



Harnessing Asynchronous Digital Simulations of Problem-based Lessons to Support Mathematics Teachers' Professional Development: A Design-based Approach

Gil Schwartz^{1,2} · Patricio Herbst¹ · Amanda M. Brown¹

Received: 6 March 2024 / Accepted: 6 October 2024
© The Author(s) 2024

Abstract

The complexity of mathematics teaching is especially evident in lessons where teachers build on students' genuine ideas, such as problem-based lessons. To enhance teachers' capacity for rich discussions in problem-based instruction, we have developed a unique approximation of practice: digital asynchronous simulations where teachers make subject-specific decisions for a virtual teacher avatar. The simulations are based on materials and principles from a practice-based professional development (PD) program, implemented with small groups of teachers. The self-paced simulation model offers flexibility and scalability, allowing more teachers to participate on their own schedules, but it lacks key affordances of collaborative PD. To examine how to leverage the affordances of collaborative, practice-based PD, this paper uses a design-based research approach to explicate the mechanisms in which digital simulations can support mathematics teachers' learning about problem-based lessons. We focus on two cycles of design, implementation, analysis, and revisions of the simulation model, drawing on data from focus groups with mathematics teacher educators, prospective teachers' performance, and teachers' reflective assignments. The analysis illustrates how two design principles – *Authenticity to the teacher's work*, and *Nuanced feedback* – were transformed to better reflect aspects of practice-based teacher learning. We argue that self-paced, asynchronous simulations with indirect feedback can effectively emulate aspects of collaborative, practice-based PD in supporting teachers' growth. The paper also contributes to the literature on mathematics teachers' noticing and decision-making, examining how the two interact in simulated environments. We suggest implications for designing practice-based asynchronous digital simulations, drawing on emerging technologies.

Extended author information available on the last page of the article

Published online: 29 October 2024

Springer

Keywords Digital simulations · Approximations of practice · Practice-based teacher learning · Problem-based instruction · Online professional development · Asynchronous professional development · Design-based research · Noticing · Responsiveness · Decision-making · Selecting · Sequencing · Framing · Student work · Geometry

The complexity of mathematics teaching is especially evident in lessons where teachers build on students' genuine ideas to develop new knowledge, such as in problem-based lessons. To enhance novice teachers' capacity for problem-based instruction, we have developed simulations where teachers make decisions for a virtual teacher avatar. While this design offers flexibility and scalability, it lacks some of the main affordances embedded in the design of collaborative practice-based teacher education environments.

Practice-Based Teacher Education (PBTE) is rooted in the notion that teachers should “learn in and from practice” (Cohen & Ball, 1999, p. 18). Unlike traditional teacher preparation and professional development (PD) programs, which predominantly focus on theoretical knowledge, PBTE enables teachers to apply theories and strategies by engaging in activities of demonstration/representation, approximation, and decomposition of practice (Grossman et al., 2009; Herbst et al., 2016).

Approximations of practice refer to teachers' enactment of practice in settings of reduced complexity. This definition includes script writing (Herbst & Milewski, 2018; Osmanoglu & Girit-Yildiz, 2024; Zazkis, 2017), role-playing activities (e.g., rehearsals, Lampert et al., 2013), and live simulations with actors (Shaughnessy & Boerst, 2018) as well as digital simulations with avatars. Digital simulations are either controlled by actors (e.g., in Mursion; Dieker et al., 2013; Mikeska et al., 2023), performed by a group of teachers (e.g., in Virtual Teaching Simulator; Park Rogers et al., 2024), or completely autonomous (e.g., in SimSchool; Gibson, 2007; and Teacher Moments; Reich, 2022; see also Herbst et al., 2020, 2022).

Approximations of practice help teachers develop their skills in controlled and carefully designed environments before applying them in actual classroom settings. This approach provides educators with opportunities for hands-on learning as well as reflection. Importantly, approximations of practice are distinct from real experiences, allowing for focused and repeated refinement of specific teaching subcomponents (Davis et al., 2017; Milewski et al., 2018).

While live simulations have been a viable professional tool for decades (e.g., role plays and microteaching; Sahu, 1984), various types of digital simulations have emerged in recent years. This growth is part of the expansion in online professional learning (see below), driven by technological advances and pandemic-induced changes in educational systems. In contrast to live simulations, digital simulations offer flexibility, scalability, sustainability, customization, and built-in analytic tools that enable instructors to collect data on learner performance and interactions, including tracking learners' progress over time (Sweeney et al., 2018).

Due to their digital and self-paced nature, these simulations can be integrated into PBTE either in a hybrid model (e.g., incorporated in a methods course; see Mikeska et al., 2023; Thompson et al., 2019) or as standalone professional learning opportunities, potentially available to larger groups. When simulations are designed to be completed individually, without peer or mathematics teacher educator (MTE) support, it is crucial to consider the feedback mechanisms they provide and how these mechanisms support teacher learning. Designers must decide whether to implement a normative (prescriptive) model, a naturalistic (descriptive) model, or balance the approaches (Scherpereel, 2015). This tension is similar to that in Massive Open Online Course (MOOC) design, where reconciling learners' experimentation with validation of their actions is critical for supporting learning and improving retention (Brennan et al., 2018).

Building on the work of Brousseau (1997), we use the term *didactic feedback* to describe normative simulation models that provide direct feedback on teacher performance, assessing their actions against specific criteria that prescribe how good teaching looks like. In contrast, we use the term *adidactic feedback* to describe models that provide indirect or "soft" feedback, reflecting rather than verifying responses, and facilitating experimentation with diverse teaching approaches. This feedback may involve visualizing teachers' choices in simulations, or prompting their noticing and reflection without direct evaluation. For example, in TeachLive/Mursion (Dieker et al., 2013), a remote operator—known as Sim-Expert—controls avatars that respond to user decisions, providing adidactic feedback to help teachers relate between actions and implications.

Our approach for utilizing simulations in PBTE follows the latter type. In this paper, we investigate an on-demand, asynchronous, adidactic simulation model, which is an adaptation of the facilitated practice-based professional learning model *StoryCircles* (Herbst & Milewski, 2018) where teachers collectively design a problem-based lesson by creating a storyboard. In *StoryCircles*, peer-interaction about mathematics instruction is an essential resource for participants' learning (Brown & Herbst, 2023; Schwarts et al., 2022, 2023). Therefore, this adaptation has led us to explore the following overarching question, guiding our research project: In the absence of peers, MTE, or *didactic* feedback, what mechanisms can practice-based digital simulations rely on to promote mathematics teachers' learning in and from practice?

The case we discuss includes a set of simulations of a problem-based geometry lesson, wherein secondary mathematics teachers make subject-specific decisions (e.g., selecting and sequencing student work for a whole-class discussion) on behalf of an avatar teacher. This paper investigates how these asynchronous simulations can harness some of the affordances of practice-based teacher collaboration. We employ a design-based research approach (as outlined in Bakker, 2018), studying two cycles of design, implementation, analysis, and revisions by using conjecture mapping (Sandoval, 2014). Our goal is to share insights into developing sustainable, scalable practice-based asynchronous digital simulations that embrace diverse teaching approaches without being prescriptive. Beyond

discussing design, we also present positive results of teachers' performance and reactions to these simulations. Our central questions are:

1. How can asynchronous didactic practice-based simulations support mathematics teachers' opportunities to learn about problem-based instruction?
2. What resources and mechanisms can be incorporated into such simulations to emulate aspects of collaborative practice-based teacher education?

By examining these questions, we contribute to discussions about the design and viability of online, asynchronous, practice-based PD models. We suggest ways in which familiar mechanisms of teacher learning can be adapted into new contexts, as well as identify novel mechanisms that are afforded by emerging technologies.

Literature Review: Situated Student and Teacher Learning

So far, we highlighted PBTE as a framework to guide our design and research of teacher learning. We now aim at grounding our work in a related, broader theory of learning in context, known as situated learning (Lave & Wenger, 1991), emphasizing how students and teachers learn through authentic, contextualized activities, often within a community of practice. Situated learning recognizes that much of what students and teachers learn is acquired through experiences and often remains tacit and context-specific. It follows that tacit knowledge (Herbst & Chazan, 2011) can be elicited during hands-on, practice-related activities, such as simulations of teaching.

Below, we review the relevant literature for our study, elaborating on how the principles of situated learning are reflected in students' problem-based learning and teachers' collaborative learning. We conclude with describing teachers' learning in digital spaces.

Problem-Based Learning

Situated learning involves active participation in meaningful and authentic activities, emphasizing the importance of social interactions in the learning process. Problem-based instruction applies these principles by offering learners opportunities to engage collaboratively in analytical thinking, reasoning, interpretation of quantitative data, and sense-making (Santos-Trigo, 2024). Solving complex, novel problems could mirror the work of mathematicians (Koichu, 2014), providing a context for students to develop and apply their knowledge. Endorsed by various reform documents (e.g., National Council of Teachers of Mathematics [NCTM], 2014; Organization for Economic Cooperation and Development [OECD], 2022), problem-based learning is recognized as an effective strategy to promote students' productive struggle and participation in explorative mathematical discourse.

Discussion-rich, problem-based lessons typically begin with students grappling with a novel, challenging problem either individually or in small groups, followed by a whole-class discussion where students publicly present their work which the teacher weaves together to reach their instructional goals¹. However, teaching in ways that build on student ideas is considered a challenging endeavor (Kooloos et al., 2022). To tackle this issue, Stein and colleagues (2008) recommend five practices for orchestrating productive mathematical discussions, the first four of them precede the whole-class discussion: Anticipating, monitoring, selecting, sequencing, and connecting student work. They maintain that teachers should select and sequence students' contributions to create "mathematically coherent storylines" (ibid., p. 44). Emphasizing pre-discussion practices, rather than solely focusing on talk moves, highlights the importance of teachers' noticing of student mathematical thinking and their related decision-making. It conveys the message that teachers should plan how to use students' ideas to reach the lesson goal, rather than using ad-hoc talk moves to manage what students happen to contribute when the discussion begins. To enhance teachers' capacity to implement such an approach, numerous PD initiatives have been established, specifically targeting support for design and implementation of problem-based instruction.

Collaborative Teacher Learning

Collaborative environments for mathematics teacher learning draw on the principles of situated learning and PBTE, offering structured opportunities for reflective practice. A recent ICMI Study addressed this topic (Borko & Potari, 2024), with its Discussion Document (International Program Committee, 2019) stating that: "Collaborative work of teachers has a long tradition in mathematics education as it is critical as a way to bring educational innovation into the everyday practice of teaching." (p. 2). Notably, the document highlights a significant limitation in this body of knowledge: the lack of research on "the relationship between [teacher] learning and collaboration" (ibid).

Existing research suggests that collaboration promotes reflection and critical inquiry, encouraging teachers to question assumptions, explore innovative approaches, and refine their teaching practices based on peer-discussions and feedback (Vangrieken et al., 2015). Furthermore, collaboration cultivates a sense of community among teachers, fostering supportive networks and relationships that can sustain and invigorate professional growth (Martin & Gobstein, 2015). These goals are achieved through carefully designed structures that support collaboration, coupled with skillful facilitation. The role of the MTE or facilitator in such contexts is to create a supportive and inclusive atmosphere and lead discussions that amplify teachers' voices and promote reflection (Schwartz et al., 2021, 2022).

