RAPID: DRL AI:
Scaffolding Automated Feedback for Teachers

Briefing Paper

Project Team

Dora Demszky, Assistant Professor, Stanford University (PI)
Janet Carlson, Professor, Stanford University (Co-PI)
Heather Hill, Professor, Harvard University (Co-PI)
Jing Liu, Assistant Professor, University of Maryland (Co-PI)
Jim Malamut, PhD Student → now Postdoctoral Scholar, Stanford University
Christine Bywater, Project Manager, Stanford University
Michele Reinhart, Instructional Coach, Stanford University
Samantha Booth, Project Manager, Harvard University
Jonathan Thompson, Project Manager, Harvard University
Hannah Rosenstein, Project Manager, University of Maryland

Partner Professional Learning Programs

Mathematical Quality of Instruction (MQI) at Harvard University
Center to Support Excellence in Teaching (CSET) at Stanford University

Background

Teachers face an abundance of technology tools and platforms and associated requests to integrate them into their teaching practice. According to a recent report, a single teacher uses an average of 148 unique ed tech tools annually (Instructure, 2022). This continual flow of new technologies and the overload of information they can create may lead to understandable resistance from teachers to embracing new tools, even if those tools, if used effectively, would be beneficial to their teaching and their students' learning (Klein, 2022). AI tools are no exception to this pattern. While well-resourced schools tend to provide support for their teachers in adapting new AI technologies, teachers in Title I schools rarely receive such support, which further widens educational inequities (Herold, 2017). Thus, the field is facing an urgent need to facilitate the effective integration of AI technologies into schools where their benefits can be maximized while reducing the burden on teachers' time—before burn-out or ineffective use of technology aggravates teacher turnover and widens achievement gaps.

To facilitate effective integration, we proposed to take advantage of two promising lines of work. First, recent studies inform how we can use AI tools to provide feedback to STEM teachers to improve their instruction. For example, Guskey & Link (2022) note that teachers desire personalized, non-judgmental,

consistent feedback to help them improve; AI tools can help provide such feedback in a way that is cost-effective and scalable. AI-powered feedback tools for teachers have demonstrated success in improving instruction and student outcomes in online STEM learning environments (Demszky et al., 2023; Demszky & Liu, 2023). Second, instructional coaches are available to most teachers, and coaching is widely regarded as one of the most promising forms of professional development (Kraft et al., 2018). There are challenges, however, with how coaching assignments are made. Coaches are often assigned more teachers than they can reach on a regular basis and as a result they typically engage in a limited number of coaching cycles with any one teacher. Embedding AI feedback in teacher coaching could solve these challenges and enhance the effectiveness of teacher professional learning. Given that teachers are the single most influential school-based factor in student success, this approach brings the potential to greatly reduce inequities by improving STEM instructional quality. Furthermore, we can address systemic inequities by bringing this resource to schools that serve the most marginalized student populations.

Objectives & Primary Outputs

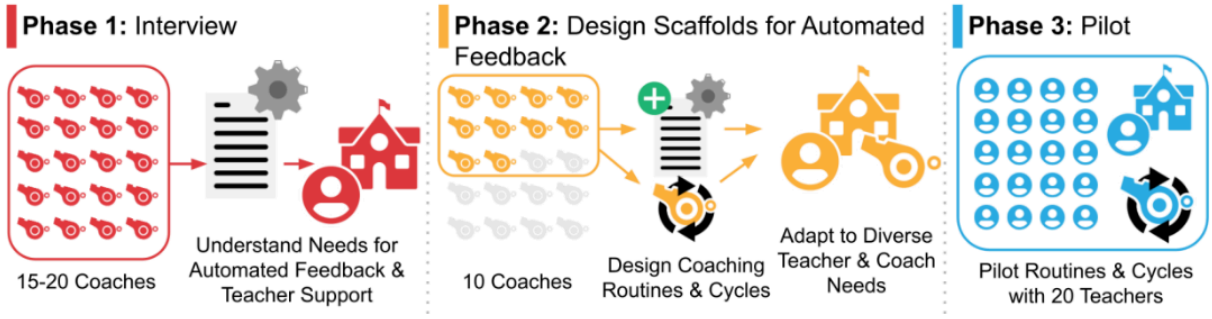
Our study aims to explore and demonstrate how AI technologies can be effectively integrated into instructional improvement routines to maximize their benefits while minimizing the time burden on teachers and coaches. Specifically, we seek to understand how coaches can enhance the impact of AI-powered teacher feedback tools by leveraging their advantages—such as cost-effectiveness, scalability, customizability, data-drivenness, and privacy—while addressing technical and time-related barriers to adoption.

