Abstract Title Page

Title: A Randomized Field Study of a Teacher Professional Development Program Using Mixed-Reality Simulation to Develop Instructional Practice

Authors and Affiliations:
Rachel Garrett (co-presenter), American Institutes for Research, rgarrett@air.org
Toni Smith (co-presenter), American Institutes for Research, tsmith@air.org
Melinda Griffin, American Institutes for Research, mgriffin@air.org
Melissa Yisak, American Institutes for Research, myisak@air.org

Abstract Body

Context:
Group discussions during mathematics lessons provide important opportunities for students to deeply engage with mathematics content and to develop communication skills and strong understanding aligned with college and career readiness objectives (Walshaw & Anthony, 2008). However, such student-centered discussions are less common than teacher-led lectures in middle grades mathematics (Stigler & Hiebert, 2004). Research has found that lack of training contributes to the challenges in successfully implementing student-driven discourse (Piccolo et al., 2008).

One innovative way for teachers to develop their instructional practices is the Mursion/TeachLivE mixed-reality classroom simulator. Teachers use a computer to connect with a simulated classroom via an online video conference platform to interact with a set of student avatars (see Appendix Exhibit 1). Middle and high school mathematics and science teachers using the simulated classroom have reported they suspended disbelief and experienced an authentic teaching experience while encountering the challenges of an actual classroom (Becht & Delisio, 2015; Straub, Dieker, Hynes, & Hughes, 2014, 2015). This opportunity to rehearse teaching practices in an authentic but low-stake environment may help teachers learn how to deliver complicated instruction (Lampert et al., 2013). However the field lacks rigorous research on impacts on classroom practice.

Purpose:
Simulated Instruction in Mathematics Professional Development—or SIM PD—is a new program being developed to meet the need for PD focused on mathematical questioning and discourse (NCTM, 2014) using the classroom simulation technology.

This presentation will share findings from a small-scale, school-level randomization study in 2018-19 to provide an early, rigorous assessment of whether the PD helps teachers develop the targeted instructional practices.

1) Can the SIM PD activities be completed as intended?
2) What are teacher perceptions of the usefulness of SIM PD?
3) What is the impact of SIM PD on classroom practice?

Program Description:
SIM PD aims to help teachers providing mathematics instruction in grades 4–7 support student learning through effective questioning and discourse practices. The study team trained school or district staff to serve as coaches that worked with school-based teacher teams. The 10-hour PD entailed two modules that each followed the same structure (see Appendix Exhibit 2): (a) a workshop to learn about the focal instructional practices and time for planning to implement in the simulated classroom; (b) a group simulation session during which teachers, with their coach, worked as a team to develop skills in implementing the practices; (c) an individual simulation session with their coach to continue refining their skills; and (d) a professional learning community (PLC) meeting to debrief and consider application to teachers' real-life classrooms.

Research Approach:
This study used a school-level random assignment design so that groups of two to six teachers within a school were randomly assigned to either implement the program during the 2018-19 school year, or participate in data collection but wait to participate in the PD until the 2019-20 school year.

To monitor and support implementation, the development team collected video recordings of the PD activities, online coaching logs completed after each PD activity, and teacher reflections completed individually after their simulation activities but before their group PLC.

To assess the impact of the program on instructional practice, the evaluation team collected video classroom observations of teachers' actual classrooms, conducted before the PD began, immediately after the PD ended in the winter of 2019, at the end of the school year in spring 2019, and once more at the start of the 2019-20 school year. The video observations were coded with 13 measures either directly taken or adapted from the Mathematics Scan (M-Scan) observation rubric (Berry et al., 2013). The videos were coded by someone external to the development team, blind to teachers' treatment condition.

Findings:
Of 51 teachers in the initial randomization sample, 44 remained in the analysis sample (see Appendix Exhibit 3). This overall attrition of 14% and group differential attrition of 4% aligns with meeting the What Works Clearinghouse "cautious" attrition assumptions for a randomization study (WWC, 2017.).

Overall the program was implemented as intended. For example, coach reports indicated that on average activities lasted within 10 minutes of their intended length of time, with the exception of the first PLC which averaged 12 minutes shorter than planned and the second group simulation session that averaged 19 minutes longer (see Appendix Exhibit 4). Coaches reported completing 95% or more of the planned segments within each PD activity (see Appendix Exhibit 5). Nineteen of the 21 teachers in the analysis sample completed all 8 PD activities.

Analysis of teacher reflections (not presented here) indicated that the teachers felt the program focused on important instructional practices for their work. Many teachers struggled initially with the technology, but almost all become accustomed to the simulated classroom and found it useful for practicing the focal strategies. All the teachers reported that they were trying to use the PD's focal strategies in their actual classrooms.

