Howell, H., Mikeska, J., Lottero-Perdue, P., Shekell, C., Kinsey, D., Lee, C., Reich, J., Maltese, A., Park Rogers, M., Cross Francis, D., Bharai, P. K., & Halder, S. (2025). Exploring the use of simulations in secondary mathematics and science methods courses. *Contemporary Issues in Technology and Teacher Education, 25*(3), 468-515.

Exploring the Use of Simulations in Secondary Mathematics and Science Methods Courses

Heather Howell & Jamie Mikeska
ETS

Pamela Lottero-Perdue
Towson University

Calli Shekell
Thiel College

Devon Kinsey
ETS

Camila Lee
Digital Inquiry Group

Justin Reich
MIT

Adam Maltese & Meredith Park Rogers
Indiana University

Dionne Cross Francis
University of North Carolina at Chapel Hill

Pavneet Kaur Bharaj
California State University Long Beach

Shreyashi Halder
ETS

This article describes the outcomes of using a scaffolded set of two digital, simulation-based approaches to support secondary mathematics and science preservice teachers (PSTs) in learning to facilitate argumentation-focused discussions. The researchers investigated PST and teacher educator (TE) impressions of the experience and collected evidence of change in PSTs' understanding of argumentation, beliefs, and skills in facilitating argumentation-focused discussions, working within mathematics and science methods across six institutions. PSTs and TEs endorsed future use of the approach. Pre/post improvement in beliefs was evident across quantitative and qualitative data sources. There was less clear evidence of improvement in understanding of argumentation, although understandings at both time points were aligned with the study's vision. Evidence of improvement in skill was mixed with quantitative results showing no improvement and qualitative results showing evidence that PSTs and TEs believed PSTs' skills had improved.

This article describes the outcomes of using a digital, simulation-based approach to supporting secondary mathematics and science preservice teachers (PSTs) in learning how to facilitate argumentation-focused discussions. The approach used a set of tools called the Online Practice Suite (OPS) that consists of two scaffolded digital teaching simulations: first, a lower intensity written response format called Teacher Moments,and second, a higher intensity interactive format called Avatar-Based Simulation. Similar studies have been conducted at the elementary level but have often relied on using a single simulation tool (e.g., Mikeska et al., 2023). A connected study on elementary-level PSTs who utilized the OPS found significant improvement in their ability to lead small and large group discussions, their beliefs about science and mathematics content teaching, and their feelings of overall preparedness (Mikeska et al., 2025).

The present study shifted focus to secondary PSTs and addressed the insufficiency of empirical research on such approaches at the secondary level (Dieker et al., 2014; Gundel et al., 2019). We hypothesized that PSTs' engagement in OPS would be closely linked to their learning and perception of the value of OPS and that their understanding of argumentation, beliefs, and ability to facilitate argumentation-focused discussions proficiently would improve (based on Chen et al., 2019; Faize et al., 2017; Luft et al., 2011; Mikeska et al., 2021; Pajares, 1992; Solar et al., 2021). We operationalized this focus in the following four research questions.

1. What evidence is there that (a) PSTs perceived the scaffolded suite of simulations to be useful and that it contributed to their learning, and (b) teacher educators perceived that PSTs learned from the suite of simulations?

2. What evidence is there of PSTs learning to better understand argumentation-based discussion?

3. What evidence is there of change in PSTs' beliefs about content instruction and their preparedness to teach?

4. What evidence is there of PSTs' improvement in ability to facilitate argumentation-focused discussions?

The unique contributions of this study include its focus on studying a scaffolded approach of using OPS at the middle school level (grades 6-8), where such tools are scant, and our efforts to design new instrumentation to measure learning.

Literature Review

We start with an overview of the literature to explain the field's perspective on the importance of involving K-12 students in argumentation and discuss the focal content areas. Then, we describe ways the ability of PSTs to lead argumentation-focused discussions may be influenced by their comprehension of argumentation, teaching skills, and personal beliefs. Finally, we discuss literature suggesting that PSTs' ability to facilitate challenging teaching practices may be enhanced by their participation in approximations of practice (Grossman et al., 2009), such as digital simulations.

Student Practice and Content

Argumentation as a Student Practice in Mathematics and Science

Argumentation is a core and effective content-learning practice for elementary (Arias et al., 2017; Fishman et al., 2017) and secondary (Conner et al., 2014; Osborne et al., 2013) students to enhance their understanding of scientific and mathematical phenomena (Bogar, 2019; Chen et al., 2019; Staples & Newton, 2016). Engaging in mathematical or scientific argumentation enables students to partake actively in the epistemic process while interacting dialogically with peers and teachers (Fishman et al., 2017; Ford, 2012). Argumentation entails social interactions (Berland & Reiser, 2009; Ford, 2008; Osborne & Patterson, 2011) among students as they collaborate and exchange ideas with their peers, resulting in an enhanced understanding of scientific and mathematical concepts (Makar et al., 2015; Noviyanti et al., 2019; Staples & Newton, 2016; Tsai et al., 2015).

The *Next Generation Science Standards* (NGSS Lead States, 2013) underscored the importance of "engaging in argument from evidence" as an essential scientific practice to understand the nature of science and construct scientific knowledge. Similarly, the *Common Core State Standards* (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) emphasized students' ability to "construct viable arguments and critique the reasoning of others" to enhance students' mathematical proficiency.

In a classroom setting, engaging in scientific and mathematical argumentation is characterized by the assertion and defense of claims (Grooms et al., 2018; Norris et al., 2007; Noviyanti et al., 2019; Rumsey &

Langrall, 2016; Staples & Newton, 2016). Students employ scientific data (Arias et al., 2017; Duschl, 2008; Osborne & Patterson, 2011) and mathematical derivation (Hanna, 2000; Peretz, 2006; Singletary & Conner, 2015). In counterarguments, individuals evaluate and critique their peers' claims, drawing on scientific justification (González-Howard & McNeill, 2019; Sampson et al., 2013) and mathematical reasoning (Conner et al., 2014; Hoffman et al., 2009) to convince others. Disciplinary differences in how the argument is enacted within mathematics and science tend to reflect the nature of the argument; the use of data and evidence is more prominent in science, where mathematics focuses on logical justification.

Across both disciplines, students may struggle to engage unless they understand the nature of argumentation (McDonald & Kelly, 2012; McNeill & Krajcik, 2007). As argumentation is a complex process encompassing structure and dialogic components, students often encounter challenges grasping its nature (Faize et al., 2017; Schwaighofer et al., 2017). For example, one study demonstrated that elementary students faced difficulties in decision-making, including when constructing evidence to support claims, amongst other practices (Choi et al., 2015). Students require guidance from teachers to understand the nature of argumentation and to improve their skills to engage in productive scientific and mathematical argumentation (Choi et al., 2015; Lampert & Cobb, 2003; Sampson et al., 2013). Argumentation is relevant across a wide range of topics in science and mathematics. Our study focuses on two, proportional reasoning and heat transfer, for middle school students (ages 11 to 14 in the United States school system).

Mathematics Content: Proportional Reasoning

Proportional reasoning is described as an important mathematical concept in middle school (Hilton et al., 2016; NGSS Lead States, 2013) and "consists of the ability to discern a multiplicative relationship between two quantities as well as the ability to extend the same relationship to other pairs of quantities" (Lamon, 2007, pp. 637–638). Proportional reasoners exhibit an intuitive grasp of covariation, enabling them to make meaningful comparisons among quantities (Fielding-Wells et al., 2014; Lesh et al., 1988).

