The Role of Automated Post-Rehearsal Feedback on Pre-Service Mathematics Teacher Questioning

Bywater, J. P., Lilly, S., Datta, D., Ansari, S., Mohammad, A., Watson, G. S., Brown, D., & Chiu, J. L.

Automated Post-Rehearsal Feedback
Pre-Service Mathematics Teacher Questioning
Teaching Simulation Tools

Rehearsal simulations create opportunities for pre-service teachers to build questioning skills before they enter the classroom, yet such experiences are rarely accompanied by detailed individualized feedback. This exploratory study considers the role of automated feedback on pre-service teacher questioning in an elementary mathematics methods course. Twenty-seven pre-service teachers used a chat-based dialogue simulator three times and were provided with automated feedback after each rehearsal. Results show that over the three rehearsal attempts there was a statistically significant increase in the number of probing and exploring questions asked by the pre-service teachers. Results suggest that teaching simulation tools that incorporate automated feedback can effectively contribute to the development of high-quality teaching practices.

Introduction

Rich and inclusive classroom conversations provide K-12 students with opportunities to develop their mathematical understanding and their mathematical identities (Aguirre et al., 2013). These conversations occur when teachers ask questions which probe and explore student thinking rather than questions which assess procedural fluency or tell students information (NCTM, 2014). To ask these types of questions, teachers need to be able to effectively anticipate and respond to student ideas (e.g., Jacobs et al., 2010) by using questions and feedback strategies appropriate to the specific learning context (Sherin, 2002; Shute, 2008).

Research indicates teachers need intentional opportunities to rehearse and build their questioning skills before they enter the classroom (e.g., Grossman et al., 2009; McGarr, 2021). Further, they need high-quality feedback in scenarios authentic to the pre-service teachers (Cohen et al., 2020), and the feedback should be immediate and actionable (Pianta & Hamre, 2009). Such experiences are limited within most teacher education programs (Forzani, 2014). When provided, they rarely include detailed individualized feedback (Shaughnessy & Boerst, 2018) and the impact of the feedback on pre-service teachers' mathematical questioning skills is rarely studied (an exception being Mikeska et al., 2024).

As part of a larger exploratory case study (Yin, 2018), we pilot an AI-based teaching simulator which provides automated question-type feedback within an elementary mathematics methods course to consider its impact on pre-service teachers mathematical questioning. Specifically, in this paper, we address the following research question: How does automated feedback impact the type of questions pre-service teachers ask during rehearsals?

Background

Of the various validated measures of mathematics teaching, only the Instructional Quality Assessment (IQA; Boston & Candela, 2018) has a specific sub-score which measures the quality of the questions teachers ask. As such, it is particularly well suited for providing validated feedback to pre-service teachers (Boston et al., 2015). The IQA draws upon the categories of question type developed by Boaler and Brodie (2004) and measures the degree to which teachers use probing and exploring questions in contrast to procedural questions. Several classifiers have been developed to classify what teachers say, including one that is aligned with the IQA (Datta et al., 2023; see Table 1). This classifier was developed by using a pre-labeled data set of mathematics teacher questions to fine-tune the RoBERTa large language model and reports a 76% accuracy, which is considered state-of-the-art (Datta et al., 2023).

Table 1

Classifications used by the teacher question type classifier

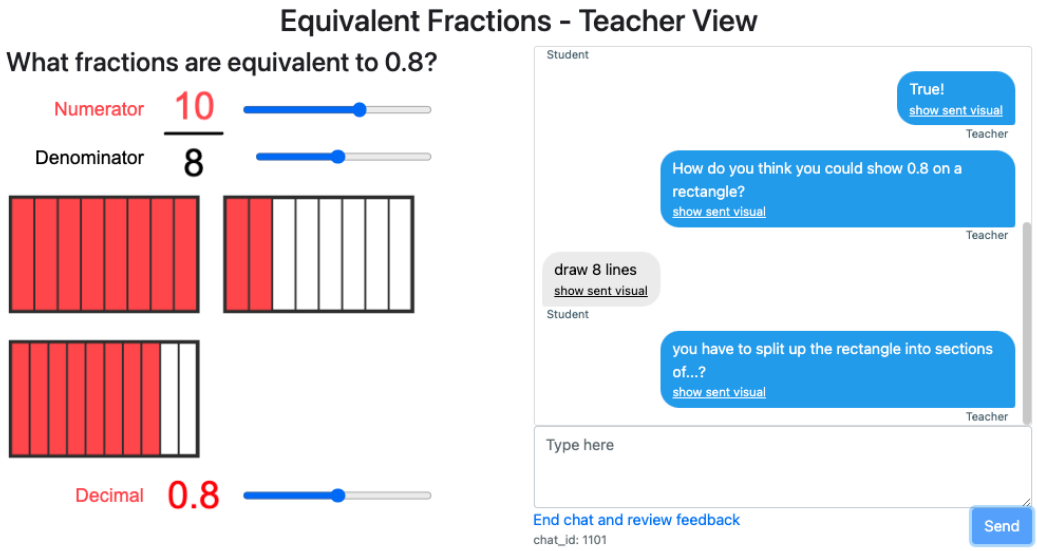
Classification	Description	Example
Probing and Exploring	Clarifies or elaborates student thinking. Focus on mathematical ideas, meanings, and connections	How did you get that answer? Explain to me how you got that expression. Why is it staying the same?
Procedural	Elicits facts, single response answers, or recall of a procedure	What is 3 x 5? Does this picture show one-half or one-quarter? What do you subtract first?
Expository	Provides cueing or mathematical information to students without engaging ideas	The answer is three, right? Between the 2?
Other	General or non-math talk; everything else.	How's your day?

