Multimodal Analytics for Collaborative Teacher Reflection of Human-AI Hybrid Teaching: Design Opportunities and Constraints

Shamya Karumbaiah[1], Pinyang Liu[2], Alisa Maksimova[3], Lea De Vylder[2], Nikol Rummel[3], Vincent Aleven[2]

[1] University of Wisconsin-Madison, USA
shamya.karumbaiah@wisc.edu
[2] Carnegie Mellon University, USA
pinyangl@andrew.cmu.edu, ldevylde@andrew.cmu.edu, aleven@cs.cmu.edu
[3] Ruhr-Universität Bochum, Germany
alisa.maksimova@cais-research.de, nikol.rummel@rub.de

Abstract. Past research shows that teachers benefit immensely from reflecting on their classroom practices. At the same time, adaptive and artificially intelligent (AI) tutors are shown to be highly effective for students, especially when teachers are involved in supporting students' learning. Yet, there is little research on how to support teachers to reflect on their practices around AI tutors. We posit that analytics built on multimodal data from the classroom (e.g., teacher position, student-AI interaction) would be beneficial in providing effective scaffolding and evidence for teachers' collaborative reflection on human-AI hybrid teaching. To better understand the design opportunities and constraints of a future tool for teacher reflection, we conducted storyboarding sessions with seven in-service teachers. Our analysis revealed that certain modalities (e.g., position v. video) might be more beneficial and less constrained than others in identifying reflection-worthy moments and trends. We discuss teachers' needs for reflection in classrooms with AI tutors and their boundaries in using multimodal analytics.

Keywords: Teachers, Multimodal Analytics, Storyboards, Reflection, Human-AI Partnerships, Collaboration.

1 Introduction

Teacher reflections on their classroom practices are shown to be effective for teachers' beliefs, practices, and reflection, and in turn, their students' learning [1]. *Reflections* are a form of deliberative thinking (c.f. [2]) that teachers do as part of either their everyday practice (reflection-in-action) or continuous professional development (reflection-on-action), with an aim to improve their classroom practices. While past empirical research on reflection highlights its benefits to teacher practices and student learning, there is a lack of consensus on how to support teachers effectively during their reflection, especially leveraging any benefits technology may have to offer [1]. Supporting

effective teacher reflection involves providing scaffolds and evidence to induce perplexity about their practices and support problem-solving [4]. Currently, deriving evidence from classroom teaching predominantly relies on videos, with technology supporting manual video annotation or capturing of reflection-worthy snippets during class [5]. Videos afford teachers the opportunity to look back on their practice, without relying on their memory during the often fast-paced nature of classroom orchestration. However, past research has also suggested that browsing through hours of video, potentially multiple times, to identify relevant evidence tends to be time-consuming [6], which may make it hard for in-service teachers due to their limited time.

In classrooms with adaptive and artificially intelligent (AI) tutors, teachers' classroom practices such as conceptual guidance, socio-emotional support, and monitoring are related to improved student learning and engagement with the AI tutor [3]. Yet, there is little research on how to support teachers to reflect on *human-AI hybrid teaching*, wherein the instructional role is shared between teachers and AI [7]. While current teacher-facing tools in AI tutors mostly include analytics on students [3], such tools could also be developed for teacher reflections, especially if they could augment teachers' intelligence about their own practices. Hence, we aim to build a reflection tool that will extract and show trends in students' learning and teachers' practices in classrooms with AI tutors and illustrate them with strategically selected examples from class sessions. However, it is underexplored how best to design such a tool.

We adopt two primary design stances: multimodality and collaboration. It is shown that multimodal access to the learning situations helps to enable teachers' reflection [8]. Accordingly, recent research in non-AI classrooms has explored how multimodal analytics from classroom data could help identify effective scaffolds for teacher reflection, what we call *reflection-worthy* moments and trends. This includes analytics built on teacher video, position, and interaction. Similarly, opportunities to collaborate with peers to co-construct better solutions are considered effective in teacher learning [9]. A review by Clarke and colleagues [1] reported that an overwhelming majority of teacher reflections are structured to be collaborative.