Recognizing covariance signifies a fundamental understanding of how the change in one quantity is related to the change in the other (Lamon, 2007).Proportional reasoning is necessary to understand a variety of mathematical contexts, such as probabilistic contexts and measurement (Ayan & Isiksal-Bostan, 2019; Tizón-Escamilla & Burgos, 2023) and other subject areas like art, economics, geography, and science (Hilton & Hilton, 2016; Howe et al., 2011).

Science Content: Heat Transfer

The concept of heat transfer is a fundamental topic for secondary science students' learning, as it presents a critical scientific process that is essential in natural materials and advanced technology artifacts (Brown, 2011; Hatzikraniotis et al., 2010; Herrmann-Abell & DeBoer, 2018). To gain an understanding of how heat transfers between objects, one must have knowledge of the relevant constructs (e.g., temperature, mass, and energy) and mechanisms (i.e., conduction, convection, and radiation) involved (Brown, 2011; Clary, 2017; Lubben et al., 1999; Widodo et al., 2020). Understanding the principles of heat transfer is relevant from a scientific standpoint and has applications in daily life, such as cooking (Purnama et al., 2023; Widodo et al., 2020).

Connecting Argumentation and Content

Participating in argumentation-focused discussions not only engages students in disciplinary practice but also helps students make sense of key disciplinary concepts. Through argumentation, students can deepen their understanding of proportional reasoning or heat transfer as they share ideas with one another, make claims, support arguments with evidence and reasoning (science) or justification (mathematics), and engage in critique. This requires that the teacher has strong content knowledge (Batanero et al., 2014; Begolli et al., 2021; Keeley, 2012; Lee, 2014) and knows how to carefully facilitate a discussion focused on argumentation — a challenging teaching practice.

Teacher Understanding, Pedagogical Skills, and Beliefs Related to Argumentation

Teacher Understanding

To facilitate productive argumentation, teachers must develop their epistemic understanding of the nature of argumentation (Charalambous, 2015; Chen et al., 2019). The epistemic understanding of argumentation refers to knowing what, why, and how to argue to generate intended knowledge, which is tied to the specific content being discussed and the context in which the argument occurs (Chen et al., 2019). Teachers' epistemic understanding of argumentation can be influenced by several key factors, including their prior knowledge, past experiences, alternative conceptions, and individual interpretations of specific ideas (see also Rudsberg et al., 2013). Hence, teachers' content knowledge and understanding serve as a foundation for productive argumentation, enabling them to scaffold their students' argumentation and support them in conceptual sense-making (see also Henderson et al., 2018).

In a case study, researchers explored how teachers' mathematical understanding guided elementary students to focus on proportional reasoning and improve their responses in argumentation (Fielding-Wells et al., 2014). In another study focusing on elementary science, researchers demonstrated that teachers' science content knowledge improved their skill in posing open-ended questions in argumentation, thereby providing opportunities for students to elaborate and explain their reasoning (Fishman et al., 2017).

Teachers' content knowledge and understanding of argumentation form a necessary but not sufficient foundation for promoting productive student argumentation. Scholars suggest that it is essential to consider other aspects that can influence teachers' pedagogy, including their perception of situations, actions (Leder et al., 2002; Mavhunga & Rollnick, 2016; Nation & Feldman, 2022; Nguyen & Tran, 2023), and instructional decision-making (Charalambous, 2015; Hunt et al., 2023; Sleep & Eskelson, 2012).

Teacher Pedagogical Skill

Teachers' pedagogical skills are critical in supporting student argumentation (Conner et al., 2014; Solar et al., 2021). To promote students' engagement in argumentation, teachers can establish a stimulating learning environment for student-centered discussions, provide clear instructions, scaffold dialogic interaction, and support students' collaborative reasoning. As a result, students form a community of learners with intellectual authority and ownership (Faize et al., 2017; González-Howard et al., 2017; Mikeska & Howell, 2020; Rumsey & Langrall, 2016). Furthermore, teachers can employ strategies that support discussion more generally, such as encouraging students' participation, encouraging students to articulate their thinking (Mikeska & Howell, 2020), addressing alternative conceptions, and posing content-related questions (Solar et al., 2021). These approaches benefit both students' conceptual understanding and their use of disciplinary practices.

However, a study by Staples and Newton (2016) revealed that secondary mathematics teachers found it challenging to engage students in argumentation due to the content demand and assessment procedure. In another study, secondary science teachers expressed concern about their lack of knowledge and skills in designing and engaging students in productive argumentation discussions (Sampson & Blanchard, 2012). These findings underscore the need for more pedagogical support for teachers.

Teacher Beliefs

Pajares (1992), in his groundbreaking paper on teachers' beliefs, suggested that if teacher learning or change is desired, researchers need to pay attention to beliefs. Further, Enochs and Riggs (1990) indicated that instructional change begins with changes in beliefs. Beliefs are personal constructs and play a pivotal role in shaping teachers' instructional decisions, classroom management, and representation of concepts (Luft et al., 2011).

We define beliefs as "embodied conscious and unconscious ideas and thoughts about oneself, the world, and one's position in it developed through membership in various social groups, which are considered by the individual to be true" (Cross, 2009, p. 326). A teacher's belief is rooted in their personal experiences and context (Schoenfeld, 1998), which can evolve and change (Thompson, 1992) when challenged or exposed to new information or perspectives (Wall, 2016). Beliefs are, however, often so strongly grounded that they are steadfast even when presented with contrasting evidence (Nespor, 1987), making them somewhat difficult to shift.

For teachers to implement an instructional innovation that involves learning to teach in a different way, such as facilitating argumentation-focused discussion, they need not only the necessary knowledge but also productive beliefs regarding the innovation (Charalambous, 2015; Li et al., 2020). Teachers' pedagogical beliefs can influence their academic targets, teaching values, and teaching practices (Khader, 2012; Shamim et al., 2022). Scholars have explored how to reshape teachers' practices by prioritizing a focus on their underlying beliefs (Cross, 2009; Juniati & Siswono, 2019; Wilkins, 2008) — a strategy referred to as the "beliefs-first approach." This method often involves immersing teachers in situations that create cognitive dissonance, where they encounter information that challenges their preexisting beliefs (Festinger, 1957). In this study's context, the immersion is in digital simulation environments. Understanding that beliefs can impact instructional practice, Smith et al. (2014) designed a measure to capture both teachers' beliefs and their potential to impact instructional practice in science.

Connected to teachers' beliefs is their self-efficacy which can be evident in their practice (Weirtheim & Leyser, 2002; Woolfolk et al., 2009). Teachers' self-efficacy encompasses both their belief in their own and in their students' abilities to achieve. When teachers have a stronger belief in their own and their students' abilities, they can demonstrate greater commitment to student success (Tschannen-Moran & Hoy, 2001).

Preparing Teachers to Facilitate Argumentation Discussions Through Approximations

Practice-based teacher education focuses on positioning essential teaching practices, referred to as "core" or "ambitious" practices, earlier in the teacher development trajectory (Ball & Forzani, 2011; Grossman et al., 2009; Windshitl et al., 2020). These practices are central to effective teaching, support student learning, and emphasize a student-centered pedagogical focus. Practice-based teacher education asserts that PSTs should have repeated and systematic learning opportunities to decompose the key components of instructional practice, analyze representations of the practice in varied contexts, and approximate its enactment (Dutro & Cartun, 2016; Grossman et al., 2009; Matsumoto-Royo, & Ramírez-Montoya, 2021). Our study focused on the last of these, approximating where the core practice is facilitating argumentation-focused discussions.