Preliminary results based on the classroom observations of all study teachers immediately following the PD (winter of 2019) indicate that the program had a positive effect on instructional practice. Using an overall measure of instructional practice aggregated across individual coding scores, preliminary results suggest an effect size of 0.97 SD in favor of the PD group (see Appendix Exhibit 6). Examining individual observation codes suggests that the areas of classroom discourse and explanation and justification of mathematical thinking were most consistently positively impacted by the PD. Finalized results for the immediate impacts and for the follow-up classroom observations will be available by the conference presentation.

Conclusions:

Our findings indicate that the program was feasible to implement, teachers received the PD positively, and classroom practice may have improved as a result of the PD. While final analyses are ongoing, these early signs of promise suggest that the program offered an effective use of the classroom simulation technology to support development of classroom discourse strategies. Next steps will include finalization of results, and continued work to refine the PD and study its implementation and effectiveness at greater scale.

References

Becht, K., & Delisio, L. (2015, June). The lived experiences of middle school and high school teachers utilizing TeachLivE as a professional development tool. In Bousfield, T., Hynes, M., Hughes, C., Straub, C., Dieker, L. & Ingrahm, K. (Eds.), *Proceedings of 3rd National TLE TeachLivE™ Conference 2015: Dissecting Education* (pp.2–6). Orlando, FL: University of Central Florida.

Berry, R. Q., Rimm-Kaufman, S. E., Ottmar, E. M., Walkowiak, T. A., Merritt, E., & Pinter, H. H. (2013). *The Mathematics Scan (M-Scan): A measure of standards-based mathematics teaching practices (Unpublished measure)*. Charlottesville, VA: University of Virginia.

Lampert, M., Franke, M.L., Kazemi, E., Ghousseini, H., Turrou, A.C., Beasley, H., Cunrard, A. & Crowe, K. (2013.) *Journal of Teacher Education, 64*(3), 226-243.

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

Piccolo, D. L., Harbaugh, A. P., Carter, T. A., Capraro, M. M., & Capraro, R. M. (2008). Quality of instruction: Examining discourse in middle school mathematics instruction. *Journal of Advanced Academics, 19,* 376–410.

Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership, 61*(5), 12–16.

Straub, C., Dieker, L., Hynes, M., & Hughes, C. (2014). *Using virtual rehearsal in TLE TeachLivE™ mixed reality classroom simulator to determine the effects on the performance of mathematics teachers. 2014 TeachLivE National Research Project: Year 1 Findings.* Orlando, FL: University of Central Florida.

Straub, C., Dieker, L., Hynes, M., & Hughes, C. (2015). *Using virtual rehearsal in TLE TeachLivE™ mixed reality classroom simulator to determine the effects on the performance of science teachers: A follow-up study (Year 2). 2015 TeachLivE National Research Project: Year 2 Findings.* Orlando, FL: University of Central Florida.

Walshaw, M., & Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Review of Educational Research, 78*(3), 516–551.

What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education. (2017, October). *What Works Clearinghouse: Standards Handbook (Version 4.0)*. Retrieved from http://whatworks.ed.gov