Primary outputs of this project include a set of routines that coaches and teachers can use to integrate AI feedback into their school improvement practice, as well as freely available training materials for each of these routines. We developed and piloted these materials with a focus on grade 4-8 math coaches and teachers serving classrooms with a large percentage of marginalized students.

Project Activities

In order to achieve the above objectives, we proposed a three-phased approach, summarized in Figure 1.

Figure 1: Overview of Proposed Activities.



As Figure 1 shows, in the first phase of this study, we proposed to interview 15-20 instructional coaches to understand how automated feedback might be integrated into the work they do with teachers. In the second phase, we proposed that teams at both MQI and CSET would design coaching cycles and conversational routines that meet the needs of diverse teachers, coaches and schools. In the third phase, 10 coaches were proposed to pilot the coaching cycles and conversational routines created in Phase II with 20 teachers. We focus on coach-teacher pairs in grades 4-8 mathematics, with an emphasis on classrooms with a high percentage of marginalized students.

Table 1 includes a summary of project objectives, proposed activities linked to each objective and completed activities. We had completed each proposed activity with some minor modifications. For example, we were able to pilot materials with 18 instead of 20 teachers due to two of the coaches only providing data for one of the teachers they support. At the same time, we collected more data than originally proposed, as we included an additional "baseline" observation for each coach-teacher pair, to capture their practice in the absence of automated feedback tools. This baseline data allows us to better understand how automated feedback impacts the interaction between the same coach-teacher pair.

Table 1: Summary of project activities.

Objective	Proposed Activities	Completed Activities
Understand needs for automated feedback and teacher support	Interview 15-20 instructional coaches	• Interviewed 20 coaches
Design scaffolds for automated feedback	Design coaching routines that integrate automated feedback responsive to different coach and teacher needs	• Developed three protocols for using automated feedback in coaching conversations • Created training materials for each protocol
Pilot coaching routines with teachers	Pilot the use of automated feedback tool and supporting protocols with 20 teachers	• Recorded 47 coaching conversations between 11 coaches and 18 teachers for a total of 23 hours • 18/47 recordings were to

		establish baselines without the use of tool or protocol ● 12/30 recordings involve the coach and teacher utilizing automated feedback ● Interviewed 10 coaches after piloting

Outcomes & Results

We began by coding the initial interviews with 20 instructional coaches. These interviews revealed some specific ways in which coaches anticipated the automated feedback tool might be useful to them and the teachers they work with. Coaches also shared concerns about ways in which it might be misused or otherwise harmful. Finally, the interviews revealed some specific suggestions for improvements to the features of the automated feedback tool (available at mpoweringteachers.stanford.edu) and ideas for implementing it into a coaching conversation.

The most frequently identified anticipated benefit of using the tool included being able to easily revisit specific moments from an observed lesson by clicking on a transcript or other visual representations. Many coaches also anticipated utility in how the tool presents lesson information “objectively” or without interpretation or evaluation. One coach went so far as to describe the tool as giving teachers an opportunity to make their own meaning of the information presented. Some efficiency benefits were also named in that the tool might eliminate the need for coaches to take as many notes or even be physically present for an observation. Some coaches discussed how this might free them up to take higher-level notes instead.

Table 2: Anticipated Benefits of Tools: Initial Interviews of Instructional Coaches.