Lesson-based PDs, which center on the design of a single lesson, offer a compelling approach for supporting problem-based instructions in collaborative

¹ Importantly, our focus is on lessons where problems are used to learn knowledge items which are curricular goals, rather than to learn problem-solving heuristics or to practice what has been learned before.

contexts. A notable model is Lesson Study, where groups of teachers engage in cycles of designing, implementing, and reflecting on a lesson to promote collective learning and professional discourse. Another model is *StoryCircles*, a lesson-based collaborative PD model inspired by Lesson Study. In *StoryCircles*, secondary mathematics teachers anticipate a lesson through iterative phases of scripting, visualizing, and arguing about it (Herbst & Milewski, 2018) in both synchronous and asynchronous online activities. Each *StoryCircles* cycle lasts several weeks and focuses on participants' attempts to improve one lesson, initially sketched in a storyboard, as they see fit.

The goal of *StoryCircles* is to foster teachers' peer-argumentation about practice (and not, for example, to direct them to teach a specific lesson), eliciting their tacit professional knowledge to make it shareable. This is achieved through engagement with samples of student work, scripting of teaching moves, consideration of potential contingencies that can emerge in the context of classroom implementation (Brown et al., 2021), and a constant examination of how instructional decisions align with a specific curricular goal. While we are currently scaling *StoryCircles* internationally through partnerships with MTEs, we also aim at making this experience available to a larger number of mathematics teachers without making so high demands on human infrastructure (peers and MTEs) and while respecting teachers' busy schedules. Our effort aligns with current trends in the field, in the post-COVID era, toward developing practice-based forms of online, asynchronous PD.

Teacher Learning in Digital Spaces

Online professional development has been available for a long time, but recent years have seen a surge in digital tools, spaces, and applications designed to support teacher learning. Paraphrasing Shakespeare, Dede (2022) notes that the pandemic has created a “sea change rather than a temporary discontinuity” (p. 117) in education, driven by broader social forces benefiting from hybridity. He observes that “our brave new world is rich and strange compared to pre-pandemic life” (p. 118), highlighting the abundance of new possibilities in online settings. Furthermore, the ongoing generative Artificial Intelligence (AI) revolution is transforming the educational landscape, offering novel pedagogies for online asynchronous professional learning. For instance, AI-based chatbots now serve as “facilitators” (Copur-Gencurk & Orrill, 2023), thought partners during lessons (Shin, 2022), or simulate students to enhance teachers' responsiveness (Son et al., 2024).

One example of the richness described by Dede (2022) is the rise of mixed-reality virtual field experiences (Bondurant & Amidon, 2021), addressing the need to train prospective teachers (PTs) in challenging conditions, for example – when no actual field experience is available. Emerging technologies such as virtual reality (VR) and augmented reality (AR) immerse future teachers in simulated classroom experiences, offering a realistic and controlled environment for practice (e.g., Ferdig et al., 2022). To date, using most of these models is challenging due to their high costs and reliance on specialized equipment, making them difficult to scale sustainably.

In contrast, a scalable and sustainable asynchronous model is online courses with modules for specific content or practices (e.g., Friesen et al., 2023; Lee et al., 2021). Compared to traditional PD, such courses offer flexibility and different degrees of adaptivity and interactivity. Still, their capacity to provide experiential learning opportunities for teachers varies.

Digital simulations, which combine scalability with professional experimentation, offer a unique model for supporting teacher learning. Typologies of simulations in other fields, such as medicine (Meller, 1997), highlight key differences between simulation models, including forms and levels of feedback, scope, length, authenticity, and interactivity. For example, simulations can range from short, specific scenarios (e.g., Reich, 2022) to entire lessons or units (see our model below). While approximations of practice are by definition not fully authentic (Grossman et al., 2009), different dimensions of inauthenticity have varying implications on participants' experiences, and authenticity may be perceived "as more a malleable attribute than as simulation's end goal" (Howell & Mikeska, 2021, p. 8). In terms of interactivity, Meller (1997) highlights the importance of passive elements in the simulation: "Each element of the simulator can be either passive, active, or interactive. A passive element usually is provided to enhance the setting or 'realism' of the simulator" (p. 194). In summary, simulations aim to imitate certain aspects of teachers' work, rather than providing the most genuine experience.

While simulations present many affordances, they remain relatively uncommon in the realm of professional learning for mathematics teachers. Moreover, most digital simulations currently focus on generic aspects like classroom management (Lindberg & Jönsson, 2023). This may be due to the difficulty of programming virtual students to respond mathematically in ways that support teacher learning, though some pioneer efforts have begun (Son et al., 2024). As the field evolves, enhancing subject-specific focus within digital simulations is essential for STEM educators (Mikeska et al., 2021). Our study addresses this need, guided by the theoretical framework presented below.

Theoretical Framework: Teachers' Subject-Specific Noticing and Decision-Making

The literature on teachers' professional noticing in mathematics education has expanded significantly in recent years, drawing upon foundational works such as Goodwin's (1994) concept of *professional vision* and Mason's (2002) ideas about *attention* and *awareness*. A large body of research examines how teachers perceive, categorize, and respond to students' mathematical thinking (Santagata et al., 2021). There is an increasing advocacy for a nuanced understanding of noticing that recognizes its situated nature (Blömeke et al., 2015), acknowledging that teachers' observations are shaped by their framing of their role and the instructional context (Louie et al., 2021; Scheiner, 2021, 2023).

Furthermore, following the work of Jacobs et al. (2010), a growing number of studies aimed at linking noticing with decision-making (e.g., Blömeke et al., 2015; Estapa & Davis, 2023; Kaiser et al., 2017; Rotem & Ayalon, 2024; Tekin-Sitrava

et al., 2024). In this vein, van Es and Sherin (2021) expanded their prior *Learning to Notice* framework (van Es & Sherin, 2002), which includes the dimensions *attending* and *interpreting*, to include also *shaping* – the intentional creation of interactions that enhance opportunities for interpreting mathematical thinking. Moreover, they suggest that *attending* is not only about identifying “noteworthy features of classroom interactions” (van Es & Sherin, 2021, p. 19) but also about deliberately disregarding less consequential classroom aspects. This expanded framework underscores the idea that noticing is an active process involving deliberate attention to specific classroom features.

Engaging in these discussions, Herbst et al. (2023) connect *Practical Rationality* (Chazan et al., 2016; Herbst & Chazan, 2011) and *teachers’ noticing* by describing subject-specific categories of perception that mathematics teachers use in problem-based lessons. This connection emphasizes the particular knowledge-at-stake in a lesson (Herbst, 2006) and what the practice of mathematics instruction makes available for the teacher to notice. The suggested categories of perception, to be presented below, build on two interrelated concepts from Practical Rationality: *instructional situations* and *instructional norms* (Herbst & Chazan, 2011), both can be considered as components of teachers’ tacit knowledge.

Instructional situations are recurring types of tasks, in mathematics lessons, characterized by subject-specific situational *norms*. In the U.S. context, instructional situations in geometry include *proof*, *construction*, and *calculation*, each of which has its own norms. Instructional norms are the implicit expectations about who should do what for work on a task to be interpreted in relation to the knowledge-at-stake (e.g., in the situation of *construction*, there is an expectation that students use construction tools to create new geometric objects). When teachers introduce novel problems, they can frame a problem explicitly within an instructional situation (“please construct a circle”, evoking the situation of *construction*) or leave it open-ended (“find a circle”). These framings can elicit different expectations regarding students’ contributions.

Herbst et al. (2023) propose that the norms of each instructional situation, along with the lesson’s goal, activate teachers’ categories of perception when they notice and make decisions about student ideas in problem-based lessons. *Normativity* is a category of perception that refers to the affordance for the teacher to perceive how well a student’s contribution aligns with the norms of the instructional situation used to frame a problem. *Serviceability* is a category of perception that refers to the affordance for the teacher to perceive how useful a student’s contribution is to achieving the instructional goal of the lesson (see Fig. 1). For instance, if a problem is framed as a construction or invokes this situation, teachers may view a hand-sketched diagram as less normative than one created using construction tools, regardless of the correctness of the proposed answer. However, a sketch that introduces relevant ideas for the lesson goal may be deemed more serviceable, even if it is less normative or incorrect, compared to a construction that does not contribute to achieving the lesson goal. It is important to note that normativity is not synonymous with correctness; normativity refers to adherence to the norms of the instructional situation, which means that a normative student work may be either correct or incorrect (see Fig. 1, top-left corner).

The problem: Suppose that you have two lines that intersect each other and two distinct points, one on each line. How can you find a circle that is tangent to those lines at those points?		
	More normative (instructional situation of <i>construction</i>)	Less normative
More serviceable (includes perpendiculars and /or equidistant segments)		
Less serviceable		
Lesson goal: Learning the Tangent Segments Theorem (Two intersecting lines are tangent to a circle if and only if the points of tangency are equidistant from the point of intersection of the lines).		

Fig. 1 Illustration of the categories of perceptions with respect to the simulated lesson's focal problem, instructional situation, and instructional goal

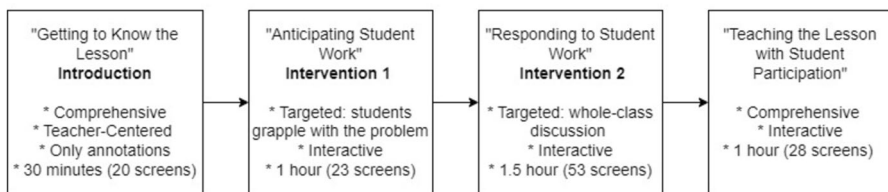


Fig. 2 The initial design of the simulations (Cycle 1)

The categories mentioned above represent mostly tacit teaching knowledge, which can be examined and elicited primarily during actual teaching or its approximations. Below, we demonstrate how these theoretical concepts were integrated into our design of particular kind of approximations of practice: asynchronous teaching simulations.

Design and Methods Overview

This section provides a brief overview of the simulations, followed by descriptions of the methods. A detailed description of the design is presented as part of the Results section, in alignment with the process of conjecture mapping.

We designed four simulations of a lesson centered on a problem (see Fig. 1) that asks students to find a circle tangent to two intersecting lines given the points of tangency. This problem ultimately leads to the *Tangent Segments Theorem* (Two

intersecting lines are tangent to a circle if and only if the points of tangency are equidistant from the point of intersection of the lines). In all simulations, the characters and the settings are represented with cartoon icons and their speech with written speech balloons. The simulations are completely digital and include programmed branches that correspond with the decisions that teachers make.