Approximations of practice offer a way for PSTs to experiment and enhance their ability to respond to demanding pedagogical scenarios (Ghousseini, 2017; Grossman et al., 2009; Kavanagh, Metz et al., 2020). By nature, approximations of practice are not real teaching in terms of context, setting, and level of risk but maintain the responsive and interactive nature of core teaching practices (Grossman et al., 2009, p. 2078). Examples of approximations include PSTs participating in structured field experiences (Shaughnessy et al., 2019), role-playing with their peers (Benedict-Chambers et al., 2020), and engaging in face-to-face or digital simulated teaching (e.g., Mikeska & Howell 2020).

Notably, digital approximations have become increasingly popular, in that they offer PSTs a controlled and manageable classroom-like experience (Campbell et al., 2020; Cohen et al., 2020; Howell & Mikeska, 2021; Lampert et al., 2013). Moreover, these digital approximations provide PSTs with opportunities for trial and error (Uluay, 2021) without impacting real classroom students. Although some digital approximations include elements such as the use of avatars, these approximations are not designed to be gamified environments. Such environments aim to address both participant learning goals and motivation, often through game-like elements such as competition and levels (Kalogiannakis et al., 2021) or through individualized adaptations (Zourmpakis et al., 2023).

Like gamified digital environments, digital approximations are designed to help participants meet learning goals (e.g., PSTs learning how to facilitate an argumentation discussion). Unlike gamification, the simulated digital approximations common in PST education and that we used in this study focus not on motivation and the inclusion of game elements but rather on creating a close-to-reality, reduced-complexity environment, in which PSTs have situated opportunities to meet those learning goals.

It is beyond the scope of this paper to list and detail all the available forms of digital approximation. Rather, we briefly describe two approaches used in the present study, which we will describe in greater detail in the method section. The first is Teacher Moments, which is a digital platform individual PSTs can use to practice generating questions for and responses to students (e.g., how they might use questioning to encourage students to critique one another's ideas; Hillaire et al., 2022; Reich, 2022). The second is Mursion®, a commercial simulation provider, which allows PSTs to hone their teaching skills by interacting with virtual student avatars (Dalinger et al., 2020; Mikeska et al.., 2021; Mikeska et al., 2022; Straub, 2018). Interacting with virtual student avatars provides PSTs with a more authentic and immersive engagement than peer rehearsals where the adults act as K-12 students (Dieker et al., 2017; Mikeska et al., 2023; Straub et al., 2014, 2015).

Within the context of approximations, PSTs attempt ambitious practices, and research has indicated that such experimentation is associated with PST progress and enhanced student performance and achievement (Arias & Davis, 2017; Ball & Forzani, 2009; Brownell et al., 2019; Campbell & Elliott, 2015; Forzani, 2014). Notably, approximations focused on argumentation have been used in elementary science and mathematics teacher education (Fishman et al., 2017; Kazemi & Wæge, 2015), where research has demonstrated a significant impact on improving PSTs' skills (Mikeska et al., 2025; Lee & Lee, 2023; McNeill et al., 2016). Other studies have demonstrated improvement in PSTs' abilities to engage students in critical thinking, constructing knowledge, critiquing ideas, and refining understanding of scientific and mathematical phenomena (Benedict-Chambers et al., 2020; Berland & Reiser, 2009; Ford, 2012; Osborne & Patterson, 2011).

Methods

In this section, we describe the study context, the specific Teacher Moments and Avatar-Based Simulation tasks, the participants, the instruments, and the analysis.

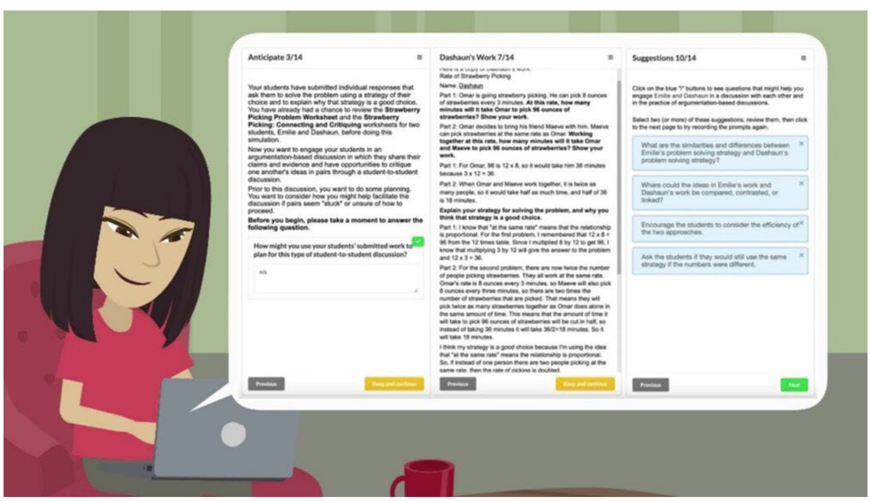
Study Context: The OPS Project Secondary Level Tasks

In this study, each TE integrated two simulated approximations of practice — Teacher Moments and the Avatar-Based Simulation — into their semester-long methods course (either in mathematics or science methods) for PSTs to practice facilitating argumentation-focused discussions (https://tarheels.live/onlinepracticesuite/). TEs prepared PSTs before and then debriefed PSTs after each simulation. No standardization of course elements was required beyond the incorporation of the two approximations and the inclusion of a preparation and debrief activity, and there was considerable variability across the set of teacher educators (TEs) in other course activities, as well as where the course fell within the university program and whether PSTs were concurrently or previously engaged in fieldwork (see Lottero-Perdue et al., 2024 for full descriptions of this variability). In what follows, we describe and introduce the mathematics and science tasks our project created in the two platforms, which are available in the project data repository at https://nyu.databrary.org/volume/1110.

Teacher Moments and the Teacher Moments Tasks

Teacher Moments is an online platform wherein PSTs can engage in short scenario-based experiences to practice responding to difficult moments in teaching (Hillaire et al., 2022; Reich, 2022). See Figure 1 for a depiction of a PST engaging in Teacher Moments. Teacher Moments is freely available through the MIT Teaching Systems Lab, and support is provided to TEs to learn to create their own tasks. We developed the Teacher Moments tasks to provide an opportunity for the PSTs to practice skills related to facilitating an argumentation-focused discussion prior to engaging in the more complex Avatar-Based Simulation task, described in the next section. Teacher Moments focused on questions to elicit student argumentation, setting the stage for the Avatar-Based Simulation, in which PSTs then have to contend with how students respond to such questions.

Figure 1
Preservice Teacher Engages With an Online Teacher Moments Scenario by Typing Responses



The Teacher Moments mathematics task is called "Rate of Strawberry Picking" and addresses proportional reasoning. The task vignette describes how two students have each individually solved a problem related to the rate of strawberry picking. The PST's role is to generate questions to encourage the pair to share and critique one another's ideas. After the PST submits initial responses, four suggestions are provided for their consideration (e.g., What are the similarities and differences in the students' problem-solving strategies?). Then, the PST has an opportunity to try responding a second time to the prompts.

The Teacher Moments science task is called "Keeping the Heat" and addresses heat transfer. In the task vignette, the class measured the temperature of hot chocolate in a paper and foam cup over time, observing the temperature change over two points in time. Each student was asked to write a claim supported with evidence about which cup was better at keeping the hot chocolate warm. The PST sees the work of two students and tries to engage these two students in an argumentation-focused discussion. As in the mathematics task, after the PST provides initial responses, they are provided suggestions and an opportunity to revise.