The AI-based Classroom Teaching Simulator (ACTS)

This exploratory study examines the impact of the automated question-type feedback provided within the AI-based Classroom Teaching Simulator (ACTS). ACTS provides pre-service teachers with opportunities to rehearse conversations with students and then provides automated post-rehearsal feedback. As shown in Figure 1, pre-service mathematics teachers can use ACTS to interact with a virtual student by exchanging written text and virtual manipulatives. For example, in Figure 1, pre-service teachers are given a task prompt and a virtual manipulative that can be changed by adjusting the interactive sliders.

Figure 1

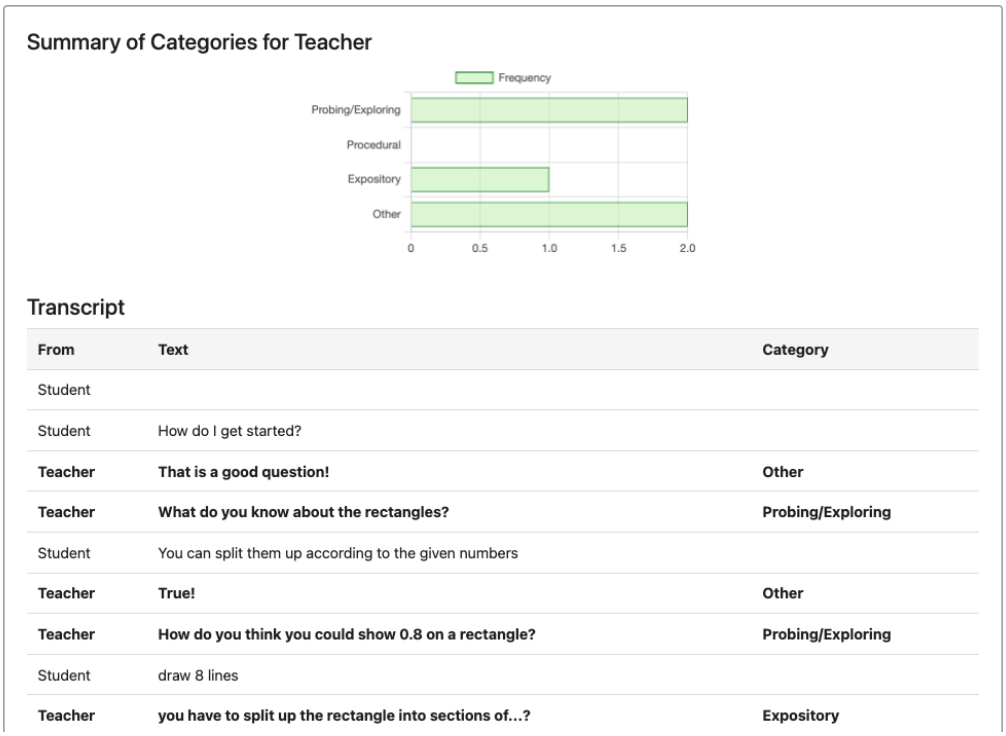
Screenshot of an example ACTS scenario



At the end of the conversation, ACTS provides immediate automated formative feedback about the conversation, which can be customized to use different classifiers. For example, in Figure 2 the feedback leverages the IQA-aligned classifier to provide both a graphical summary and a line-by-line analysis of the conversation transcript.