While multimodal analytics can open new possibilities, multimodal sensing could raise stakeholder concerns over the issues of privacy and the possibility of creating surveillance systems monitoring students' and teachers' behavior [10]. Therefore, it is critical to examine teachers' views on constraints on the ways to collect and use multimodal data. Moreover, data ownership and the ethical use of data need to be established with stakeholders to avoid the potential misuse of student and teacher data [11]. Similarly, our design stance on collaboration may raise additional social constraints for teachers while sharing their personal recordings with a colleague (e.g., awkwardness in viewing themselves on screen). While evidence from classrooms that don't match teachers' beliefs about themselves or their students may trigger effective moments of reflection with a colleague, it may need a supportive culture [12].

Thus, the two research questions that guide our analysis are: 1) *What are teachers' needs for reflecting on their practices in classrooms with AI tutors?* 2) *What are their boundaries and perceptions in using multimodal analytics for collaborative reflection of human-AI hybrid teaching?*

2 Methods

To understand teachers' needs, preferences, and boundaries for multimodal collaborative reflection of human-AI hybrid teaching, we designed a speed dating study with storyboards. The speed dating method allows us to validate teachers' needs, discover new ideas through participants' observations and suggestions, test and understand the boundaries of the tool through riskier solutions, and receive direct feedback from the primary user group to identify possible design directions.

We designed twelve storyboards to document cohesive storylines with different class scenarios with AI tutors and corresponding teacher interactions with the reflection tool. Specifically, each storyboard consists of three panels with the following outline: a description of the problem or class scenario, how a (future) tool reacts to the scenario (i.e., what data it collects and/or how it collects data), and what the tool presents to the user. We identified topics for storyboards that illustrate our key design stances: multimodality and collaboration (see Section 1). The topics correspond to the content of the reflection (e.g., teaching practice, student learning), multimodality (e.g., sensors, privacy), collaboration (e.g., with whom to reflect, data sharing), and reflection method (e.g., when to reflect, data to display, platform).

Seven teachers participated in the speed dating study. They were all math teachers from middle and high schools in the United States (6) and Croatia (1), with teaching experience ranging from 5 to 21 years. They taught a range of math topics such as algebra, statistics, geometry, and calculus. Six of them use some form of educational technology in their everyday teaching. Each session lasted for around 60 minutes. In total, we extracted 189 such quotes from the seven sessions. After grouping all the quotes by the topics, we formed 3 top-level and 14 mid-level clusters.

3 Results

3.1 Aligning Multimodality with Reflection Needs

Our analysis reveals teacher needs and priorities for reflection that can inform the modalities that could maximize the benefits for teachers (see discussion in Section 4).

Teachers need data on their *own* classroom practices to reflect on how effective they are. Three teachers wanted to know whether their interventions and any new activity they try in the classroom is effective for student learning. One teacher expressed an interest in seeing how their help-giving supports students' learning with the AI tutor. While explaining struggles with their current approach, they said: "Usually, they'll ask me a question, and I'll answer it. And I'll come back in a few minutes to see [if] they are answering more questions. Are they still stuck in the same spot? Are they still making the same mistake? Then maybe my help wasn't too effective, and maybe I've got to come up with another way to explain that problem."

Teachers need data that would help them reflect on their *students' needs* for support. Teachers said that they needed data to reflect on how students are doing: do several people in the class struggle with something that should be explained in more detail?

Three teachers suggested that other content-related data, such as common misconceptions among their students, would be helpful in reflecting on the improvements they need to make. Similarly, they wanted to know which students they were helping the most and whether they were ignoring anyone who may need them.

Data about student *well-being and stress* could be helpful if it were accurate and the causes were reliably interpreted. The majority of the teachers (4 out of 7) were interested in reflecting on their students' well-being. While data-driven evidence of student well-being was appreciated by teachers, one teacher raised concerns about the accuracy of inferring students' mental states and the factors associated with them.