The Avatar-Based Simulation and the Avatar-Based Simulation Tasks

For this study, the Avatar-Based Simulation utilized the Mursion® middle school classroom of five student avatars shown in Figure 2. Each PST facilitated an up to 20-minute argumentation-focused discussion with the avatars. The avatars are able to respond in real-time because there is a human-in-the-loop (an actor called, in Mursion® terminology, a simulation specialist), who voices and operates all five avatars and undergoes extensive training to ensure consistency across different PSTs. This training is essential (Bondie et al., 2021), as it helps the simulation specialists learn to embody each avatar, manipulate their facial expressions and gestures, give each an identifiable voice, and shift from avatar to avatar quickly. Training also covers task contents, the student ideas in the tasks, how those ideas may change during the discussion, and how to respond to different approaches that PSTs might use.

The Avatar-Based Simulation mathematics task, "Hungry Huskies," addresses proportional reasoning, specifically the different rates at which two dogs eat food and how long it will take for the dogs to eat a certain weight of food. Prior to the argumentation-focused discussion, two student groups (a group of two and another group of three) solved the problem, identifying unit rates. Each group's work and reasoning had both strengths and ways in which it could be improved. The PST's goal is to support students in comparing approaches and reaching a consensus about how unit rates may be combined when solving a problem that includes more than one rate.

The Avatar-Based Simulation science task, called "Keep it Cold," addresses heat transfer and involves an investigation that demonstrates how cold water in a paper cup warms up at a faster rate than cold water in a foam cup. Again, students are assumed to have worked in groups, and each group produced a written model and explanation prior to the discussion. The PST's goal is to engage student avatars in critiquing one another's models and have them agree on a consensus model.

Participants

We first recruited six TEs — three mathematics TEs and three science TEs — via professional connections, selecting based on interest and availability and maximizing variability in their personal and institutional demographic characteristics. TEs incorporated one use of the Teacher Moments activity and one use of the Avatar-Based Simulation activity into their normal coursework for all enrolled PSTs for the fall 2022 semester-long course in either mathematics or science methods.

Figure 2
Preservice Teacher Engages in a Small Group Discussion With Five Student Avatars

We invited all PSTs (48) in each of the TEs' methods courses (three mathematics sections and three science sections) to participate in the study during the fall 2022 semester, which required engaging in pre/post measures and postactivity surveys. Of these, 39 agreed to participate (83% rate of participation). Table 1 includes demographic and background information for PSTs and TEs. Details of how the TEs integrated the activities into their courses are described in Lottero-Perdue et al. (2024).

Instruments

The instruments that we used to address the four research questions, as well as the item types used for each, are shown in Appendix A. Specific item questions that we used in this study are listed in Appendix B, which also identifies the instrument and number of responses coded for each. Pre and post timepoints occurred at the beginning and end of the semester, respectively; surveys administered after the Teacher Moments and Avatar-Based Simulation varied depending on the TE's preference for implementation dates.

Table 1
Demographic and Background Information for PSTs and TEs

Characteristic	All PSTs (*n* = 39) *n* (%)	All TEs (*n* = 6) *n* (%)
Gender		
Female	27 (69%)	4 (67%)
Male	11 (28%)	2 (33%)
Prefer not to respond	1 (3%)	0 (0%)
Race/Ethnicity [a]		
Asian or Asian American	4 (10%)	0 (0%)
Black/African American	1 (3%)	0 (0%)
Hispanic/Latino	2 (5%)	1 (17%)
White/Caucasian	35 (90%)	5 (83%)
Year in Teacher Education		
Year 1 – Bachelor's Program	0 (0%)	-
Year 2 – Bachelor's Program	1 (3%)	-
Year 3 – Bachelor's Program	14 (36%)	-
Year 4 – Bachelor's Program	19 (49%)	-
Year 5 – Bachelor's Program	1 (3%)	-
Year 1 – Master's Program	2 (5%)	-
Year 2 – Master's Program	2 (5%)	-
Current Major(s) [a]		
Elementary Education	3 (8%)	-
Middle School Education Math	3 (8%)	-
Middle School Education Science	8 (21%)	-
Secondary Education Math	18 (46%)	-
Secondary Education Science	15 (38%)	-
Special Education	3 (8%)	-
Natural Sciences (e.g., Biology, Chemistry, Earth, Physics)	6 (15%)	-
Previous Experience Using Simulated Classrooms		

Characteristic	All PSTs (n = 39) n (%)	All TEs (n = 6) n (%)
Never	30 (77%)	-
Once or twice	5 (13%)	-
Three or four times	4 (10%)	-
Five or more times	0 (0%)	-
Previous Experience Leading Mathematics Discussions With Secondary Students		
Not at all	25 (64%)	-
A little (once or twice)	7 (18%)	-
Some (three or four times)	6 (15%)	-
A lot (five or more times)	1 (3%)	-
Previous Experience Leading Science Discussions With Secondary Students		
Not at all	23 (59%)	-
A little (once or twice)	7 (18%)	-
Some (three or four times)	6 (15%)	-
A lot (five or more times)	3 (8%)	-
Cumulative GPA		
3.5 or above	25 (64%)	-
3.0-3.49	7 (18%)	-
2.5-2.99	7 (18%)	-

[a] Participants could select more than one response.

Note. PST = preservice teacher; TE = teacher educator. A total of 39 PSTs and six TEs participated in the study.

Pre- and Postsurveys

This survey aimed to assess PSTs' beliefs about effective mathematics or science instruction using the Mathematics Beliefs Instrument (MBI; Smith et al., 2012) or the Teacher Beliefs about Effective Science Teaching questionnaire (TBEST; Smith et al., 2014). These are three-factor scales built from Likert survey items (48 items for MBI and 21 items for TBEST). The surveys also included Likert, yes/no, and open-response questions that we developed by adapting questions used on the National Survey of Science and Mathematics Education (Banilower et al., 2018) to measure PSTs' beliefs and understanding about discussion and argumentation and their preparedness to teach mathematics and science (content and pedagogy). Additional questions were included in the postsurvey, in which PSTs were asked to reflect retrospectively on how their beliefs, understandings, or preparedness changed over the semester.

Pre- and Postenacted Measure

Given at the same time points as the pre- and postsurveys, PSTs completed identical mathematics or science pre- and postenacted measures. The purpose of the enacted measure was to have an assessment of PST performance before and after OPS with respect to facilitating argumentation discussions, rather than relying solely on self-report data. The pre- and postenacted measures were implemented within the Teacher Moments platform, using a different vignette than either the Rate of Strawberry Picking or Keeping the Heat task. Task features were similar to the Teacher Moments task described earlier, although abbreviated because it was used for measurement purposes (e.g., the second-round questions were omitted).

Teacher Moments and Avatar-Based Simulation Task Surveys

PSTs and TEs completed a survey after each of the Teacher Moments and Avatar-Based Simulation learning cycles, asking them to reflect on the learning cycle, the simulation, and on discussion and argumentation. Our focus in this study was on items in the surveys related to what PSTs learned from engaging in Teacher Moments or Avatar-Based Simulation and the utility or appropriateness of Teacher Moments or Avatar-Based Simulation for PST teacher education.

Post-TE Interview

At the end of the semester, TEs participated in a semistructured interview, asking them to reflect on their instructional decisions and PST learning. We used two interview responses to triangulate survey data around the extent to which TEs felt their PSTs made progress with respect to argumentation-focused discussions during the semester.