Anticipated Benefit	**Number of Coaches Who Mentioned**
Being able to click and scroll through the transcripts and to revisit specific times and moves	10
The visual representations	9
Tool presents information without automatic interpretation or suggestion ("objective")	7
Eliminates the need for coach to scribe by hand or collect low-inference data	7
Allows coach to "observe" without being physically present	4
Transcripts might feel less intimidating to teachers who are reluctant to watch their own recordings	1
Could be a useful way to support teachers who are missing out on coaching due to coaches being assigned to more emergency situations/teachers with more pressing	1

needs	
Could be used to build a library of effective coaching transcripts for learning purposes	1
Gives teachers opportunities to make meaning of the information presented	1
Supports the shift of attention away from "is [what the teacher did] good or bad?" to something more like 'how did the students respond when the teacher did...?'	1

The two most prevalent concerns from coaches were around protecting the privacy of teachers and students and ensuring the accuracy of recorded lessons. Additionally, some coaches spoke of dangers related to piling extra responsibilities on already overworked teachers. Others anticipated limitations related to only being able to record whole-class conversations and missing opportunities for useful data from small group or partner discussions. Coaches spoke to various potential problems related to misplaced incentives such as perpetuating oversimplified views of “academic language”, privileging of easy-to-detect talk moves over others, or otherwise providing a platform for overly-narrow high-stakes evaluation of teachers.

Table 3: Anticipated Concerns of Tools: Initial Interviews of Instructional Coaches.

Concern	Number of Coaches Who Mentioned
Recording and privacy concerns	8
Accuracy concerns, including for people with accents that are more difficult for the model to recognize. Also including bias in how talk moves are detected. How to build/maintain trust despite known accuracy issues.	8
Avoid making the platform feel like extra unnecessary work/making technology feel accessible	5
This appears to only be able to work on whole class discussions or other participation structures where only 1 person is speaking at a time. How might this work if students are working in small groups or pairs? How many mics are needed?	5
Transcripts don't include things like gestures and intonation which can matter in observations	3
This could be misused (e.g. for evaluation and incentives corrupt measures)	3
The ordering of the data. What should be the first thing teachers/coaches see? Which metrics should be grouped together?	3
"No talk move" might be an overstatement since there are lots of important math talk moves not covered in the 3 currently available	2

Equity of voice does not always mean quantity of words in participation, could be misleading and/or incentivize bad practices	1
Could perpetuate oversimplified views of "academic language"	1

A number of suggestions for improving the automated feedback tool were also made during the initial interviews. We were able to implement some of these features immediately, before the pilot began, and others by the beginning of the following school year. Some features that were developed immediately included word cloud filters for both teachers' and students' most frequent words and allowing users to manually edit and annotate the transcripts. Other suggested features added during the year were support for linking transcripts and feedback to multiple accounts and giving users the ability to upload their own transcripts to receive automated feedback. Other suggestions, such as creating more metrics for other talk moves, are still in development. The participants also shared ideas for ways to support coaches and teachers in using the tool effectively. Many of these were incorporated into the three protocols we developed for potential ways to utilize automated feedback in coaching conversations.

We developed three different protocols to scaffold the incorporation of automated feedback into coaching cycles. This development tried to anticipate different use cases such as minimal or "lite" use, teacher self-reflection, coaches who are or aren't able to physically observe a lesson, etc. The goal of creating three sets of protocols was to provide options for supporting coaches and teachers in utilizing the tool and better understand which aspects of the tool the participants took up. Each protocol provides options for teachers and coaches to reflect on the data presented in the tool and make pedagogical decisions about teaching practice. The protocols are now freely available on our website: https://www.mpoweringteachers.com/projects/2748-application-evaluation.

To analyze the kinds of coaching practices encouraged by each protocol, we applied Boguslav's (2024) Framework for Describing Coaching Discourse to code the moves suggested by each protocol. We are currently in the process of applying the codes from Boguslav's framework (2024) to each video of coaching conversations to get a better understanding of the ways in which the tool and the extent to which the supporting protocols were used in practice.

Significance

Our study is the first to provide insights into integrating AI-driven feedback into teacher coaching, contributing to a broader understanding of the challenges and opportunities in incorporating AI within existing instructional processes. The protocols we developed will support professional learning organizations, as well as district-based coaches and teachers, who aim to adopt automated feedback in their practice. Ultimately, our project will reveal how AI can be harnessed to improve teacher effectiveness and student learning in real-world educational settings in a scalable way.