3.2 Boundaries of Multimodal Data Collection and Sharing

Besides the financial cost, there could be logistical (e.g., device setup and maintenance) and social (e.g., awkwardness in hearing one's own voice, the feeling of surveillance, and data sharing) constraints in introducing multimodal sensors to classrooms.

Cameras are perceived as too invasive and untrustworthy. The majority of the teachers (5 out of 7) emphasized that cameras are not allowed in their classrooms. The two common reasons cited were privacy and trust. As one teacher explained, "You feel you don't want your kids to be filmed during school, and as a teacher, I feel like I don't want to be filmed during school. I'd like to have some bigger level of trust." Still, two teachers expressed their inclination to have their classes videotaped, especially in their bigger classrooms. However, the goal was not to reflect on their classroom practices. Rather, they were interested in cameras as a surveillance tool to observe students who tended to be off-task and "mischief." This shows that teachers may have contradictory notions of privacy and invasiveness.

Collecting data on teacher positions, student-AI interaction, and demographic is acceptable. As an alternative to cameras, teachers (5 out of 7) suggested that the AI tutor log data would be useful and "non-invasive," to monitor student learning and progress. Log data was also considered helpful in providing evidence for further investigations on the concept students may be struggling with. To reflect on their own practices, collecting data on their position was perceived as non-invasive and helpful. One teacher said, "I think this [teacher position data] is very helpful - like teachers tend to not emphasize or tend to support a kind of students than the other." Relatedly, all teachers objected to having any sensor that may distract the class. Half the teachers were comfortable sharing their students' demographic information, as long as it was one-time and not too time-consuming.

Physiological data is seen as a way to detect what they cannot otherwise monitor, but there are concerns about their accuracy and lack of context. Having multimodal data, potentially including physiological data, meant for teachers that they can reflect on students' learning and mental states that are not accessible in other ways. One teacher noted how younger students sometimes are not very good at expressing their feelings. However, teachers were concerned that the accuracy of the physiological data collected may be questionable and that "[they] could be very easily skewed based on something that happened in a previous period that you're not even aware of."

Teachers want to be able to decide who to share their data with, and some are wary of their colleagues' potential misuse and misinterpretation of them. Four teachers objected to sharing their class data with their school administrators. One of them said, "The supervisor is a person that doesn't know what's going on, and then still tries to micromanage things, and that will make the teacher want to leave." In contrast, teachers were relatively more comfortable in sharing their data with a colleague. Opinions on this split. Three teachers suggested that it could provide great opportunities to reflect together on their teaching practices and learn from each other. Yet, the other three teachers expressed concerns about sharing class data as they feel uneasy about the situation of comparing themselves and their class to the others. Nevertheless, one thing that all the teachers agreed on, is to have control over their own class data. If someone wants access to their data, they should ask for permission first.

3.3 Opportunities and Constraints in Multimodal Collaborative Reflection

In addition to understanding teacher needs for reflection and potential boundaries in data collection and sharing, our analysis revealed teachers' desires and constraints about how they would like to reflect with multimodal data and the role of collaboration.

Collaborative reflections are more effective, but the nature of collaboration may vary. The majority of teachers (4 out of 7) expressed interest in reflecting collaboratively with colleagues, recognizing its benefits in improving their classroom practices. One participant said, "You know it's good to have someone else look at it from a different perspective to help you with reflections." While some teachers shared a preference for paired reflection on specific problems, others were comfortable with discussing best practices more generally as a bigger group. Two participants, in particular, suggested doing reflection with teachers in the same subject or department. In contrast, another teacher stated that student differences might limit how useful data from another class could be, despite being the same grade and subject.

Effective collaborative reflection needs a positive relationship. Teachers also expressed some constraints in making collaborative reflection work and highlighted how important it is to have a positive relationship with the collaborator. One teacher reasoned out the potential tension arising from viewing collaboration as competition. Only one teacher preferred doing reflections alone. They expressed feeling socially awkward when suggesting improvements to a teacher who may not be doing well.