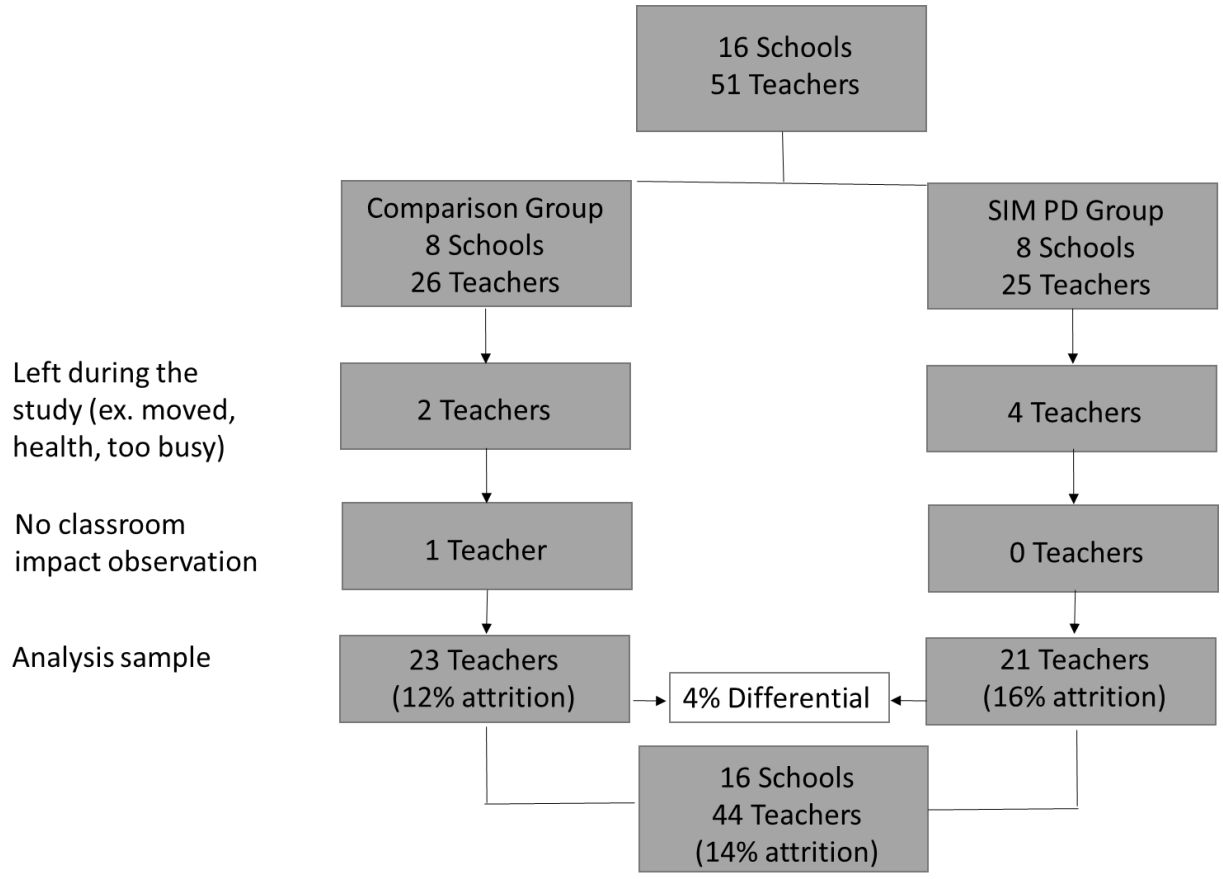
Appendix

Exhibit 1. A Teacher Practicing With the Simulator and an Interactor Operating the Avatars

Exhibit 2. Description of SIM Activities in Each Module

Activity	Description
Workshop (2 hours)	• Introduction to the instructional practice that is the focus of the module • Preparation for work in the simulated classroom
Group Simulation (1.5 hours for Module 1; 1 hour for Module 2)	• [Module 1] Introduction to the avatars • Collaborative work to implement the instructional practice that is the focus of the module: – Teachers take turns implementing the practice – The coach pauses the simulation frequently to encourage reflection
Individual Simulation (1 hour)	• Individual work to implement the instructional practice that is the focus of the module: – The coach regularly pauses activity for teacher reflection
Professional Learning Community (PLC) Meeting (45 minutes)	• Discussion of learning during the module and consideration for how the practice can be implemented in a real classroom.

Exhibit 3. School and Teacher Sample Over the Course of the Study



Note: Calculations based on study records.

Exhibit 4. Planned and Actual Program Activity Length in Minutes

Program Activity	Planned Time	Actual Time Module 1	Actual Time Module 2
Workshop	120	127.5	115.7
Group Simulation	90 for Module 1 60 for Module 2	99.4	79
Individual Simulation	60	63.1	63
PLC	45	33	44

Note: Calculations based on coaching log reports.

Exhibit 5. Percentage of Activity Segments Completed, by Program Activity and Module

Activity	Module 1		Module 2	
	Number of Segments per Activity	Percentage of Segments Completed Across Activities	Number of Segments per Activity	Percentage of Segments Completed Across Activities
Workshop	7	98%	7	100%
Group Simulation	6	98%	5	95%
Individual Simulation	5	96%	5	97%
PLC	4	100%	4	100%

Note: The percentage of segments completed includes all segments across all instances of an activity for implementation. For example, among eight group workshops for Module 1 that each included seven segments, 98% of the 8 x 7 = 56 intended segments were completed. Calculations based on coaching log reports.

Exhibit 6. Preliminary Estimates of SIM PD Impacts on Classroom Practice Observation Scores

	Preliminary Impact Estimates	
	B	SE
Overall Score		
	0.97***	(0.23)
Sub Scores, by Domain		
Use of Representations		
Teacher translation among representations	0.60*	(0.32)
Student translation among representations	1.42***	(0.39)
Cognitive Demand		
Teacher encourages conceptual connections	0.35	(0.34)
Mathematical Discourse Community		
Teacher's role in discourse	2.49***	(0.43)
Mathematics community through student talk	1.19***	(0.33)
Teacher questions	0.69**	(0.27)
Explanation and Justification		
Presence of explanation and justification	1.70***	(0.30)
Depth of explanation and justification	0.84***	(0.27)
Problem Solving		
Student engagement with the math	-0.14	(0.29)
Students grapple with the math	0.51**	(0.21)

Note: Each row represents a separate regression. Each regression contains 44 teachers (21 treatment and 23 control) in 16 schools. Classroom observation scores are based on codes from the MSCAN observation rubric (Berry et al., 2013), with observation scores standardized within time point based on the control group distribution. Program impacts estimated using two level hierarchical models with teachers nested within schools. All models control for a baseline measure of the dependent variable, teacher years of experience, elementary grade indicator, and district randomization blocks.

*** p<0.01, ** p<0.05, * p<0.1