Figure 2

Screenshot of an example ACTS feedback



Transcript

From	Text	Category
Student		
Student	How do I get started?	
Teacher	**That is a good question!**	**Other**
Teacher	**What do you know about the rectangles?**	**Probing/Exploring**
Student	You can split them up according to the given numbers	
Teacher	**True!**	**Other**
Teacher	**How do you think you could show 0.8 on a rectangle?**	**Probing/Exploring**
Student	draw 8 lines	
Teacher	**you have to split up the rectangle into sections of...?**	**Expository**

Methods and Analysis

As part of a larger exploratory case study (Yin, 2018), we define the case as 27 pre-service teachers who are enrolled in a single section of an elementary mathematics methods course at a large mid-Atlantic university and who engaged with the ACTS simulator during their course. We chose an exploratory case study methodology as we are considering what is happening within this single, bounded context, and our research question asks “how”, meaning it does not require control over the pre-service teachers’ behavior, and focuses on events that occur in-the-moment (Yin, 2018, p. 9).

To engage with the ACTS simulator, the 27 pre-service teacher participants were assigned partners; one played the “teacher” role while the other played the “student” role. One researcher partnered with the remaining student. When each conversation concluded, the participant playing the “teacher” role received automated feedback from ACTS about the types of questions they had asked. This was repeated with the pairs switching roles until each participant had played the “teacher” role three times.

To test for changes in the counts of each question type across the three attempts we used a growth curve model with random intercept and random slope coefficients (see Figure 3).

Figure 3

Growth curve model with random intercept and random slope coefficients

$$y_{ti} = \pi_{0i} + \pi_{1i}(attempt)_{ti} + e_{ti}$$
$$\pi_{0i} = \beta_{00} + r_{0i}$$
$$\pi_{1i} = \beta_{10} + r_{1i}$$

Results

This section includes quantitative findings about the types of questions asked in each attempt. First, the total counts of each question type asked at each attempt are shown in Table 2 and in Figure 4.

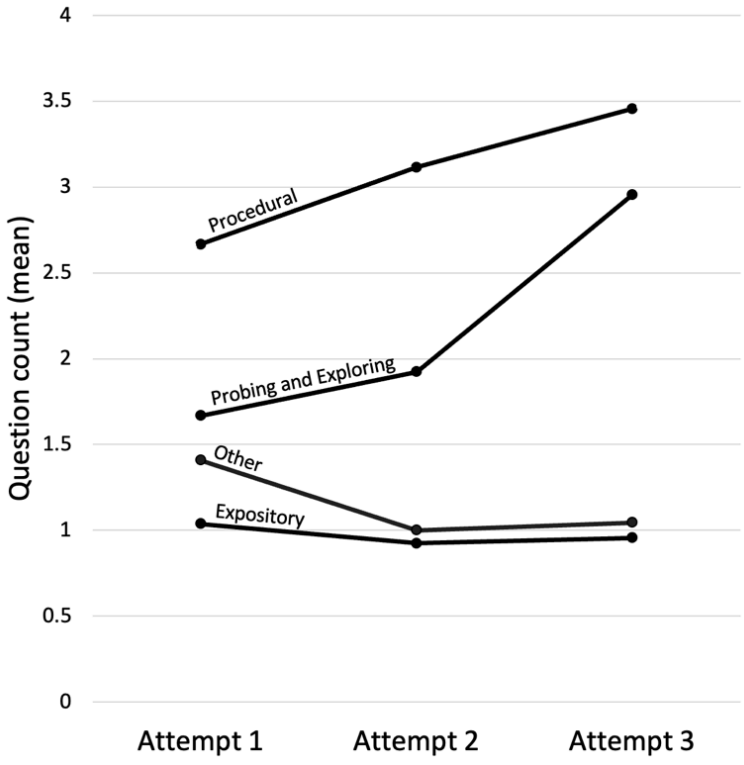
Table 2

The mean and standard deviation of the counts of each type of question asked at each attempt

Question Type	**Count Mean (SD)**		
	Attempt 1	**Attempt 2**	**Attempt 3**
Probing and Exploring	1.667 (1.617)	1.923 (1.671)	2.955 (1.864)
Procedural	2.667 (1.981)	3.115 (2.747)	3.455 (2.540)
Expository	1.037 (1.055)	0.923 (0.891)	0.955 (0.844)
Other	1.407 (1.047)	1.000 (1.058)	1.045 (0.722)
n	27	26	22

Figure 4

The mean count of each type of question asked at each attempt



These results show increases in the probing and exploring and procedural categories and small decreases in the expository and other categories. The results from testing the slope of each question, and whether this slope is statistically significant, is then shown in Table 3.

Table 3

Test results for the slope for each question type

Question Type	Slope	Std Error	df	t-value	p-value
Probing and Exploring	0.620	0.236	47	2.625	0.012*
Procedural	0.310	0.203	47	1.523	0.135
Expository	-0.066	0.115	47	-0.570	0.572
Other	-0.200	0.118	47	-1.692	0.097

*p < 0.05

The growth curve model found a statistically significant ($p < 0.05$) increase in the number of probing and exploring questions asked. No statistically significant change was found in the remaining three categories (Table 3).