Two simulations are comprehensive, namely, depict the entire lesson (see Fig. 2): (a) “Getting to know the lesson,” where participants are introduced to how work on the problem could lead to the theorem through a teacher-centered lesson and are tasked with annotating the lesson’s representation, scripting ideas to improve it; and (b) “Teaching the lesson with student participation”, where participants made decisions for a teacher avatar in key moments of the lesson. The remaining two simulations target specific phases in the lesson: (c) “Anticipating student work” focuses on the phase when teachers make sense of students’ attempts to grapple with the problem, which leads to the teacher’s selecting and sequencing of student work²; and (d) “Responding to student work,” centering on various ways to respond to students’ mathematical contributions in whole class discussions.

In terms of interactivity, the “Getting to Know the Lesson” simulation involves participants annotating a lesson representation and solving the mathematical problem, but does not allow them to influence the lesson’s direction. In contrast, the other three simulations require participants to make decisions, with the system responding dynamically through programmed branches, following a ‘choose your own adventure’ style. This includes using second-person language to engage participants (“as *you* look at the student work...”). The design of these simulations evolved through iterative experimentation with teachers and MTEs, as detailed below.

Design-Based Research and Conjecture Mapping

To organize and link the design of the simulations with its theoretical background and the empirical evidence we collected, we use conceptual tools from Design-Based Research (DBR). DBR aims to achieve both practical improvements and theoretical refinements through iterative cycles of design, enactment, analysis, and revision. To provide adequate argumentative grammar (Kelly, 2004) for DBR, *conjecture mapping* helps explicate the links between design, theoretical conjectures, and observed outcomes (Sandoval, 2014).

By using conjecture maps, researchers explicate not only “what works” but also unpack how it works, specifying the theory of action underlying the design. Conjecture maps consist of the following components: (1) *High-Level Conjecture*: a general statement about how an intervention intends to support some form of learning; (2) *Embodiment*: the physical artifacts used (e.g., instruments, software, media), the task structure (what participants are asked to do), and the social structure; (3) *Mediating Processes*: the processes by which design features are expected to activate learning; and (4) *Desired Outcomes*: the expected results of the mediating processes, clearly stated to allow measurement.

² Note that this is a different use of the word “anticipating” compared to the Five Practices terminology.

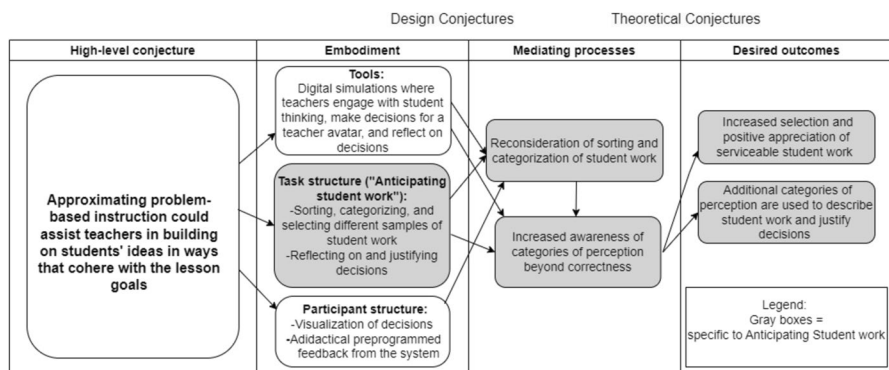


Fig. 3 An initial conjecture map, focused on the “Anticipating Student Work” simulation

The links between Embodiment elements and Mediating Processes are called Design Conjectures, whereas the links between Mediating Processes and Desired Outcomes represent Theoretical Conjectures (see Fig.3). We used conjecture mapping over two implementation cycles to demonstrate how our design refined to better support mathematics teachers’ professional growth.

Data Collection and Analysis

Our analysis is grounded in two cycles conducted in 2023: one with expert MTEs and the other with PTs³. We piloted the simulations with these groups to gain diverse, complementary perspectives. The cycles differed in nature: with the MTEs, we conducted focus groups to discuss design aspects, without examining their simulation performance. This allowed us to explore design conjectures—that is, the interplay between simulation *embodiment* and *mediating processes*. In Cycle 2, we used simulation data and reflective assignments to test theoretical conjectures (how the *mediating processes* relate to the *desired outcomes*). These distinct goals required varied methods for data collection and analysis, detailed below.

Cycle 1: Expert MTEs

Cycle 1 data originates from consultations with ten U.S.-based MTEs, who interacted with the simulations individually (one per week) and attended weekly focus group meetings (two groups of five MTEs each) in early 2023. The MTEs were recruited via emails sent in professional networks. All of them had previous experience with, and knowledge about, problem-based instruction, but they were not familiar with using asynchronous teaching simulations. In total, we conducted ten two-hour focus group sessions with MTEs.

³ Our design aims to apply for both prospective and practicing teachers.

All meetings were recorded and transcribed for analysis. The first step was a thematic analysis following Clarke and Braun (2017), aimed at identifying both problematic and productive aspects of the simulation's design. We coded each session separately, identifying overarching issues across all simulations and specific issues related to particular design elements within each simulation. We also compared the two parallel sessions for each simulation to examine the prevalence and fluctuations of issues between the two focus groups.

In the second round of analysis, we refined the emerging issues according to the conjecture map elements. This step was inspired by the approach described by Russell et al. (2022), where bottom-up themes were used to elucidate or revisit map elements and their interconnections. This process yielded several preliminary categories of design decisions, such as immersiveness versus reflection (the need to decide when to pause engagement in the simulation to promote teacher reflection); ensuring prerequisite knowledge (the need to confirm that teachers understand the mathematics before experimenting teaching it); balancing simulation sections (the need to manage passive, active, and interactive parts); authenticity of teaching decisions (the need to make teaching decisions realistic); authenticity and effectiveness of student work (the need to offer representative and diverse student work samples, that align with the simulation goals); coherence (the need for consistency within each simulation and across the set); technical issues; facilitated vs. non-facilitated use (relates also to the question of what teachers can learn given this is a standalone, self-paced activity); and content and structure suggestions related to the Tangent Circle lesson (how the problem is worded, how much time is given for groupwork, etc.).

We note that (a) some categories were specific to certain simulations, while others were cross-cutting, and (b) discussions on aspects that worked well were backgrounded, in the analysis, to focus on the revision process. Consequently, to answer our research questions, we used the categories to suggest two cross-cutting themes that informed our revisions, related to PBTE principles: *Lack of Authenticity and Context*, and *Lack of Feedback*.

Cycle 2: Prospective Teachers

In Cycle 2, the simulations were integrated into a 5-week period within a methods course at a university in the Northwestern region of the U.S. The simulations were assigned as individual homework tasks during the first half of the course, with data collection focused on PTs' simulation performance and subsequent reflective assignments. We analyzed data from 11 PTs (5 male, 6 female) who completed all simulations and reflections. Our objective in this study differed from that of the focus groups with MTEs: In Cycle 2 we aimed to gather evidence of the simulations' effectiveness in supporting teachers' learning. Therefore, we compared PTs' performance in the comprehensive simulation "Teaching the lesson with student participation" at the beginning of the course (pre-interventions, see Fig. 10) and four weeks later (post-interventions), and we analyzed their post-simulation reflective assignments (average length of 1 paragraph per participant * 6 assignments * 11 participants).

We used mixed methods to analyze changes in PTs' performance between pre- and post-interventions. We scored closed-ended items regarding decisions about student work (e.g., selecting and sequencing) for correctness, normativity, and serviceability (details in the Embodiment section below). These scores were analyzed using a two-tailed Wilcoxon test to determine the significance of observed changes. We analyzed open-ended items, which asked for justifications for decisions, using thematic analysis guided by predefined categories of perception (correctness, normativity, and serviceability). We also conducted a thematic analysis of PTs' reflective assignments to explore how they relate to aspects identified as problematic in Cycle 1 and subsequently revised (e.g., the authenticity of the teacher's work and nuanced feedback), as well as to other elements of the conjecture map.

Integration of Analyses

By examining reactions from two cycles, we gathered multifaceted feedback for our design from two key target communities of the simulations: MTEs (as potential choosers of the simulations as curriculum materials for their students) and PTs (as end-users). MTEs provided insights into the design conjectures, while PTs' reflections and performance data informed our understanding of possible learning outcomes. This dual approach enables a comprehensive articulation and testing of the overall theory of action of the design.

Results

Below, we describe two cycles of design, enactment, analysis, and revisions of the simulations, drawing on empirical evidence. All names used are pseudonyms. Due to space constraints, the evidence shared focuses mostly on the design of two simulations: (a) the targeted *Anticipating Student Work* simulation (intervention 1), where teachers are asked to peruse and classify samples of student work and select some of them for presentation in a whole-class discussion, and (b) the comprehensive *Teaching the Lesson with Student Participation* simulation (specifically the phase in which teachers are asked to monitor, select, and sequence student work when teaching the entire lesson). The comprehensive simulation serves as a context to examine how the targeted simulation shapes teachers' decisions when selecting and sequencing. In accord with this focus, the conjecture map has some elements that are particular to the *Anticipating Student Work* simulation.

Cycle 1: Design Phase

In this section, we lay out the salient elements of the simulations. In Fig. 3, the white boxes represent features relevant for the entire set of simulations, while gray boxes delineate aspects specific to the *Anticipating Student Work* simulation. Herein, we describe the conjecture map categories.

High-Level Conjecture

As mentioned above, we define problem-based lessons as those in which students' work on a novel problem leads to the collective discovery of a curricular goal. In this context, our high-level conjecture is that approximating problem-based instruction could assist teachers in developing skills needed to utilize students' ideas to arrive at the lesson goal.

Embodiment

All digital simulations use cartoon representations of teaching that were created using the Lesson *Depict* software, focusing on the Tangent Circle Problem. Two simulations offer a comprehensive view of the entire lesson, while the other two target specific phases and are considered interventions (see Fig. 2). In the “Getting to Know the Lesson” simulation participants annotate a lesson representation and solve a mathematical problem. This simulation is the only one presented using Anotemos, a media annotation software application where users can view pieces of media together and comment on them using pins, clips, screen markings, and more (Fig. 4). Here, viewers see the entire lesson, where a teacher demonstrates, without any student input, how a consideration of the problem posed leads to the lesson goal. Participants are then asked to use pins to mark moments and then script how the lesson could be more student-centered (see Figs. 4 and 5). The goal of this simulation is to help participants get oriented to the lesson by: first, familiarizing themselves with the arc of the lesson, including grappling with the mathematics, identifying potential solutions and pitfalls, and connecting the problem to the lesson goal (the theorem); second, by annotating a less desirable version of the lesson, they develop the need to improve it.