References

Boguslav, A. (2024). Parsing Coaching Practice: A Systematic Framework for Describing Coaching Discourse. AERA Open, 10. https://doi.org/10.1177/23328584241263861

Demszky, D., Liu, J., Hill, H., Jurafsky, D., & Piech, C. (2023). Can Automated Feedback Improve Teachers' Uptake of Student Ideas? Evidence From a Randomized Controlled Trial In a Large-Scale Online Course. *Educational Evaluation and Policy Analysis*.

Demszky, D., & Liu, J. (2023). M-Powering Teachers: Natural Language Processing Powered Feedback Improves 1:1 Instruction and Student Outcomes. L@S '23: Proceedings of the Tenth ACM Conference on Learning @ Scale.

Guskey, T. R., & Link, L. J. (2022). What Teachers Really Want When It Comes to Feedback. Educational Leadership, 79(7), 42-48.

Herold, B. (2017, June 12). Poor Students Face Digital Divide in How Teachers Learn to Use Tech. Education Week. https://www.edweek.org/leadership/poor-students-face-digital-divide-in-how-teachers-learn-to-use-tech/2017/06

Instructure. (2022, August 24). EdTech top 40 report shares the latest on the usage of digital tools during the 2021-22 school year. Instructure. https://www.instructure.com/resources/blog/edtech-top-40-report-shares-latest-usage-digital-tools-during-2021-22-school-year?redirectSource=LP&redirectPath=%2Ftop40

Klein, A. (2022, March 8). Tech Fatigue Is Real for Teachers and Students. Here's How to Ease the Burden. Education Week. https://www.edweek.org/technology/tech-fatigue-is-real-for-teachers-and-students-heres-how-to-ease-the-burden/2022/03

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

Appendix

Interview Protocol

Hi, thank you for participating in our interview/focus group. My name is [Name], and I am part of the team of researchers from Stanford, Harvard and the University of Maryland studying how automated feedback could potentially be integrated into teacher coaching, within existing processes and in ways that compliment the core competencies and practices of a coach. *[Team member's name] is also in the session and will be observing and taking notes.* We will be asking you some questions to understand your successes and challenges with coaching and the ways you could see using an automated feedback tool to support your role as a coach and mitigate the most common challenges you face in coaching.
Before we get started, did you have a chance to read the consent form and do you have any questions? Can you please indicate your consent verbally.

PART 1: Background (10 min)

1. Please tell us a little bit about yourself- your professional role, and the teachers and students you serve
    a. *Probing:* Tell us how you became a coach?

PART II: Current Coaching Experiences (20 min)

2. We want to know about you as a coach and about your coaching role
    a. Can you tell me how you work 1:1 (more specific: can you tell us what a typical coaching cycle looks like?)
        i. Probing: Is there a particular coaching approach or coaching tools/resources at the site or district level that guide your work?
    b. Do you currently use video, audio, or any other technology to support your coaching work? If so, what? How?
        i. *Probing:* What coaching tools do you use to support your coaching?
    c. How many teachers do you coach 1:1? How is that determined?
        i. Probing: TIme constraints, teachers interests, etc.
    d. What does a typical day and week look like for you in your current coaching role?
3. What do you do well as a coach?
4. What aspects of your current coaching role do you find most frustrating or challenging?
    a. *Probing question* - Why?
5. What are a few things you would change about your current role that you think would make you a more effective coach?
    a. *Probing question*- How would these changes help?

PART III: Scenarios and Reactions (25 min)

Framing: Technology is transforming everyone's job, and we want to hear from coaches about how you feel technology might be best used in your role. Technology should never replace coaches, and we understand the value your relationship with your teachers has on your ability to coach. We're curious to understand your ideas around a platform that uses an automated feedback tool to analyze teacher transcripts or recordings. Teachers record a lesson, upload it to a platform, and get feedback on aspects of classroom dialogue – for instance, the amount of teacher vs. student talk time, teachers' use of open-ended questions, or teacher 'uptake' of student ideas. Teachers receive information about the frequency of use of these talk moves, and also see examples of how they used them with students. This feedback is non-evaluative and, because it is generated automatically, more objective than typical classroom observations.