Data Analysis

In this section, we analyzed (1) pre/post-belief measures, (2) open responses and closed (Likert and yes/no) answers to survey and interview questions, and (3) our enacted measure.

Pre/Post Beliefs Scales (Included in the Pre/Post Surveys)

First, for each PST, we averaged the values given in response to the 4-point Likert scale items related to each construct (general, content, and argumentation preparedness) to produce a single score for each PST at each time point. Then, we utilized a paired samples *t*-test to compare means from the pre- and postmeasurement time points for these three constructs. For the TBEST and MBI, we calculated scores on each subscale per the TBEST and MBI guidelines for each PST at both time points and then used paired *t*-tests to compare from the pre- to the postmeasurement time point.

Survey Responses

First, we used descriptive statistics to share frequencies of responses — both number and percentage — for closed and Likert scale items. Then, for each of the open response items, two or more project team members engaged in iterative qualitative analysis. Four of these items, identified in Appendix B, involved the use or adaptation of a coding tree from a previous project (Mikeska et al., 2023). The remainder involved developing and applying new codes. Percentage agreement and one-way average measures ICC are reported in Appendix B when double coding was used. With most large response sets (77 or more responses), approximately 10% of the responses were group coded, and 20% were double coded, with the remainder individually coded, provided the percent agreement was over 80% or ICC of 0.7 or higher. For some smaller response sets, only group or double coding was employed.

Pre/Post Enacted Measure

We developed a scoring guide for the pre- and postenacted measure comprised of seven indicators, which were derived from an Avatar-Based Simulation scoring rubric used in other parts of the larger research project and implemented in prior research (Go Discuss Project, 2021). The indicators were as follows:

1. Proximity to substantive content;
2. Invites students to share or explain ideas;
3. Asks students to compare and contrast, make connections to past experiences, and/or link ideas;
4. Encourages direct student interactions;
5. Invites students to persuade others about their point of view;
6. Invites students to justify their own ideas or others' ideas; and
7. Invites students to critique their own ideas or others' ideas.

The scoring team identified four responses in each of the two tasks that captured relevant information of these indicators. Other responses that served more to set the stage for the participant were not included in the scoring. For each PST, each indicator was scored using a 4-point scale (0-3 points) accounting for both the frequency and quality of the indicators across the four scored responses.

Quality generally referred to the degree to which the response used language that was aligned with the indicator definition, while frequency referred to how many of the four responses contained evidence for that indicator. A score of 0 indicated the absence of that indicator in the PST's response. A score of 1 indicated a lack of quality and consistency but not a complete absence. A score of 2 indicated a lack in either quality or consistency but not both, and a score of 3 indicated the presence of quality and consistency in the PST's response. This resulted in a total score ranging from 0 to 21.

Our team completed group coding on 20% of the PSTs' responses, and then three raters coded an additional 20% of the responses in pairs, reaching pairwise percent agreement of 78%, 77%, 75%, and one-way average ICCs of 0.95, 0.95, and 0.94, respectively. The remaining 60% of the PSTs' responses were coded individually, although the raters convened at the end of the scoring process to discuss and address any uncertain scores**.** We then used paired *t*-tests to compare PSTs' scores from the pre- to the postmeasurement time point on each indicator.

Analysis Across the Instruments

Some items on the PST and TE Post Survey and TE Interview asked for reflection back, making them a form of retrospective self-report. Other items, such as asking PSTs to define argumentation, were asked multiple times at different time points, allowing for comparison. In general, our pragmatic approach is to note where data support strong conclusions and to complement with analysis that is less able to support strong conclusions but able to give a general trend among multiple triangulated sources of weaker evidence (see limitations).

In some cases, such as for Research Question (RQ) 1, triangulation with TE data was limited simply to examining responses systematically for confirming or disconfirming evidence. While we did check systematically for differences between the mathematics and science cohorts, we found few differences, and this was not a primary focus of the study. Therefore, we report mathematics and science cohort results separately only for the two beliefs scales that were completely different.

Results

This section is a summary of results by research question, drawing on both qualitative and quantitative data to provide a picture, for each research question, of the collection of evidence available around improvement. The discussion then highlights the most striking results, including those relating to teachers' beliefs, and describes limitations of the study.

Results for RQ 1: PSTs' Perceived Usefulness of the Suite

Our first research question addressed the evidence relating to PSTs' and TEs' perceptions of the usefulness of OPS and of what PSTs learned. We answer this by first exploring PSTs' self-reported main takeaways from the OPS, how appropriate PSTs and TEs felt that the OPS was for PST education, and the teacher roles PSTs believed that the OPS supported.

PSTs' Main Takeaways from OPS

Of the 77 PST responses to the open-ended survey question asking their main takeaway from the OPS (Table 2), the most frequently applied code was the awareness of teacher practices (n = 21 code applications; 27%), applied to statements such as "Be specific in your question while trying to get students to critique each other's work" (science PST, Teacher Moments survey), highlighting the importance of teacher questioning as a practice. A close second (19; 25%) was PST statements reflecting learning about the importance of preparing for a discussion, applied to statements such as "The better your preparation is, the more smoothly your lesson or discussion will be" (mathematics PST, Avatar-Based Simulation survey).

Table 2
PSTs' and TEs' Responses on PST Learning From OPS

Code	Description	All PST Responses (n = 77) n (%)	All TE Responses (n = 12) n (%)
Awareness of Teacher Practices	Describes an awareness of the importance of particular teacher practices (e.g., questioning, probing, teacher talk moves).	21 (27%)	7 (58%)
Preparation Importance	Describes a general importance of preparing for the discussion and/or interaction with students.	19 (25%)	5 (42%)
Other	Describes a different main takeaway or learning that is not one of the previous codes.	11 (14%)	0 (0%)
Practice	Describes the use of the simulated discussion activity as a way to practice particular instructional skills, improve instructional skills, and gain experience for teaching real students or a way to gather information about one's practice.	10 (13%)	1 (8%)
Importance of Knowing and Responding to Students	Describes the importance of knowing and/or responding to the students and their thoughts, ideas, perspectives, and/or behaviors.	7 (9%)	2 (17%)
Recognize Discussion Importance/ Challenge	Describes the importance and/or challenge of including discussion in a math/science classroom. The response could be describing the overall role the teacher plays in facilitating the discussion.	7 (9%)	1 (8%)
Vague/Not Related	Describes a main takeaway or learning that is too vague to categorize.	3 (4%)	2 (17%)
Importance of PSTs' Emotional Regulation as Learners	Describes using and/or developing their socioemotional skills (e.g., confidence, patience) and/or the importance of emotional regulation.	2 (3%)	0 (0%)
Flexible	Describes being flexible when preparing for and/or when leading the discussion. For example, being prepared to have unexpected questions/thoughts come up during the discussion.	2 (3%)	1 (8%)

Code	Description	All PST Responses (*n* = 77) *n* (%)	All TE Responses (*n* = 12) *n* (%)
None	Describes no main takeaway or learning from the simulated discussion experience.	0 (0%)	0 (0%)
Evidence of shift in beliefs		26 (34%)	7 (58%)
Evidence of shift in skill	Describes a main takeaway or learning that is about becoming better at the work of leading an argumentation-focused discussion, or a component of it.	19 (25%)	6 (50%)
Evidence of understanding argumentation	Describes a main takeaway or learning that speaks to having learned about what argumentation is or coming to understand it better.	1 (1%)	0 (0%)
Total Number of Code Applications		182	43
Note. Multiple codes could be applied to each response. PST = preservice teacher; TE = teacher educator.			