The other three simulations were developed using the Qualtrics software and include numerous interactive elements. While these simulations have some sections that simply show how the lesson unfolds, they primarily require participants to make

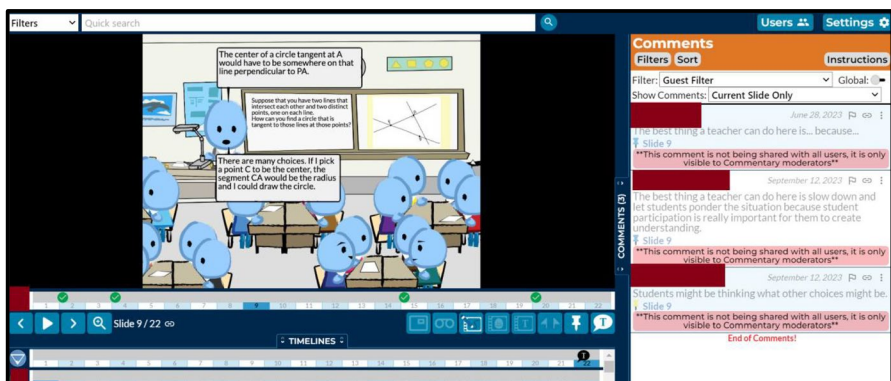


Fig. 4 Screen 9 from the “Getting to Know the Lesson” simulation, with teachers’ pinned comments

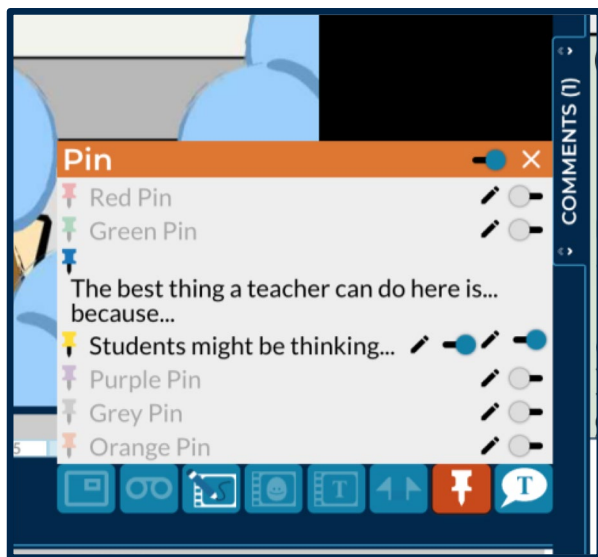


Fig. 5 Pins prompting teachers to script ways in which the lesson can be improved

decisions as if they were teaching the lesson, with the system responding based on programmed branches. Each simulation takes 0.5 to 1.5 hours to complete, and includes 20 to 53 screens (see Fig. 2).

We now delve into the details of the Anticipating Student Work simulation. A key aspect of this simulation is its use of carefully designed samples of student work, showcasing variations in correctness, normativity (alignment with norms of the situation of construction), and serviceability (usefulness for arriving at the Tangent Segments Theorem). Each student work sample was pre-coded with a triplet of "1" or "0" to denote its correctness, normativity, and serviceability. For example, Sigma's work in Fig. 6a was apriori coded [1,0,1] because it is correct,

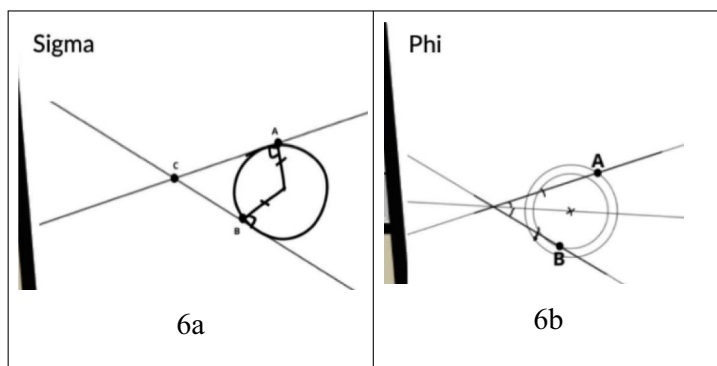


Fig. 6 Samples of student work with various features

non-normative, and serviceable, whereas Phi's work in Fig. 6b was apriori coded [0,1,0] because it is incorrect, normative, and non-serviceable. The samples demonstrated various combinations of these codes to cultivate awareness of different features and challenge assumptions, such as the belief that correct work must necessarily be normative.

The *Anticipating Student Work* simulation begins with participants viewing the start of the lesson, illustrated with a storyboard, to situate them in the phase where students are grappling with the problem while a teacher (represented by the participant's avatar) walks around the room taking notes. Participants then sort student work into bins in various ways, starting with six samples and two

The activity	Illustrating screens
0. Context: participants are situated in the specific phase of the lesson (the beginning); they see a reminder of the class's previous knowledge and the teacher presenting the problem.	
1. First sorting: participants are shown six samples of student work and are asked to sort them into bins, name the bins, and explain their sorting to hypothetical colleagues. Then they are asked how many bins they would prefer to use, between 3-5, and they sort the six samples accordingly.	Imagine that as you look at the student work you mentally sort it into two bins. Drag each piece of student work to the bin where you think it should go.
2. Second sorting: participants are now shown 18 samples of student work (including the previous six samples) and are asked to sort them into their existing bins. Then they are asked how many bins they would prefer to use, between 3-8, they resort the 18 samples accordingly, and explain their sorting to hypothetical colleagues.	Here, we have provided you with the number of bins indicated needing on the previous screen. Please show how you would sort the images of the student work into the number of bins you selected. (Order does NOT matter.)
3. Reflection: participants share their thoughts on the choices they had made when sorting student work, including comparing the rounds of sorting and noting which samples were harder to fit into bins.	Besides having a different number of bins, what differences do you see in the way you sorted the 6 pieces of student work, between the first and the second activities?

Fig. 7 The structure of the anticipating student work simulation

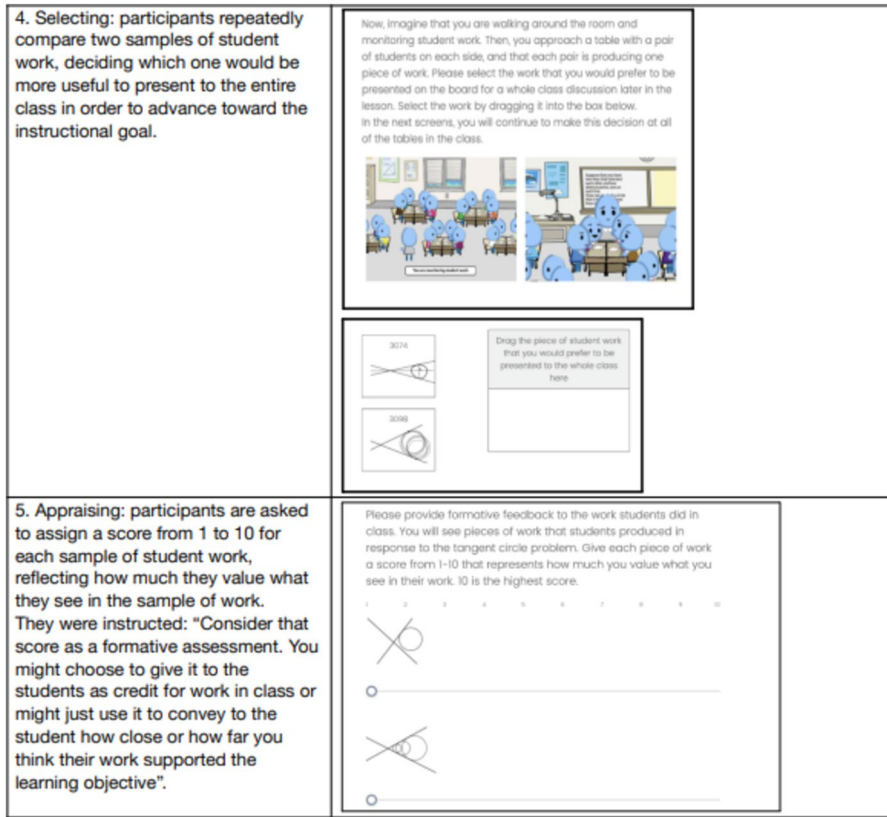


Fig. 7 (continued)

bins, gradually increasing the number of bins and increasing the number of work samples from six to 18 (see the detailed instructions and illustrations in Fig. 7). Later, participants reflect on their decisions, select student work, and provide formative assessment to students (see Fig. 7).

In terms of *participant structure*, the simulations engage teachers individually, providing didactic feedback by visualizing the decisions (e.g., participants see how they sorted the samples of student work) and offering opportunities to revise these decisions. Importantly, no direct feedback is given regarding the student work samples (e.g., what are considered right or wrong answers for sorting pieces of student work into bins, or selecting student work).

Mediating Processes

Our design conjecture predicted that participants would initially sort student work based on correctness or normativity (e.g., correct/incorrect or formal/informal). By

increasing the number of sorting bins (from 2 to 3–5) and samples (from 6 in the first activity to 18 in the second activity), we aimed to raise awareness of the possibility of additional considerations which might elicit awareness of serviceability. These activities were designed to prompt teachers to reconsider their initial sorting and question their criteria. We hypothesized that this mediating process would be supported by reflective questions (the third activity), allowing participants to articulate the reasons behind their sorting.

Desired Outcomes

The desired outcomes include enhancing participants' focus on serviceability, resulting in increased selection of serviceable student work and positive appraisals of such work (in the fourth and fifth activities, respectively). The theoretical conjecture posits that as participants *notice* additional categories of perception, these categories will influence their *selection* of student work. Ultimately, our goal in designing this simulation was to better understand the connections between teachers' noticing of student work and decision-making related to the practices of selecting and sequencing.

Cycle 1: Enactment and Analysis

In Cycle 1, the simulation was implemented with a group of MTEs who individually engaged with one simulation per week. Additionally, each week, two groups of five MTEs participated in focus groups with the design team to discuss the learning opportunities and limitations of the simulations. These focus group meetings used slides to direct participants' attention on particular parts of the simulation, incorporating both broad questions ("What could teachers learn from this simulation?") and focused questions ("Do you have any comments about the number of samples of student work used?").

In this section, we highlight two cross-cutting themes (hereafter design tensions) identified in both groups. We demonstrate their connection to elements in the conjecture map and explain how the analysis of focus group discussions informed revisions of the *Anticipating Student Work* simulation. The two focus groups are designated A and B, and their composition was based on the days of the week when MTEs were available to meet. Meeting numbers, for each focus groups, are designated with a number between 1 and 5 that follows the A/B notation. Talk turns are represented with numbers after the underscore (e.g., A3_137 is the 137th turn talk in the third meeting of focus group A).

Design Tension I: Lack of Authenticity and Context

A recurring theme in both focus groups was the concern over the simulation's authenticity in accurately reflecting teachers' daily work, specifically how closely simulation activities mirror authentic decision-making in teaching. Notably,

MTEs were not concerned about the photorealism of the characters in the simulations (to read more about this issue, see Brown et al., 2023). However, the activities involving sorting student work into bins received some criticism for perceived lack of relevance to the lesson. Despite those concerns, MTEs recognized the potential benefits of the activities:

There could just be a little more emphasis on the idea of selecting around building a storyline [...] I don't know if it was quite emphasized enough, [...] the idea of building towards [...] a coherent storyline based on students' ideas, students' thinking [...], that seemed implicit to me (Isaac, A3_120–132).

I see the value in just focusing on – can you make sense of what the students have done? – [...] This is the first step in order to think about the storyline and all of that. First, I really have to make sense of what the students have done, and not just superficially (Julia, A3_145–153).