Discussion

Results indicate that detailed and high quality automated post-rehearsal feedback had a statistically significant increase in the quantity of probing and exploring questions the pre-service teachers in this study asked. Despite limitations such as the small sample size and exploratory case study design, this study suggests such feedback can play an important role in the design of simulation tools for developing teaching practices. These findings are in line with prior research which supports the importance of providing immediate and actionable feedback (Pianta & Hamre, 2009) to pre-service teachers to develop their questioning skills (e.g., Grossman et al., 2009; McGarr, 2021) in mathematical contexts similar to those that pre-service teachers will face in their future classrooms (Cohen et al., 2020).

As a preliminary step to further investigation, implications of this study include the importance of supporting pre-service teachers to engage in simulation tools through their teacher education programs and working with instructors of pre-service teachers to design and implement simulation tools that fit the needs of their courses. From this exploratory study, future research will include expanding the sample size to compare the impact of ACTS feedback on preservice teachers' mathematical questioning across different class sections and types of mathematical activities as well as qualitative research considering how pre-service teachers understand the impact of the ACTS feedback on their questioning. This work is important to ensure that pre-service teachers are supported to develop the questioning skills for their future classrooms that will provide opportunities to K-12 students that support both their mathematical understanding and their mathematical identities.

References

Aguirre, J., Mayfield-Ingram, K., & Martin, D. (2013). *The impact of identity in K-8 mathematics: Rethinking equity-based practices*. The National Council of Teachers of Mathematics.

Boaler, J., & Brodie, K. (2004). The importance, nature and impact of teacher questions. In D. E. McDougall & J. A. Ross (Eds.), *Proceedings of the 26th Conference of the Psychology of Mathematics Education* (pp. 773-781). Toronto: OISE/UT.

Boston, M., Bostic, J., Lesseig, K., & Sherman, M. (2015). A comparison of mathematics classroom observation protocols. *Mathematics Teacher Educator, 3*(2), 154-175. https://doi.org/10.5951/mathteaceduc.3.2.0154

Boston, M. D., & Candela, A. G. (2018). The Instructional Quality Assessment as a tool for reflecting on instructional practice. ZDM: International Journal of Mathematics Education, 50(3), 427-444.

Cohen, J., Wong, V., Krishnamachari, A., & Berlin, R. (2020). Teacher coaching in a simulated environment. *Educational evaluation and policy analysis, 42*(2), 208-231. https://doi.org/10.3102/0162373720906217

Datta, D., Bywater, J. P., Phillips, M., Lilly, S., Chiu, J. L., Watson, G. S., & Brown, D. E. (2023). Classifying mathematics teacher questions to support mathematical discourse. In *International Conference on Artificial Intelligence in Education* (pp. 372-377). Cham: Springer Nature Switzerland.

Forzani, F. M. (2014). Understanding “core practices” and “practice-based” teacher education: Learning from the past. *Journal of Teacher Education, 65*(4), 357-368. https://doi.org/10.1177/0022487114533800

Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers and Teaching: Theory and Practice, 15*(2), 273-289. https://doi.org/10.1080/13540600902875340

Jacobs, V. R., Lamb, L. L., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. Journal for research in mathematics education, 41(2), 169-202.

Mikeska, J. N., Howell, H., & Kinsey, D. (2024). Teacher Educators’ Use of Formative Feedback During Preservice Teachers’ Simulated Teaching Experiences in Mathematics and Science. International Journal of Science and Mathematics Education, 1-29.

McGarr, O. (2021). The use of virtual simulations in teacher education to develop pre-service teachers’ behaviour and classroom management skills: implications for reflective practice. Journal of Education for Teaching, 47(2), 274-286.

National Council of Teachers of Mathematics (2014). Principles to actions: Ensuring mathematical success for all. Reston, VA: National Council of Teachers of Mathematics.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. Educational researcher, 38(2), 109-119.

Shaughnessy, M., & Boerst, T. (2018). Designing simulations to learn about pre-service teachers’ capabilities with eliciting and interpreting student thinking. In Research advances in the mathematical education of pre-service elementary teachers (pp. 125-140). Springer, Cham.

Sherin, M. G. (2002). A balancing act: Developing a discourse community in a mathematics classroom. Journal of mathematics teacher education, 5, 205-233.

Yin, R.K. (2018). Case study research and applications: Design and methods, 6th Edition. SAGE.

Acknowledgements

This material is based upon work supported in part by 4-VA, a collaborative partnership for advancing the Commonwealth of Virginia, by the Robertson Foundation, and by the National Science Foundation under Grant Nos. 2315436 and 2315437. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of 4-VA, the Robertson Foundation, or the National Science Foundation.