4 Discussion

Several instances in our storyboarding sessions highlighted the "state of perplexity or doubt" and the need for "subsequent inquiry" that teachers express during reflection about their classroom practices. Such deliberative thinking separates the kind of support teachers need for reflection from what is available now in teacher-facing orchestration tools in AI tutors. Overall, all teachers expressed a strong interest in collecting classroom data to reflect on how effective their classroom practices are in general and how well they meet individual student needs. However, teachers varied (sometimes even

contradicted one another) in their perceptions of how useful or constraining the different modes of data collection would be in supporting their reflection needs. The majority, if not all, teachers shared the opinion that cameras are too invasive. This finding raises serious concerns about the over-reliance on videos for teacher reflections. Our findings revealed a few alternatives for classrooms with AI tutors (e.g., position, interaction log, physiological data) that tend to be acceptable to teachers, some of which may also apply to traditional classrooms. Other ethical concerns around sensors collecting biometric data did not come up in the study as relevant for the teachers. However, it needs to be taken into account that these approaches are highly debated. Lastly, since the majority of the teachers in this study found collaboration to be effective in general and already a part of their everyday practice, our tool should foster collaborative. As per teachers' strong preference for data control, the tool needs to be designed to allow teachers to own and share their data with those they trust. While data-driven reasoning was appreciated by all teachers, they also realized that a few cases might be beyond an AI tool's capabilities (e.g., external factors that may be causing student stress). Past research on AI-based teacher tools has noted the benefit of human-AI synergy, wherein teachers make sense of a situation alerted by a tool using their rich contextual knowledge and gathering more information by talking to the students [6]. Hence, an open question for future design activities is how to design for effective AI-teacher complementarity in reflection tools.

References

1. Clarke, S., Gates, Z., Foster, M., &Others. (2020) *Mapping the Design Space for Teacher Learning Through Reflection*. https://repository.isls.org//handle/1/6535.
2. Dewey, J. (1910). *How We Think*. Lexington, MA: D.C. Heath and Company.
3. Holstein, K., McLaren, B. M., Aleven, V. (2018) Student learning benefits of a mixed-reality teacher awareness tool in AI-enhanced classrooms. *Artificial Intelligence in Education*.
4. Kori, K., Pedaste, M., Leijen, Ä., Mäeots, M. (2014). Supporting reflection in technology-enhanced learning. *Educational Research Review.*
5. Fleck, R., Fitzpatrick, G. (2009) Teachers' and tutors' social reflection around sensecam images. *International Journal of Human-Computer Studies*.
6. Santagata, R. Bray, W. (2016) Professional development processes that promote teacher change: the case of a video-based program focused on leveraging students' mathematical errors. *Professional development in education*.
7. Karumbaiah, S., Borchers, C., Shou, T., Falhs, AC., Liu, P., Nagashima, T., Rummel, N., Aleven, V. (2023) A Spatiotemporal Analysis of Teacher Practices in Supporting Student Learning and Engagement in an AI-enabled Classroom. *Artificial Intelligence in Education*.
8. Porayska-Pomsta, K. (2016) AI as a methodology for supporting educational praxis and teacher metacognition. *International Journal of Artificial Intelligence in Education*.
9. Darling-Hammond, L., Richardson, N. (2009) Teacher learning: What matters? *Educational Leadership*.
10. Cukurova, M., Giannakos, M., Martinez-Maldonado, R. (2020) The promise and challenges of multimodal learning analytics. *British Journal of Educational Technology*.
11. Ogan, A. (2019). Reframing classroom sensing: Promise and peril. *Interactions*.
12. Prilla, M., Blunk, O., Chounta, I. A. (2020) How does collaborative reflection unfold in online communities? An Analysis of Two Data Set. *Computer Supported Cooperative Work*.