These comments highlight MTEs appreciation for sorting activities to enhance noticing but also express a desire for increased relevance, emphasizing the need for a clear connection between noticing tasks and subsequent decision-making activities. This pointed to the weakness of the theoretical conjecture we had made linking reconsideration of sorting to increased category awareness and selection. Isaac argued that if teachers are tasked with selecting student work without explicit guidance on crafting a mathematical storyline, their awareness of additional categories may not influence their decision-making. Julia acknowledged the value of sensemaking about student work as a preliminary step but suggested it should be followed by consideration of the broader storyline.

Interestingly, both groups suggested framing the sorting activities as a precursor to selecting and sequencing student work, with a tighter connection to the five practices framework (Stein et al., 2008). They expressed reservations about the selection activity (fourth activity in Fig. 7), stating it does not reflect real-life teacher selection processes:

Part of the challenge about it is [...] It just feels a little artificial for what happened in a classroom, because in a classroom you would scan everybody's before you chose three to share, or something close to that. And so, picking from two repeatedly, you don't get the whole picture. [...] It just feels removed from what you'd actually do in a classroom (Chloe, A3_923–930).

The fifth activity, of appraising student work employing formative assessment, was heavily criticized. For example, Layla stated: "I would have a problem looking at this [student work] and assigning a number to it instead of providing [verbal] formative feedback" (B3_904). We were interested in this task as a way to gauge the extent to which teachers would value student work that was serviceable but non-normative compared to work that was normative but non-serviceable. Our choice of formative (rather than summative) assessment as the context attempted to acknowledge that none of the student work included complete answers to the problem. However, our attempt to elicit a numerical appreciation

of student work within a formative (rather than summative) assessment context, appeared confusing and distracting to the MTEs due to its lack of authenticity. This is evident in the following discussion:

Ian: The formative frame is problematic, because when I give formative feedback it's almost never going to be a numerical value of any kind. [...] If you ask me to evaluate their work or grade their work, then sure I can assign a number to it if I want. But the frame was, how much I value what I see in their work, and my problem was, I found something to value in all the work. So that's my like asset-based training that's getting in the way of this particular activity (A3_1028–1033).

Julia: Giving it a number doesn't make sense to me (A3_1037).

Chloe: I'm really glad to hear Ian and Julia saying that because I struggled with this so much, I sat here looking at it for 10 minutes, and I gave up. I was like, "It doesn't make sense to me to score it" [...] it's their work, it's their attempt, it's their thinking. I guess I'd give them all tens, you know, because they've all put in some really good thinking (A3_1038–1046).

This discussion highlighted the problematic design issue of asking teachers to make decisions that conflict with their teaching ideology, thereby undermining presence and engagement ("I sat here looking at it for 10 minutes, and I gave up"). It underscored the need to find an alternative way to measure desired outcomes by confining simulation activities to actions that teachers typically do or envision themselves doing in real-life and to activities that are situated in practice. It reminded us that practice-based simulations should authentically reflect real teaching experiences, grounded in practitioners' logic.

Redesign to Increase Authenticity

To enhance the practice-based, authentic, and purposeful nature of the activities, we revised the simulation to serve as a step-by-step heuristic for deliberate selecting and sequencing student work aimed at crafting a mathematical storyline (Stein et al., 2008). Instructions for the sorting activities were updated to emphasize that sorting could scaffold informed selecting and sequencing. The last two activities were replaced with a final task where teachers select and sequence four pieces of work out of the 18 they previously sorted, and elaborate on how they plan to connect them in a whole-class discussion. While the initial selection and appraisal activities promised simpler data for precise evaluation of desired outcomes, the new final task better aligns with our approach to teacher learning. Categories of perception can still be examined by analyzing participants' rationales for selecting specific student samples. Additionally, these categories can be explored by comparing the student work participants choose when simulating teaching the full lesson before and after the intervention simulations—a revision we introduced for the next cycle.

Design Tension II: Lack of Feedback

MTEs expressed concern that novice teachers might struggle to gain meaningful insights from the simulations if completed individually, without the benefit of peer or instructor feedback. Ian (A3_732) remarked, “You only know what you know”, while Zara (A3_778) added, “They’re [novice teachers] not at a point of being able to process a lot of nuance between student work”, emphasizing the need for explicit discussion of the subtleties in the design of student work samples for effective learning. Mila noted the difficulty of changing sorting heuristics without seeing counterexamples (B3_687). These comments relate to a potential limitation of the design conjecture, which suggests that participants are expected to increase their awareness of student work features through changes in the number of bins and increased variation in student work. Jacob explained (B3_728–732) why he thought this design does not guarantee improved noticing by novices:

It didn’t get any better going from the two groups – “They got it”, “They didn’t get it” – to four groups where they “definitely got it”, “They sort of got it”, “They didn’t really get it”, “They definitely didn’t get it”.

Using the conjecture map language, Jacob’s comment suggested that repeated engagement in sorting student work may not necessarily result in reconsiderations. Importantly, Jacob also attributed much more efficacy to the focus group discussion, compared to doing the simulation individually, highlighting that the current participant structure (which does not include peer interaction) might need to be updated. This comment was particularly noteworthy as it emphasized the advantages of peer-discussions in perturbing established thought patterns, making us recognize that we still do not have an effective mechanism to replace the social learning aspects of PBTE, mainly peer-interactions. The MTEs not only identified problematic links in our design logic, but also offered valuable suggestions for improvements, such as: “seeing other people’s classification schemes [...] is a wonderful way to broaden your horizons about ways to re-categorize work” (Ian, A3_732).

Redesign to Emulate Peer-Feedback

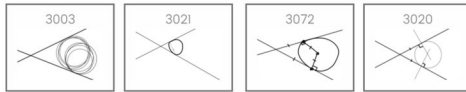
We welcomed the idea of incorporating peer-feedback to replicate the productive peer-interactions integral to collaborative PBTE, including our *StoryCircles* model. However, we encountered a significant challenge during the revision process: How could we discourage users from perceiving peer-feedback as definitive solutions dictating what “should have been done”? How could we strike a balance between the value of added feedback for learning and our approach of avoiding prescriptivism?

To address this tension, we presented the sorting examples as contributions from other teachers (see Fig. 8). This approach emphasizes that there is not a single correct way to sort and evaluate student work, thereby encouraging diverse perspectives. Figure 9 illustrates the updated conjecture map following the revisions of Cycle 1.

Here are some examples of sortings that we have seen other teachers do. Please note that their sorting is by no means "the correct" answer. This is just an opportunity to consider how others may have done this work and contrast it with your own.

Example 1:

Bin 1: Perpendiculars



Bin 2: Angle bisectors

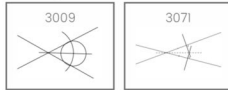


Fig. 8 A screen using peer-feedback, added to the Anticipating Student Work simulation

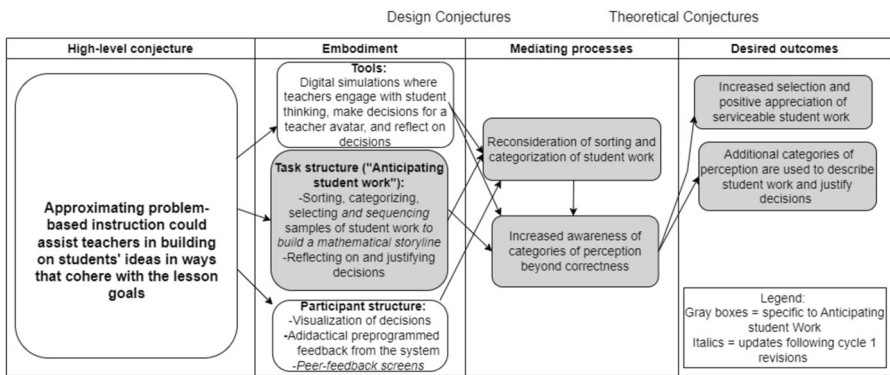


Fig. 9 Updated conjecture map following the revisions of Cycle 1

Cycle 2: Enactment and Analysis

In the Fall semester of 2023, we implemented the next iteration of the simulations with a group of PTs as part of a methods course, one of the target audiences for these simulations. This time, our focus was on examining whether participants' performance aligned with our theoretical conjectures.

Before presenting the results, we briefly describe additional revisions to the simulations that extended beyond those discussed for *Anticipating Student Work*. We introduced various forms of adidactic feedback to the Responding to Student Work and Teaching the Lesson with Student Participation simulations. Additionally, we added frames and accompanying questions to *Getting to Know the Lesson* to clarify potential solutions to the problem. Also, an important change was made to improve the measurement of desired outcomes with greater face validity: Participants took the *Teaching the Lesson with Student Participation* comprehensive simulation twice (Fig. 10): once after completing the introductory simulation (pre-interventions) and again after completing the two intervention simulations (post-interventions). Below,

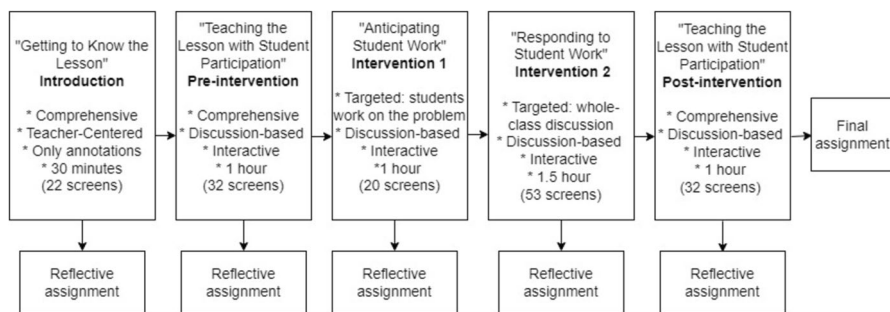


Fig. 10 The simulations and reflective assignments in cycle 2

we present results from comparing the pre and post comprehensive simulations, focusing on noticing, selecting, and sequencing student work, along with participants' reflections on the revised *Anticipating Student Work* simulation.

Analysis of Teachers' Performance in the Pre- and Post-Interventions

By asking teachers to simulate the teaching of the entire lesson twice and treating the *Anticipating Student Work* simulation as an intervention, we were able to make direct comparisons and provide evidence that some desired outcomes were achieved. In the *Teaching the Lesson with Student Participation* simulation, participants faced several key teaching decisions, each leading to a different branch. They first decided how to frame the problem, such as whether to provide a diagram with marked points. Then, they viewed eight samples of student work (aligned with their framing choice; each branch had different samples), taking notes as they monitored the class. Later, they selected and sequenced three out of eight samples for a whole-class discussion.

The eight samples varied in correctness, normativity, and serviceability, coded with triplets of ones and zeros accordingly. These codes facilitated the creation of scores for each construct to measure participants' improvement in selecting serviceable work between pre- and post-interventions, one of the desired outcomes. With each sample having "0" or "1" codes for each construct, participants' selection of three samples of student work yielded a score between 0 and 3. Additionally, we conducted a thematic analysis of responses to open-ended questions to identify changes in participants' justifications for selecting and sequencing.

