discourse. Rather than treating observations as binary correctness events, behavior-in-context evidence is modeled as weighted probabilistic updates that estimate each learner's latent proficiency while preserving uncertainty. The resulting multimodal learner model supports both evaluation and downstream pedagogical intervention, producing (1) quantitative construct estimates, (2) exemplary interaction moments linked to those estimates, and (3) LLM-supported natural-language summaries derived from structured evidence. Preliminary evaluation indicates high faithfulness and usefulness of model outputs (mean ratings $> 4.4/5$ across components).

3.2 Actionable Feedback

To address RQ2, this work conceptualizes *actionability* as extending beyond feedback content (what to improve) to include feedback structure (how guidance is organized) and its social-affective delivery (how it is conveyed). The goal is to support learners' agentic and self-regulated development of active listening skills by combining actionable guidance with reflective dialogue that promotes understanding and autonomy. Building on the multimodal learner model produced by TME, feedback is delivered through a LLM-mediated conversational interface structured by the PEARLS framework and its advocacy-inquiry method [3], as shown in Fig. 1. The system presents construct estimates and exemplary moments, invites learners to reflect on their perspectives, and provides personalized actionable guidance. In this design, the LLM operates within predefined pedagogical constraints rather than as an open-ended generative agent.

To evaluate the effectiveness of this intervention, a randomized between- and within-subjects experiment with 90 college students is planned for Spring 2026. Participants will be randomized at the group level to either an active control condition involving prompted self-reflection or an experimental condition receiving the dialogic feedback intervention. All groups will complete two comparable collaborative sessions; the first session serves as baseline, and feedback or self-reflection is delivered between sessions. Outcomes will assess pre-post changes between conditions in (1) self-awareness of active listening [1], (2) motivation toward active listening [13], and (3) behavioral change measured through TME across tasks.

4 Expected Contributions

This work contributes to AIED and CPS research by advancing pedagogically grounded multimodal AI for supporting complex collaboration skills. *Theoretically*, it operationalizes active listening as a measurable construct linking active listening theory with multimodal behavioral indicators. *Methodologically*, it introduces the Theory-Measurement-Evaluation (TME) framework, connecting theory-informed indicators, multimodal evidence, and uncertainty-aware learner modeling. *Pedagogically*, it demonstrates how explainable assessments can be translated into actionable dialogic interventions that support self-regulated collaboration skill development.

References

1. Brownell, J.: Listening: Attitudes, Principles, and Skills. Routledge, New York, 7 edn. (2023). https://doi.org/10.4324/9781003316794, https://www.taylorfrancis.com/books/9781003316794
2. Cress, U., Rosé, C., Wise, A.F., Oshima, J. (eds.): International Handbook of Computer-Supported Collaborative Learning, Computer-Supported Collaborative Learning Series, vol. 19. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-65291-3
3. Eppich, W., Cheng, A.: Promoting excellence and reflective learning in simulation (pearls): development and rationale for a blended approach to health care simulation debriefing. Simul. Healthcare: J. Soc. Simul. Healthcare **10**(2), 106–115 (2015). https://doi.org/10.1097/SIH.0000000000000072
4. Graesser, A.C., Fiore, S.M., Greiff, S., Andrews-Todd, J., Foltz, P.W., Hesse, F.W.: Advancing the science of collaborative problem solving. Psychol. Sci. Public Interest **19**(2), 59–92 (2018)
5. Huang, X., Ochoa, X.: Charting the development of collaboration skills through collaborative learning analytics systems. J. Learn. Anal. **12**(1), 338–366 (2025). https://doi.org/10.18608/jla.2025.8523
6. Huang, X., Ochoa, X., Gopalakrishnan, M., Zhang, S.: Idea tracing analysis (ITA):a temporal-structural methodology for modeling collaborative discourse and contextualizing multimodal interactions. In: Proceedings of the International Conference on Computer Supported Collaborative Learning (CSCL) (2026)
7. Huang, X., Ochoa, X., Hiterer, D.: The TME framework: multimodal learner modeling for active listening skills in collaborative problem solving. In: Proceedings of the 27th International Conference on Artificial Intelligence in Education(AIED) (2026)
8. Kaliisa, R., Misiejuk, K., López-Pernas, S., Khalil, M., Saqr, M.: Have learning analytics dashboards lived up to the hype? a systematic review of impact on students' achievement, motivation, participation and attitude. In: Proceedings of the 14th learning analytics and knowledge conference, pp. 295–304 (2024)
9. Ochoa, X.: Multimodal learning analytics-rationale, process, examples, and direction. Handbook Learn. Anal. **2**, 54–65 (2022)
10. Ochoa, X., Huang, X., Charlton, A.: Unpacking the complexity: why current feedback systems fail to improve learner self-regulation of participation in collaborative activities. J. Learn. Anal. **11**(2), 246–267 (2024)

11. Rosé, C.P., McLaughlin, E.A., Liu, R., Koedinger, K.R.: Explanatory learner models: why machine learning (alone) is not the answer. Br. J. Edu. Technol. **50**(6), 2943–2958 (2019)
12. Strauß, S., Rummel, N.: Promoting regulation of equal participation in online collaboration by combining a group awareness tool and adaptive prompts. but does it even matter? Int. J. Comput.-Support. Collab. Learn. **16**(1), 67–104 (2021)
13. Worthington, D.L., Bodie, G. (eds.): The Sourcebook of Listening Research: Methodology and Measures. Wiley, Hoboken, NJ (2017)
14. Yang, M., Carless, D.: The feedback triangle and the enhancement of dialogic feedback processes. Teach. High. Educ. **18**(3), 285–297 (2013). https://doi.org/10.1080/13562517.2012.719154, http://www.tandfonline.com/doi/abs/10.1080/13562517.2012.719154