TEs' responses to the corresponding survey item triangulated these findings, with their report of what the PSTs learned most often coded as awareness of teacher practices (7; 58%) and preparation importance (5; 42%). For example, Dawn (pseudonyms used throughout), a mathematics TE, stated, "I think they learned that in order to have students share and critique ideas, they have to provide specific prompts with a common goal (for example, ask the students to find similarities and differences in their solutions rather than just explain their solutions to each other)," reflecting a focus on the practices of sharing and critiquing ideas.

Each identified takeaway was also coded concurrently with respect to showing evidence of one or more of our three main constructs (*understanding*, *beliefs*, and *skill*; Table 2), if applicable. The most frequently categorized takeaways were beliefs, with 34% of codes (26) applied to PST responses in that area. For example, one PST stated "My main takeaway would be the true power of argument-based discussion. I am very surprised as to how valuable it is in a classroom setting" (mathematics PST, Teacher Moments survey). Another mentioned the importance of the teaching practice of eliciting student ideas, saying, "It is critical to elicit ideas from students, to encourage their thinking and sharing" (science PST, Avatar-Based Simulation survey).

About 25% of codes (n = 19) were categorized as a shift in PSTs' skills, while nearly none (1; 1%) were categorized as shifts in understanding of argumentation. Again, TEs' report of PSTs' learning was distributed similarly, most frequently naming beliefs (7; 58%) and skills (6; 50%) and with no responses coded as mentioning the change in PST understanding of argumentation.

Appropriateness of the OPS Simulations and Likeliness to Recommend

Overall, most of the PSTs' (*n* = 75; 97% of responses) and TEs (11; 92% of responses) responses to the Likert scale item indicated that they believed the simulations (Teacher Moments and Avatar-Based Simulation) were *somewhat* or *very appropriate* to incorporate into teacher education methods courses. Most PSTs responded to the yes/no survey item indicating they would recommend using the Teacher Moments simulation (35; 92% of PSTs) or Avatar-Based Simulation (32; 82% of PSTs) in a future section of their secondary mathematics or science methods course.

How the OPS Supported Their Role as a Teacher

Table 3 shows the findings for the main patterns in the PSTs' perceptions about how the Teacher Moments and Avatar-Based Simulation activities supported their role as teachers. Most PSTs indicated that simulations provided a way for them to practice (*n* = 51; 66%) facilitating a discussion or to practice specific discussion skills, such as asking probing questions or eliciting student ideas. PSTs also described how the practice afforded by the simulations helped them to be better prepared for teaching in real classrooms, to apply what they learned from the experience to their teaching in real classrooms or to gather information about their own practice. Collectively, this finding suggests that the PSTs were likely to recognize how the simulations provided them with opportunities to build their teaching skills.

Table 3
PSTs' Survey Responses About How the Simulations Supported Their Role as a Teacher

Code	Code Description	All PST Responses (*n* = 77) *n* (%)	TM Responses (*n* = 38) *n* (%)	ABS Responses (*n* = 39) *n* (%)
Practice	Describes how the simulations supported them: by providing a way for them to practice facilitating a discussion or specific discussion skills, in being prepared for teaching in (real) classrooms, in applying what they learned to their teaching (in real) classrooms, and/or in gathering	51 (66%)	23 (61%)	28 (72%)

Code	Code Description	**All PST Responses (n = 77) n (%)**	**TM Responses (n = 38) n (%)**	**ABS Responses (n = 39) n (%)**
	information about one's practice.			
Understand discussion or recognize discussion importance/ challenge	Describes how the simulation supported them in building their understanding about/awareness of discussion, its importance and/or the associated challenges.	8 (10%)	3 (8%)	5 (13%)
Identify Areas for Growth	Describes how the simulations supported them in identifying areas for growth/weakness.	4 (5%)	3 (8%)	1 (3%)
Supported Learning/ Teaching about Content	Describes how the simulations supported them in learning/teaching about conceptual ideas or topics.	4 (5%)	1 (3%)	3 (8%)
Supported Self-Reflection	Describes how the simulations supported them in reflecting on their practice (either in the moment or in the future).	4 (5%)	1 (3%)	3 (8%)
Awareness of Teacher Practices	Describes how the simulation supported them in building an awareness of the importance of particular teacher practices (e.g., questioning, probing, teacher talk moves).	3 (4%)	1 (3%)	2 (5%)
Emotional Regulation	Describes how the simulations supported them in using and/or developing their socioemotional skills (e.g., confidence, patience) and/or the importance of emotional regulation.	3 (4%)	2 (5%)	1 (3%)

Code	Code Description	All PST Responses (n = 77) n (%)	TM Responses (n = 38) n (%)	ABS Responses (n = 39) n (%)
Other	Describes a different way in which the experiences supported their role as a teacher that is not one of the other codes (e.g., helping students find their voice; using problems with a real-life connection).	3 (4%)	3 (8%)	0 (0%)
Vague/ Unclear	Response is too vague or unclear to understand.	3 (4%)	3 (8%)	0 (0%)
Did Not Support Role as Teacher	Describes how the simulations did not support them in their role as a teacher.	1 (1%)	1 (3%)	0 (0%)
Evidence of Shift in Skill	Describes how the simulations supported them in becoming better at the work of leading an argumentation-focused discussion, or a component of it.	29 (38%)	17 (45%)	12 (31%)
Evidence of Understanding Argumentation	Describes how the simulation supported them in learning about what argumentation is or coming to understand it better.	2 (3%)	1 (3%)	1 (3%)
Evidence of Shift in Beliefs	Describes how the simulation supported them in shifting their beliefs about the importance of discussion and/or argumentation, about instruction, or about efficacy.	4 (5%)	0 (0%)	4 (10%)

Note. PST = preservice teacher; TM = Teacher Moments; ABS = Avatar-Based Simulation.

Findings also showed that approximately a third of the PST responses (29; 38%) noted how the simulations supported them in becoming better at the skill of leading an argumentation-focused discussion or a component of this type of discussion. For example, one PST noted how the "Teachers Moments experience supported me in my role as a mathematics teacher because it taught me how to teach students important skills like critiquing each other's work" (mathematics PST, Teacher Moments survey). Another PST stated how it "helped me learn more about facilitating questions as a science teacher" (science PST, Teacher Moments survey).

Results for RQ 2: PST Understanding of Argumentation-Focused Discussions

Descriptions of PST-reported characteristics of high-quality discussions focused on argumentation at each of the pre- and postmeasurement time points are shown in Table 4. Around half of the PSTs reported students providing rationale, evidence, or reasoning to justify their own or others' claims, ideas, or strategies as a characteristic of high-quality discussions focused on argumentation at both the pre- (n = 21; 54%) and postmeasurement time points (n = 19; 50%). Similarly, PSTs commonly mentioned promoting sharing or disagreement in a safe space by attending to specific normsand/or habits of mind at the premeasurement time point (21; 54%), with a slight but not statistically significant decrease at the postmeasurement time point (17; 45%). Other reported characteristics are listed in Table 4.