Our analysis shows the achievement of some of the desired outcomes. Using the scores outlined above, we conducted a two-tailed Wilcoxon test to compare participants' decisions on selecting student work in the pre- and post-interventions. The results (see Table 1) indicate that in the post-intervention, while normativity and

Table 1 Results of a two-tailed Wilcoxon signed rank test to identify changes in correctness, normativity, and serviceability scores

	Correctness	Normativity	Serviceability
<i>P</i> -value (Asym. Sig., 2-tailed)	0.157	1.00	0.026 (< 0.05)
<i>Z</i>	1.414	0	2.232

correctness scores did not significantly change (as expected), the serviceability score increased significantly (p -value = 0.026, $Z = 2.232$). These findings are supported by the thematic analysis of the justifications provided.

In the initial round, PTs cited reasons for selecting and sequencing such as:

Danny: Start with one that has common errors then roll into one that adjusted their approach to talk about the relations it made and then end with one that's correct.

Gregory: I want to start with clearing any misconceptions.

These examples illustrate the prevailing emphasis on the correctness of student work ("common errors", "misconceptions"), and limited discussion of students' conceptions. They reflect the expectation that a sequence of samples should incorporate common errors, leading to a correct solution. In the post-interventions, the emphasis changed:

Max: None of them are correct but I am hoping they notice something about the points that they are choosing.

Zachary: [...] I then want to introduce Sigma's work because although it is not precise, I'm hoping the class can still learn from their idea that the tangent line is perpendicular to the radius.

These excerpts illustrate that participants continued to consider correctness and normativity when making their selections ("not precise") but minimized their relevance. Notably, participants mentioned other aspects found in the student work samples, linking them to crucial mathematical concepts necessary for achieving the lesson goal.

A surprising finding was that some participants shifted from an approach to sequencing that progressively *approximates* the correct solution (Zoe, pre-interventions: "It is important to see the *progression* of detail") to selecting student work that highlights *complementary aspects* of the solution (Zoe, post-interventions: "They are each bringing *something different* to the table"). This unexpected outcome has now been incorporated into our conjecture map, as we hypothesize that it may be influenced by the structure of the sorting activities.

Others transitioned to showcasing various methods of solving. This shift suggests a stronger outcome than initially anticipated: participants not only noticed and selected more serviceable work, but their justifications also indicated opportunities for generative connections between student work samples. Overall, sequencing heuristics that prioritize broader participation over correctness seem to emerge from this shift.

Results from Participants' Reflective Assignments

Further evidence supporting the effectiveness of the design was identified in the analysis of participants' reflective assignments written following the *Anticipating*

Student Work simulation. We looked for mentions of previously identified design challenges, such as authenticity and feedback, as well as any emerging issues. Results show that participants had minimal concern with authenticity, presence, or motivation. Some participants spontaneously referenced the connections we aimed to establish between sorting activities and subsequent selecting and sequencing, as well as their utilization of the added peer-feedback:

Diana: This activity [Anticipating Student Work] made it much easier for me to select work as to create a mathematical story than last week's simulation [refers to the pre-intervention]. The structure of the "boxes" made sense, and it allowed me to come up with common threads that could then be elaborated on. I liked getting to see how others sorted student work and how I could borrow from the reasoning of other educators, not necessarily changing how I was sorting work but taking that aspect of the problem into consideration when I was formulating my responses to student work.

Danny: After seeing how others did it and more student work was added [referencing the shift from six to 18 samples of student work] I saw the value in labeling my bins with a more focused ideology which would not only make them easier to understand but every student's work would easily fit into one category or another. Also taking in the account of time and only being able to show a handful of work, choosing work that builds off one another was beneficial because it would save time and slowly display the concepts I would want the student to recognize.

We also identified explicit mention of our design conjectures, in terms of the increased number of bins and the enhanced variability of student work:

Zachary: It's easy to mentally sort students' work in two "bins", oftentimes representing "right" and "wrong". However, throughout the simulation I found it to be more beneficial to sort students' work in 3+ "bins" based on their similarities and/or differences. Incorrect work can still be a helpful component of a math storyline.

Zelda: I really like this activity because I got to see my growth in responses and how I organized the student's work throughout the activity [...] At first my organization was quite vague; it was between how much detail was being presented.

However, we also encountered an unexpected concern about labeling students that warrants consideration in future revisions:

Eugenia: I also wonder how necessary it is to sort the student work into bins like these, essentially putting labels on students even if it is just for this one lesson.

This comment illustrates a caveat of self-paced, digital simulations. If shared in collaborative PBTE, it could prompt a fruitful teacher discussion, with peers and

MTEs questioning Eugenia's interpretation of labels. We are currently seeking suitable platforms to facilitate such exchanges among participants, inspired by successful implementation of large-scale peer-feedback in MOOCs (Gamage et al., 2021).

Overall, the analysis of reflections demonstrates that the adapted simulation successfully navigated the delicate balance of presenting additional approaches to sorting student work without portraying them as definitive solutions, and that the didactic feedback from the system (e.g., sorting 18 samples instead of six) has also contributed to participants' awareness of additional categories. Furthermore, upon reviewing the reflections, we recognized that our intended outcome might have been too modest. PTs described a shift in their attitudes toward problem-based instruction, recognizing its complexity and developing strategies for both planning ahead and responding in the moment. Our next steps involve analyzing data from practicing teachers to further refine the conjecture map, with a focus on emerging outcomes that reveal potential learning opportunities afforded by the simulations, which we did not foresee.

Discussion

This study contributes to understanding how asynchronous digital simulations with didactic feedback, adapted from a collaborative practice-based model, can support mathematics teachers' opportunities to learn about problem-based instruction. The paper also explores what resources and mechanisms can be integrated into simulations to emulate aspects of collaborative PBTE. Using conjecture mapping (Sandoval, 2014), we illustrate how our design of the simulations evolved to better reflect the vision and principles driving the original PD program, as well as to increase responsiveness to educators' and learners' needs. Since the mechanisms of teacher learning in subject-specific digital simulations are not yet well specified (Mikeska et al., 2021), this paper makes a contribution by detailing methods and design principles for identifying and developing teachers' situated, tacit knowledge. Moreover, it provides proof of concept that PTs' noticing and decision-making can be enhanced to be more responsive to students' serviceable mathematical ideas. Below, we discuss how our main results engage with the existing literature and extend it.

Mathematics Teachers' Opportunities to Learn about Problem-Based Instruction in Practice-Based Asynchronous Digital Simulations

The results indicate that PTs improved their ability to notice and select student work that is serviceable, focusing on the lesson goal and setting aside notions of correctness and normativity. This supports previous studies highlighting (a) the difficulty of in-the-moment noticing and responsiveness (Kilic & Dogan, 2022; Osmanoglu & Girit-Yildiz, 2024), (b) the potential for enhancing strength-based noticing (Scheiner, 2023), and (c) the effectiveness of practicing these skills within an approximation of practice context (e.g., Estapa & Davis, 2023).

Our unique design, which is both asynchronous and uses didactic feedback, allowed PTs to increasingly build on students' serviceable ideas while also voicing their recognition of the complexity of teaching. An unexpected finding was that PTs developed heuristics for selecting and sequencing student work that might rely on the structure of the simulations, namely, sorting samples of student work into bins, naming the bins, and then selecting one piece of student work from each bin. Some of them shifted their view of sequencing from gradually approximating the correct solution to selecting student work that highlights complementary aspects of the solution. This connection suggests a new insight into the relationship between what teachers notice and what they decide to do. It seems that the *structure* of the sorting (noticing) activities eventually shaped PTs' decisions when selecting and sequencing. Thus, we suggest further examining the claim that the nature of the noticing activity shapes teachers' ensuing decision-making. This is a novel insight regarding how noticing relates to decision-making, which could inform future design of simulations and research on what people learn from them.

Resources and Mechanisms that can Emulate Aspects of Collaborative PBTE

A key design principle identified through our research is that, for practice-based simulations to be impactful, they must authentically imitate aspects of teachers' work, drawing on resources such as student work, lesson plans, and authentic problems of practice. Simulations should be coherent, with each part logically leading to the next, and must appear reasonable and relevant to teachers. This aligns with Howell and Mikeska's (2021) interpretation of authenticity in the context of approximations of practice.

Another critical aspect is the role of peer-feedback in collaborative PBTE. Given that our current model does not support asynchronous interactions between participants to maintain engagement, our findings suggest that incorporating resources like milieu feedback screens can enhance teacher learning within self-paced, standalone simulations. Although we did not conduct a controlled random trial to confirm this, evidence from PTs' reflections indicated that peer-feedback was instrumental mechanism in their evolving sense-making of student work samples. Future research could enhance this by employing artificial interactive "peers" (Copur-Gencturk & Orrill, 2023), trained with the data from teacher responses to the first generation of the simulations.

In terms of mechanisms, our study lays the groundwork for articulating learning mechanisms in digital simulations by emphasizing the importance of *flow-enhancing experiences*. These experiences involve teachers being immersed in action without being asked to reflect or make meta-comments on their decisions in real-time. Additionally, receiving feedback from the system, such as milieu feedback and visualizations of decisions, plays a crucial role. While teacher reflection during peer discussions in PBTE is a prominent mechanism for professional growth (Brandenburg et al., 2017; Schwarts & Karsenty, 2020), we propose that within *approximations* of practice, *experiencing* is key for elucidating and enhancing tacit knowledge, and reflection plays a secondary role. Specifying the connections, within digital simulations, between

experience and reflection – key constructs in educational literature – is of high importance for theorizing teachers' learning in approximations of practice.

Our study highlights the utility of conjecture mapping in adapting collaborative PBTE models into online self-paced formats. We propose that when adapting synchronous collaborative PDs into an asynchronous individual PD, an initial step could comprise mapping the original program, ensuring the inclusion of only salient elements (Sandoval, 2014), and distinguishing between elements that could be easily translated into digital settings, elements that needed to be transformed, and elements that should be replaced. We advocate for gathering diverse data sources, from multiple stakeholders, to analyze different elements and connections within the conjecture map.

Limitations and Future Work

Conjecture mapping allowed us to clarify hypotheses, identify relationships, and distinguish between expected and unexpected outcomes (Sandoval, 2014). Most importantly, it provided a schematic view of the logic underlying the simulations, providing us with argumentative grammar to discuss how the design elements relate to each other. However, this approach has limitations, such as our lack of hard evidence that the mediating processes are indeed the reason for the observed changes in PTs' decisions. Moreover, we lack details on the discussions about features of student work in the PTs' methods course, and thus the methods course is not part of the map. The small number of design cycles, as well as number of participants, are another limitation. Nevertheless, the latter limitation allowed us to generate a blueprint for the desired outcomes with a larger sample, data that we are currently collecting and analyzing. In particular, we aim to make theoretical contributions regarding the interrelationships between noticing, selecting, and sequencing, that align with recent calls to relate to the situated nature of noticing (Scheiner, 2021, 2023; van Es & Sherin, 2021). Furthermore, we aim to theorize learning mechanisms within practice-based asynchronous digital simulations.