From previous studies, we found that teachers had difficulty translating descriptive automated feedback into actionable improvements to their practice when they did so in isolation, so we'd like to consider ways to integrate these tools in complementary ways to how coaches are already working with teachers.

For the next part of our interview, I am going to show you an example of data from (not real data!) a platform that analyzes teachers' transcripts or a classroom recording with analysis using an automated feedback tool. I will then ask you to take some time looking at sample feedback generated from the data in order to respond to a few questions. Do you have any questions before we begin?

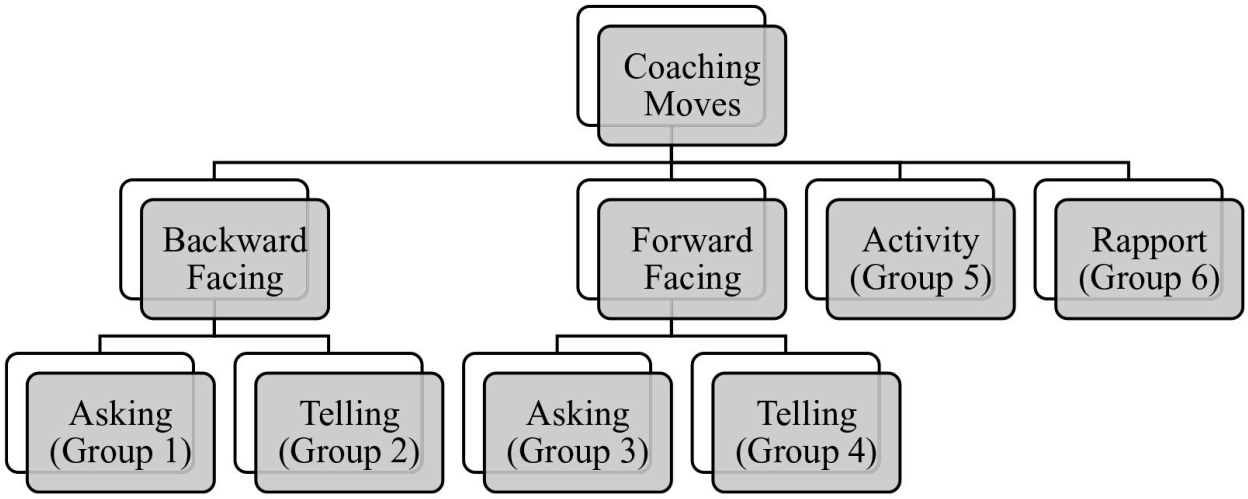
a. Here are some samples of automated analysis and feedback the platform produced from the recording you took of Teacher A's class. You can see that this teacher is getting feedback on uptake, reasoning, and questioning. Take a moment to look at these samples. (*Give participant 3-4 minutes to explore, and ask questions)*
    i. What questions do you have about the analysis and feedback examples?
    ii. What do you like about what you see? Why?
    iii. What do you find less useful? Why?
    iv. Would this data, in combination with your own notes from your observation of his class, be helpful for a coaching conversation with the teacher? Explain
        1. *Probing:* Is there anything missing that you think would be helpful to include?
b. Now that you have seen examples of automated feedback,
    i. What barriers do you see in using this platform?
    ii. Can you imagine a teacher you coach recording themselves and uploading it to the app for you to view?
    iii. Are there other ways you might use the recordings and data as part of your coaching work?
        1. *Probing: In 1:1 coaching scenarios, in PLC team meetings, to understanding teachers' instruction and plan for professional development*

PART IV Conclusion (5 min)

1. Do you have any other thoughts to share?
2. Do you have any questions for me?
3. Once this platform is developed, would you be interested in piloting this platform with teachers?
    a. Do you have colleagues who might be interested?

Thank you for your time! The compensation for the $50/hr Amazon gift card will be emailed to you upon the completion of the study at the email address you've provided to us.

Coding Framework



Boguslav, A. (2024). Parsing Coaching Practice: A Systematic Framework for Describing Coaching Discourse. AERA Open, 10. https://doi.org/10.1177/23328584241263861