Table 4
Comparing PSTs' Survey Responses on Understanding of Argumentation Pre-Post

Code	**Description** **Describes argumentation-focused discussions as a way…**	**All Pre PST Responses (n = 39) n (%)**	**All Post PST Responses (n = 38) n (%)**
Justification	…for students to explain their thinking, provide rationales and/or provide evidence or reasoning to justify their and/or others' claims, ideas, and/or strategies.	21 (54%)	19 (50%)
Attending to Norms	…to promote sharing and/or disagreement in a safe space by attending to specific norms and/or habits of mind.	21 (54%)	17 (45%)
Sharing/ Discussion	…for students to share and/or discuss their claims, ideas, and/or strategies.	12 (31%)	15 (39%)
Evaluate/ Critique	…for students to debate amongst themselves through critiquing, evaluating, and/or comparing/contrasting one another's work.	10 (26%)	12 (32%)

Code	**Description Describes argumentation-focused discussions as a way…**	**All Pre PST Responses (*n* = 39) *n* (%)**	**All Post PST Responses (*n* = 38) *n* (%)**
Learning/ Understanding	…to promote a broad sense of student conceptual learning or understanding.	5 (13%)	8 (21%)
Student Led/Teacher as Guide	…for students to lead or facilitate the discussion and/or for teachers to help guide a student-led discussion.	3 (8%)	6 (16%)
Questions	…to include questions asked by either the teacher or other students.	1 (3%)	8 (21%)
Consensus	…for students to build towards and/or come to a consensus.	1 (3%)	4 (11%)
General Engagement	…to promote student engagement and/or keeping students on task.	1 (3%)	2 (5%)
Other	…that does not fall within any of the previously listed codes.	1 (3%)	2 (5%)
Attending to Thinking	…to promote different ways of thinking (e.g., critical thinking, thinking systematically, thinking deeply, etc.).	1 (3%)	1 (3%)
Vague	…that is too vague to categorize.	1 (3%)	0 (0%)
Persuade	…for students to persuade or convince one another that they are correct and/or another student is incorrect or correct.	0 (0%)	1 (3%)
Note. PST = preservice teacher.			

Twenty of the 38 PSTs (53%) self-reported a change in understanding of argumentation and were asked to describe the nature of that change and to what they attributed the change. Descriptions of change in understanding are shown in Table 5. The most prevalent description (9; 45% of codes applied) was recognition of the importance of argumentation, which arguably might be described as more of a shift in understanding something about argumentation's role in student learning than one in understanding what argumentation is.

The second most frequent description (6; 30%) was of shifts that articulated a change in how the PST envisioned argumentation, such as "Argumentation is not to convince others who is right or wrong but to discover and explore different ideas and make them relevant to the learning" (mathematics PST, postsurvey). Most of the PST attributions of change in understanding were to activities associated with the OPS (15; 75%), including simulations and associated assignments, with no clear attributions to activities outside of the OPS.

Table 5
PST Descriptions of Change in Understanding Over the Semester

Code	Description Reports Change Associated With…	All PST Responses (n = 20) n (%)
Importance	…recognizing the importance of argumentation.	9 (45%)
Vision of argumentation	…their understanding of what argumentation is, or developing a vision of it in instruction, or details that vision.	6 (30%)
Benefits	…seeing the payoff or effectiveness of engaging in argumentation for the students.	4 (20%)
Facilitate argumentation	…knowledge or confidence in facilitating argumentation or provides details about doing so.	4 (20%)
Increased exposure	…increased exposure to argumentation discussion or a that this is the first time they considered argumentation.	3 (15%)
Student-focused	…argumentation as related to a student-centered classroom or less teacher-driven instruction.	1 (5%)
Total Code Applications		27

Note. PST = preservice teacher

Results for RQ 3: PST Change in Beliefs: Content Instruction and Teaching Preparedness

As shown in Table 6, on the pre/post scales corresponding to general, content-specific, and argumentation-specific preparedness, we observed significant increases in beliefs about preparedness on all three scales. On the TBEST (science) instrument, all three factors shifted in the direction that was anticipated (increase for Factor 1, decrease for Factors 2 and 3), but the change was statistically significant only for Factor 2. For the MBI (mathematics) instrument, all three factors shifted in the direction that was anticipated (increase), and the shift was statistically significant on all factors.

PST responses about how important they believe it is for students to have opportunities to engage in argumentation showed an increase pre-post, with the percentage describing it as *very important* increasing from 24 (62%) to 33 (87%) and the percentage describing it as only *somewhat unimportant* (the lowest level selected) dropping from 2 (5%) to 0 (0%). By the end of the semester, all (38; 100%) categorized it as at least *somewhat important*.

Table 6
Pre/Post Comparison of PSTs Surveys Responses on Beliefs

Factor/ Aspect	Pre		Post		Paired Samples *t*-test					Effect Size
	Mean	*SD*	Mean	*SD*	Mean Diff.	*SD* Diff.	*t*	*df*	Sig. (2-tailed)	
Changes in Beliefs (Preparedness): All (*n* = 38)										
General Preparedness	2.64	0.62	2.98	0.48	-0.34	0.66	-3.14	37	**0.003[a]**	-0.51
Content Preparedness	2.40	0.69	3.05	0.59	-0.65	0.77	-5.19	37	**<.001[a]**	-0.84
Argumentation Preparedness	2.62	0.72	3.23	0.44	-0.62	0.68	-5.55	37	**<.001[a]**	-0.90
Changes in Beliefs (Effective Instruction)										
Science (TBEST) (*n* = 18)										
Factor 1: Learning Theory Aligned	5.42	0.34	5.49	0.45	-0.07	0.49	-0.57	17	0.575	-0.14
Factor 2: Confirmatory	3.71	0.85	3.29	0.96	0.42	0.69	2.57	17	**0.020[a]**	0.61
Factor 3: All Hands on All the Time	3.59	0.81	3.11	1.18	0.48	0.98	2.09	17	0.052	0.49
Math (MBI) (*n* = 20)										
Factor 1: Curriculum	3.11	0.43	3.33	0.46	-0.22	0.26	-3.84	19	**0.001[a]**	-0.86
Factor 2: Learner	3.11	0.37	3.30	0.48	-0.18	0.33	-2.52	19	**0.021[a]**	-0.56
Factor 3: Teacher	3.47	0.44	3.62	0.39	-0.15	0.29	-2.33	19	**0.031[a]**	-0.52

[a] Bold text indicates the indicators that demonstrated statistically significant change.
Note. PST = preservice teacher; MBI = Mathematics Beliefs Instrument; TBEST = Teacher Beliefs about Effective Science Teaching. These measures were completed at both the pre- and postmeasurement time points. However, only 38 PSTs completed the measure at both time points. Therefore, we only included the 38 PSTs who had the full set of pre and post data in this comparison.

Our coding reflected that PSTs' ideas about the importance of argumentation were ones that we also considered to be critical. Most frequently mentioned (Table 7) were students sharing ideas with peers and hearing ideas from others (54; 18%), learning to justify ideas (51; 17%), and learning content more deeply by having learned it via argumentation (40; 13%). There were no statistically significant changes over the four time points at which the question was asked, pre- and postmeasurement, indicating that while individual PSTs' responses may have varied, there were no overall trends in the types of reasons reported.