Concluding Remarks

Our work bridges the gap between the literature on collaborative practice-based teacher learning, which relies on peer-interactions, and the emerging literature on teacher learning in and from digital simulations. We sketch a possible path for scaling and sustaining collaborative PBTE, in ways that align with their fundamental tenets. This is a promising avenue for the “rich and strange” (Dede, 2022, p. 118) post-COVID era of teacher education. However, although we successfully devised a mechanism to perturb participants' ideas akin to peer-interactions, questions remain: Does it make sense for teachers to explore collaborative problem-based instruction via an individual platform? Should teachers learn about collaborative instruction only within a collaborative setting? And, how could the definition of teacher collaboration expand in the context of recent technological advances? Answering these questions should guide future designs.

Looking ahead, our objective is to further develop resources and mechanisms for teachers' learning from digital simulations, fostering a versatile design capable of

catering to diverse teacher populations. This endeavor may entail the integration of AI-agents or generative-AI-produced student responses, streamlining the simulation design process, and amplifying the breadth of possible human-computer interactions beyond what is currently preprogrammed. Envisioning the potential for avatars trained on large learning models to offer approximations of real-world classroom dynamics, we recognize that this technological leap must be accompanied by a thorough examination of its theoretical implications. In essence, our journey toward advancing digital simulation capabilities is not only about technological progress but also – and perhaps mostly – about delineating the educational theory underpinning these innovations, and how it informs teachers' practice.

Acknowledgements We deeply thank Irma Stevens and Kristin Lesseig for their support of this work. We also thank the mathematics teacher educators whose feedback was instrumental for the revisions of the simulations.

Funding Open access funding provided by Weizmann Institute of Science. The analysis and writing of this paper were supported by James S. McDonnell grant 220020524. The opinions expressed are solely those of the authors and do not necessarily represent the views of the Foundation.

Data Availability The data supporting the findings of this study are available upon reasonable request but are not publicly accessible to protect the privacy of the research participants.

Declarations

Ethical The research reported in this manuscript was approved by the Institutional Review Board of the University of Michigan.

Conflict of interest There is no conflict of interest regarding this submission.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bakker, A. (2018). *Design research in education*. Routledge. <https://doi.org/10.4324/9780203701010>
- Blömeke, S., Gustafsson, J., & Shavelson, R. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift für Psychologie*, 223, 3–13.
- Bondurant, L., & Amidon, J. (2021). Virtual field experiences as an opportunity to develop preservice teachers' efficacy and equitable teaching practice. In K. Hollebrands, R. Anderson, & K. Oliver (Eds.), *Online learning in mathematics education* (pp. 317–334). Springer. https://doi.org/10.1007/978-3-030-80230-1_16
- Borko, H., & Potari, D. (2024). *Teachers of mathematics working and learning in collaborative groups: The 25th ICMI Study*. New ICMI Study Series: Springer.

- Brandenburg, R., Glasswell, K., Jones, M., & Ryan, J. (Eds.). (2017). *Reflective theory and practice in teacher education*. Springer Singapore.
- Brennan, K., Blum-Smith, S., & Yurkofsky, M. M. (2018). From checklists to heuristics: Designing MOOCs to support teacher learning. *Teachers College Record*, 120(9), 1–48. <https://doi.org/10.1177/016146811812000904>
- Brousseau, G. (1997). *Theory of didactical situations in mathematics: Didactique des Mathématiques 1970–1990* (N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield, Eds. & Trans.) Kluwer. <https://doi.org/10.1007/0-306-47211-2>
- Brown, A. M., & Herbst, P. G. (2023). On designing better structures for feedback in practice-based professional development: Using failure to innovate. *Journal of Mathematics Teacher Education*, 26(5), 581–605. <https://doi.org/10.1007/s10857-023-09588-1>
- Brown, A. M., Stevens, I., Herbst, P., & Huhn, C. (2021). Confronting teachers with contingencies to support their learning about Situation-Specific Pedagogical decisions in an online context. In K. Hollebrands, R. Anderson, & K. Oliver (Eds.), *Online learning in Mathematics Education. Research in Mathematics Education* (pp. 291–316). Springer. https://doi.org/10.1007/978-3-030-80230-1_15
- Brown, A. M., Bardelli, E., Herbst, P. G., & Dimmel, J. K. (2023). Examining the potential of Cartoon-based simulations for studying Mathematics teachers' handling of Student emotions: A replication study. *Implementation and Replication Studies in Mathematics Education*, 3(2), 243–274.
- Chazan, D., Herbst, P. G., & Clark, L. M. (2016). Research on the teaching of mathematics: A call to theorize the role of society and schooling in mathematics instruction. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of research on teaching* (pp. 1039–1097). American Educational Research Association. https://doi.org/10.3102/978-0-935302-48-6_17
- Clarke, V., & Braun, V. (2017). Thematic analysis. *The Journal of Positive Psychology*, 12(3), 297–298.
- Cohen, D. K., & Ball, D. L. (1999). *Instruction, capacity, and improvement (CPRE Research Report Series RR-43)*. Consortium for Policy Research in Education, University of Pennsylvania Graduate School of Education. <https://doi.org/10.1037/e382692004-001>
- Copur-Gencturk, Y., & Orrill, C. H. (2023). A promising approach to scaling up professional development: Intelligent, interactive, virtual professional development with just-in-time feedback. *Journal of Mathematics Teacher Education*. <https://doi.org/10.1007/s10857-023-09615-1>
- Davis, E. A., Kloser, M., Wells, A., Windschitl, M., Carlson, J., & Marino, J. C. (2017). Teaching the practice of leading sense-making discussion in science: Science teacher educators using rehearsals. *Journal of Science Teacher Education*, 28(3), 275–293. <https://doi.org/10.1080/1046560X.2017.1302729>
- Dede, C. (2022). The coming sea-change in teacher education. *Journal of Technology and Teacher Education*, 30(2), 117–125.
- Dieker, L. A., Rodriguez, J. A., Lignugaris, B., Hynes, M. C., & Hughes, C. E. (2013). The potential of simulated environments in teacher education: Current and future possibilities. *Teacher Education and Special Education*, 37(1), 21–33. <https://doi.org/10.1177/0888406413512683>
- Estapa, A., & Davis, J. (2023). Prospective teachers' instructional decisions and pedagogical moves when responding to Student thinking in Elementary mathematics and Science lessons. *International Journal of Science and Mathematics Education*, 21(5), 1703–1724. <https://doi.org/10.1007/s10763-022-10304-3>
- Ferdig, R. E., Kosko, K. W., & Gandolfi, E. (2022). Using the COVID-19 pandemic to create a vision for XR-based teacher education field experiences. *Journal of Technology and Teacher Education*, 30(2), 239–252.
- Friessen, M., Kapp, F., Barzel, B., Dreher, A., Holzäpfel, L., Larrain, M., & Hiemenz, A. (2023). Online professional development for teaching algebra: Towards the design of an asynchronous, adaptive self-learning module. In P. Drijvers, C. Csapodi, H. Palmér, K. Gosztonyi, & E. Kónya (Eds.), *Proceedings of the Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13)* (pp. 3395–3402). Alfréd Rényi Institute of Mathematics and ERME. <https://hal.science/CERME13/hal-04421614v1>. Accessed 10 Apr 2024
- Gamege, D., Staubitz, T., & Whiting, M. (2021). Peer assessment in MOOCs: Systematic literature review. *Distance Education*, 42(2), 268–289. <https://doi.org/10.1080/01587919.2021.1911626>
- Gibson, D. (2007). SimSchool and the conceptual assessment framework. In D. Gibson, C. Aldrich, & M. Prensky (Eds.), *Games and simulations in online learning: Research & development frameworks* (pp. 308–322). Idea Group.
- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96(3), 606–633.

- Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, 111(9), 2055–2100. <https://doi.org/10.1177/016146810911100905>
- Herbst, P. G. (2006). Teaching geometry with problems: Negotiating instructional situations and mathematical tasks. *Journal for Research in Mathematics Education*, 37(4), 313–347.
- Herbst, P., & Chazan, D. (2011). Research on practical rationality: Studying the justification of actions in mathematics teaching. *The Mathematics Enthusiast*, 8(3), 405–462.
- Herbst, P., & Milewski, A. (2018). What StoryCircles can do for mathematics teaching and teacher education. In R. Zazkis & P. Herbst (Eds.), *Scripting approaches in mathematics education: Mathematical dialogues in research and practice* (pp. 321–364). Springer. <https://doi.org/10.1007/978-3-319-62692-5>
- Herbst, P., Chazan, D., Chieu, V. M., Milewski, A., Kosko, K., & Aaron, W. (2016). Technology mediated mathematics teacher development: Research on digital pedagogies of practice. In M. L. Niess, S. Driskell, & K. Hollerands (Eds.), *Handbook of research on transforming mathematics teacher education in the digital age* (pp. 76–105). IGI Global. <https://doi.org/10.4018/978-1-5225-0120-6.ch004>
- Herbst, P., Boileau, N., Shultz, M., Milewski, A., & Chieu, V. M. (2020). What Simulation-based mentoring may afford: Opportunities to connect theory and practice. In E. Bradley (Ed.), *Games and simulations in Teacher Education. Advances in game-based learning* (pp. 91–114). Springer. https://doi.org/10.1007/978-3-030-44526-3_7
- Herbst, P., Shultz, M., Bardelli, E., Boileau, N., & Milewski, A. (2022). How can teaching simulations help us study at scale the tensions mathematics teachers have to manage when considering policy recommendations? *Educational Studies in Mathematics*, 110(1), 1–21. <https://doi.org/10.1007/s10649-021-10118-0>
- Herbst, P., Brown, A., Chazan, D., Boileau, N., & Stevens, I. (2023). Framing, responsiveness, serviceability, and normativity: Categories of perception teachers use to relate to students’ mathematical contributions in problem-based lessons. *School Science and Mathematics*, 123(7), 398–413. <https://doi.org/10.1111/ssm.12600>
- Howell, H., & Mikeska, J. N. (2021). Approximations of practice as a framework for understanding authenticity in simulations of teaching. *Journal of Research on Technology in Education*, 53(1), 8–20. <https://doi.org/10.1080/15391523.2020.1809033>
- International Program Committee. (2019). *Teachers of mathematics working and learning in collaborative groups: Discussion document*. <http://icmistudy25.ie.ulisboa.pt/discussion-document/>. Accessed 10 Apr 2024
- 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. <https://www.jstor.org/stable/20720130>
- Kaiser, G., Blömeke, S., König, J., Busse, A., Döhrmann, M., & Hoth, J. (2017). Professional competencies of (prospective) mathematics teachers—cognitive versus situated approaches. *Educational Studies in Mathematics*, 94(2), 161–182. <https://doi.org/10.1007/s10649-016-9713-8>
- Kelly, A. E. (2004). Design research in education: Yes, but is it methodological? *Journal of the Learning Sciences*, 13(1), 115–128. https://doi.org/10.1207/s15327809jls1301_6
- Kilic, H., & Dogan, O. (2022). Preservice mathematics teachers’ noticing in action and in reflection. *International Journal of Science and Mathematics Education*, 20(2), 345–366. <https://doi.org/10.1007/s10763-020-10141-2>
- Koichu, B. (2014). Reflections on problem-solving: Problem solving in mathematics and in mathematics education. In M. N. Fried, & T. Dreyfus (Eds.), *Mathematics & mathematics education: Searching for common ground* (pp. 113–135). Springer.
- Kooloos, C., Oolbekink-Marchand, H., van Boven, S., Kaenders, R., & Heckman, G. (2022). Building on student mathematical thinking in whole-class discourse: Exploring teachers’ in-the-moment decision-making, interpretation, and underlying conceptions. *Journal of Mathematics Teacher Education*, 25(4), 453–477. <https://doi.org/10.1007/s10857-021-09499-z>
- Lampert, M., Franke, M. L., Kazemi, E., Ghouseini, H., Turrou, A. C., Beasley, H., Cunard, A., & Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, 64(3), 226–243. <https://doi.org/10.1177/0022487112473837>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

- Lee, H. S., Hudson, R., Casey, S., Mojica, G., & Harrison, T. (2021). Online curriculum modules for preparing teachers to teach statistics: Design, implementation, and results. In K. Hollebrands, R. Anderson, & K. Oliver (Eds.), *Online learning in mathematics education* (pp. 65–93). Springer. https://doi.org/10.1007/978-3-030-80230-1_4
- Lindberg, S., & Jönsson, A. (2023). Preservice Teachers Training with avatars: A systematic literature review of human-in-the-Loop simulations in Teacher Education and Special Education. *Education Sciences*, 13(8), 817. <https://doi.org/10.3390/educsci13080817>
- Louie, N., Adiredja, A. P., & Jessup, N. (2021). Teacher noticing from a sociopolitical perspective: The FAIR framework for anti-deficit noticing. *ZDM – Mathematics Education*, 53(1), 95–107. <https://doi.org/10.1007/s11858-021-01229-2>
- Martin, W. G., & Gobstein, H. (2015). Generating a networked improvement community to improve secondary mathematics teacher preparation: Network leadership, organization, and operation. *Journal of Teacher Education*, 66(5), 482–493. <https://doi.org/10.1177/0022487115602312>
- Mason, J. (2002). *Researching your own practice: The discipline of noticing*. Routledge. <https://doi.org/10.4324/9780203471876>
- Meller, G. (1997). A typology of simulators for medical education. *Journal of Digital Imaging*, 10, 194–196. <https://link.springer.com/article/10.1007/BF03168699>
- Mikeska, J., Howell, H., Dieker, L., & Hynes, M. (2021). Understanding the role of simulations in K-12 mathematics and science teacher education: Outcomes from a teacher education simulation conference. *Contemporary Issues in Technology and Teacher Education*, 21(3), 781–812.
- Mikeska, J. N., Howell, H., & Kinsey, D. (2023). Do simulated teaching experiences Impact Elementary Preservice teachers' ability to facilitate argumentation-focused discussions in Mathematics and Science? *Journal of Teacher Education*, 74(5), 422–436. <https://doi.org/10.1177/00224871221142842>
- Milewski, A., Herbst, P., Bardelli, E., & Hetrick, C. (2018). The role of simulations for supporting professional growth: Teachers' engagement in virtual professional experimentation. *Journal of Technology and Teacher Education*, 26(1), 103–126.
- National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. Author.
- Organization for Economic Cooperation and Development. (2022). *PISA 2022 mathematics framework*. <https://pisa2022-maths.oecd.org/ca/index.html>. Accessed 10 Apr 2024
- Osmanoglu, A., & Girit-Yildiz, D. (2024). Examining how prospective mathematics teachers' instructional visions align with their responding practices through scripting tasks. *International Journal of Science and Mathematics Education*, 22, 1411–1434. <https://doi.org/10.1007/s10763-023-10435-1>
- Park Rogers, M., Ogundapo, T., Namakula, E., Lane, K., Cross Francis, D., Bharaj, K., Ataide Pinheiro, P., Maltese, W., Mikeska, A., & Shekell, J. N. C. (2024). *Preservice teachers' facilitation of argumentation: Exploring their attention to and perceived complexity of students' thinking*. Paper presented at the National Association for Research in Science Teaching (NARST).
- Reich, J. (2022). Teaching drills: Advancing practice-based teacher education through short, low-stakes, high-frequency practice. *Journal of Technology and Teacher Education*, 30(2), 217–228.
- Rotem, S. H., & Ayalon, M. (2024). Constructing coherency levels to understand connections among the noticing skills of pre-service mathematics teachers. *Journal of Mathematics Teacher Education*, 27, 579–605. <https://doi.org/10.1007/s10857-023-09574-7>
- Russell, J. L., DiNapoli, J., & Murray, E. (2022). Documenting professional learning focused on implementing high-quality instructional materials in mathematics: The AIM-TRU learning cycle. *International Journal of STEM Education*, 9(1), Article 46. <https://doi.org/10.1186/s40594-022-00362-y>
- Sahu, A. R. (1984). Microteaching: Some research studies and research questions. *International Journal of Mathematical Education in Science and Technology*, 15(6), 727–735. <https://doi.org/10.1080/0020739840150608>
- Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences*, 23(1), 18–36. <https://doi.org/10.1080/10580406.2013.778204>
- Santagata, R., König, J., Scheiner, T., Nguyen, H., Adleff, A. K., Yang, X., & Kaiser, G. (2021). Mathematics teacher learning to notice: A systematic review of studies of video-based programs. *ZDM – Mathematics Education*, 53(1), 119–134. <https://doi.org/10.1007/s11858-020-01216-z>
- Santos-Trigo, M. (2024). Problem solving in mathematics education: Tracing its foundations and current research-practice trends. *ZDM – Mathematics Education*, 56, 211–222. <https://doi.org/10.1007/s11858-024-01578-8>
- Scheiner, T. (2021). Towards a more comprehensive model of teacher noticing. *ZDM – Mathematics Education*, 53(1), 85–94. <https://doi.org/10.1007/s11858-020-01202-5>

- Scheiner, T. (2023). Shifting the ways prospective teachers frame and notice student mathematical thinking: From deficits to strengths. *Educational Studies in Mathematics*, 114(1), 35–61. <https://doi.org/10.1007/s10649-023-10235-y>
- Scherpereel, C. M. (2015). Making good decisions: Having confidence in simulations in higher education. *The International Journal of Management Education*, 13(3), 381–389. <https://doi.org/10.1016/j.ijme.2015.02.002>
- Schwartz, G., & Karsenty, R. (2020). Can this happen only in Japan? Mathematics teachers reflect on a videotaped lesson in a cross-cultural context. *Journal of Mathematics Teacher Education*, 23, 527–554. <https://doi.org/10.1007/s10857-019-09438-z>
- Schwartz, G., Pöhler, B., Elbaum-Cohen, A., Karsenty, R., Arcavi, A., & Prediger, S. (2021). Novice facilitators' changes in practices: From launching to managing discussions about mathematics teaching. *Journal of Mathematical Behavior*, 64, 100901. <https://doi.org/10.1016/j.jmathb.2021.100901>
- Schwartz, G., Stevens, I., & Herbst, P. (2022). It's a different mindset here: Facilitation challenges in a practice-based professional development. In A. E. Lischka, E. B. Dyer, R. S. Jones, J. Lovett, J. Strayer, & S. Strayer (Eds.), *Proceedings of the 44th annual meeting of the North American chapter of the international group for the psychology of mathematics education* (pp. 1461–1469). Middle Tennessee State University.
- Schwartz, G., Herbst, P., & Brown, A. (2023). How do mathematics teachers learn to create a mathematical storyline in problem-based lessons? In M. Ayalon, B. Koichu, R. Leikin, L. Rubel, & M. Tabach (Eds.), *Proceedings of the 46th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 187–202). PME 46. https://drive.google.com/file/d/1m_8PUJC4Sd4VoyTbn0WgWXfOnh9TSZGv/view. Accessed 10 Apr 2024.
- Shaughnessy, M., & Boerst, T. (2018). Designing simulations to learn about preservice teachers' capabilities with eliciting and interpreting student thinking. In G. Stylianides & K. Hino (Eds.), *Research advances in the mathematical education of pre-service elementary teachers—An international perspective* (pp. 125–140). Springer. https://doi.org/10.1007/978-3-319-68342-3_9
- Shin, D. (2022). Teaching mathematics integrating intelligent tutoring systems: Investigating prospective teachers' concerns and TPACK. *International Journal of Science and Mathematics Education*, 20(8), 1659–1676. <https://doi.org/10.1007/s10763-021-10221-x>
- Son, T., Yeo, S., & Lee, D. (2024). Exploring elementary preservice teachers' responsive teaching in mathematics through an artificial intelligence-based chatbot. *Teaching and Teacher Education*, 146, 104640. <https://doi.org/10.1016/j.tate.2024.104640>
- Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning*, 10(4), 313–340. <https://doi.org/10.1080/10986060802229675>
- Sweeney, J., Milewski, A., & Amidon, J. (2018). On-ramps to professional practice: Selecting and implementing digital technologies for virtual field experiences. *Contemporary Issues in Technology and Teacher Education*, 18(4), 670–691.
- Tekin-Sitrava, R., Özel, Z., İşıksal-Bostan, M., & Yemen-Karpuzcu, S. (2024). How does online professional development program enriched with collaborative discussion develop teachers' noticing skills? *International Journal of Science and Mathematics Education*. <https://doi.org/10.1007/s10763-024-10461-7>
- Thompson, M., Owho-Ovuakporie, K., Robinson, K., Kim, Y. J., Slama, R., & Reich, J. (2019). Teacher moments: A digital simulation for preservice teachers to approximate parent–teacher conversations. *Journal of Digital Learning in Teacher Education*, 35(3), 144–164. <https://doi.org/10.1080/21532974.2019.1587727>
- van Es, E. A., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, 10(4), 571–596.
- van Es, E. A., & Sherin, M. G. (2021). Expanding on prior conceptualizations of teacher noticing. *ZDM – Mathematics Education*, 53(1), 17–27. <https://doi.org/10.1007/s11858-020-01211-4>
- Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, 15, 17–40. <https://doi.org/10.1016/j.edurev.2015.04.002>
- Zazkis, R. (2017). Lesson play tasks as a creative venture for teachers and teacher educators. *ZDM – Mathematics Education*, 49, 95–105. <https://doi.org/10.1007/s11858-016-0808-6>

Authors and Affiliations

Gil Schwartz^{1,2}  · Patricio Herbst¹  · Amanda M. Brown¹ 

✉ Gil Schwartz
gil.schwartz@weizmann.ac.il

Patricio Herbst
pgherbst@umich.edu

Amanda M. Brown
amilewsk@umich.edu

¹ Educational Studies, Marsal Family School of Education, University of Michigan, 610 East University Avenue, MI 48109 Ann Arbor, USA

² Science Teaching Department, Weizmann Institute of Science , Herzl St. 234, Rehovot 7610001, Israel