Table 7
PST Report of Why Argumentation Is Important by Time Point

Mode Description	**Pre-Responses (*n* = 39) *n* (%)**	**TM Responses (*n* = 38) *n* (%)**	**ABS Responses (*n* = 39) *n* (%)**	**Post Responses (*n* = 38) *n* (%)**	**Total Across Time Points (*n* = 154) *n* (%)**
Sharing ideas with peers and hearing others' ideas	7 (18%)	17 (45%)	17 (44%)	13 (34%)	54 (18%)
Justification of claims/ideas/strategies or explanation of reasoning	14 (36%)	14 (37%)	9 (23%)	14 (37%)	51 (17%)
Support learning of content by learning it via argumentation	12 (31%)	10 (26%)	10 (26%)	8 (21%)	40 (13%)
Develop argumentation skills including those that will apply outside of the classroom	7 (18%)	10 (26%)	12 (31%)	4 (11%)	33 (11%)
Appreciate multiple perspectives on an issue, idea, or approach	10 (26%)	8 (21%)	5 (13%)	6 (16%)	29 (10%)
Other	3 (8%)	5 (13%)	3 (8%)	11 (29%)	22 (7%)
Apply content knowledge in the context of building arguments	8 (21%)	3 (8%)	5 (13%)	4 (11%)	20 (7%)
Have opportunities to evaluate or critique others' arguments	6 (15%)	3 (8%)	3 (8%)	8 (21%)	20 (7%)
Reflect on students' own learning or thinking	3 (8%)	2 (5%)	4 (10%)	4 (11%)	13 (4%)

Mode Description	Pre-Responses (*n* = 39) *n* (%)	TM Responses (*n* = 38) *n* (%)	ABS Responses (*n* = 39) *n* (%)	Post Responses (*n* = 38) *n* (%)	Total Across Time Points (*n* = 154) *n* (%)
Engage in argumentation because it is a critical disciplinary practice	4 (10%)	1 (3%)	4 (10%)	3 (8%)	12 (4%)
Vague	0 (0%)	0 (0%)	2 (5%)	1 (3%)	3 (1%)

Eighteen of the 38 PSTs (47%) who completed the postsurvey self-reported a change in their perception of the importance of argumentation, triggering follow-up questions to describe the change. Descriptions of change in perception are shown in Table 8. The most prevalent description (9; 50%) was seeing the benefit/payoff of argumentation in terms of how it supports students, such as "I now believe that argumentation in science class can be very beneficial to students and can help them apply evidence" (science PST, presurvey). The second and third most frequent descriptors were better understanding what argumentation is (5; 28%) in ways that elevated its importance and, more generally, simply recognizing its importance (5; 28%), with no additional detail provided. The majority of descriptions given of the activities to which PSTs attributed change were activities associated with the OPS (12; 67%), including simulations and associated assignments; three were attributed to activities outside of the OPS (3; 17%), and three were not classifiable (6; 17%).

Table 8
Self-Reported Descriptions of Change in PSTs Perception of Argumentation

Code	Description	All PST Responses (*n* = 18)[a]
Recognize importance - see the payoff/effectiveness (in what it does for students)	Reports change associated with seeing the payoff or effectiveness of engaging in argumentation for the students. *It is sufficient to mention the benefit to students without specifying the nature of the benefit.*	9 (50%)
Recognize importance - understanding makes me see its important	Reports change due to a difference in their understanding of argumentation. *This includes referencing a lack of understanding initially, even if the more robust understanding is not specified.*	5 (28%)
Recognize importance - general	Reports change associated with recognizing the importance of argumentation without explicitly going into detail.	5 (28%)

Code	**Description**	**All PST Responses (n = 18)[a]**
Increased exposure/considered for the first time	Reports change associated with increased exposure to argumentation discussion or a that this is the first time they considered argumentation. *References to doing activities during the course counts as reference to exposure.*	3 (17%)
Recognize importance - became more connected to content	Reports change associated with becoming more connected to the mathematics/science content.	2 (11%)
Other/Vague	Reports change that is not one of the previous codes or that is vague/unclear.	1(6%)

[a] Only 18 of the 38 PSTs who completed the postsurvey responded “yes” to the question “Have your perspectives about the importance of argumentation for secondary students changed over the course of the semester?” on the postsurvey; therefore, only 18 PSTs were given the following up prompt to describe how their perspectives changed.

Results for RQ 4: PST Improvement in Ability to Facilitate Argumentation-Focused Discussions

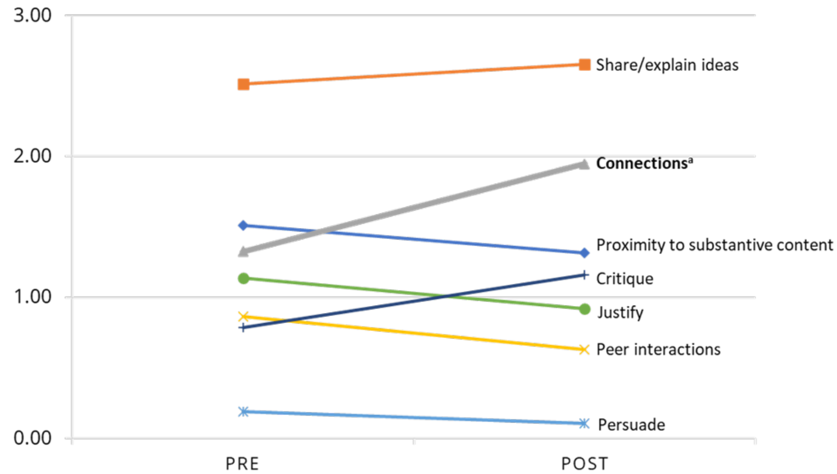
Pre/Post Teacher Moments Scores

Findings from the scoring of the pre- and postenacted measure reflected neither directionality consistent with expectations (an increase on all indicators) nor strong evidence of improvement from pre to post (Figure 3). When examining responses across all PSTs who had a full set of pre- and postmeasurement data (n = 37), only one indicator, Connections, had a statistically significant increase from pre to post. This measure was completed at both the pre and postmeasurement time points. However, only 37 PSTs completed the measure at both time points. Therefore, we only included the 37 PSTs who had the full set of pre and postdata in this comparison. No decreases reached the level of significance. The data achieved normality only for the overall score, suggesting that interpretation of the indicator scores requires caution.

Self-Report of Improvement in Skill

When asked how their level of preparedness to implement elements of high-quality discussion focused on argumentation changed over the course of the semester, the most frequent PST response was “more prepared by the end of the semester” (n = 115; 61%), followed by “a lot more prepared by the end of the semester (52; 27%), and then by “about the same level of preparedness by the end of the semester” (23; 12%). None (0; 0%) reported feeling less prepared.

Figure 3
Scores by Indicator for the Pre/Post Teacher Moments Enacted Measure



[a] Indicator that demonstrated statistically significant change (Connections).

Slightly more than half of PST responses (89; 53%) responding to what they attributed this change in level of preparedness reported supports for their improvement that were coded as clearly associated with the OPS experience. Other PSTs reported supports that were coded as clearly not associated with OPS (17; 10%), such as their fieldwork or student teaching experiences, which may have taken place concurrently or prior to the semester of study. For example, one PST stated, "…my field experience helped me be more prepared" (science PST, postsurvey). The remaining PSTs reported supports were too vague to categorize as being clearly associated or not associated with the OPS (62; 37%); for example, "knowing everyone needs to speak their opinion" (science PST, post survey).

Teacher Educator Report of Improvement in PSTs Skill

When asked about how much progress PSTs made on elements of argumentation-focused discussion on a 4-point Likert scale (*no progress*, *minimal progress*, *some progress*, *a lot of progress*), the most frequent response for any of the five dimensions was *some progress* (22; 73%). The next most frequent response wa *a lot of progress* (5; 17%) and *minimal progress* (3; 10%). None of the TEs responded with *no progress* (0; 0%).

We used TEs' interview responses, in which TEs were asked to expand on the responses they had given in the most recent survey (Avatar-Based Simulation time point), to triangulate the assumption that the progress over the semester was connected to the use of the OPS. Three TEs — Angela, Barry, and Francisco — noted general improvement. For instance, Francisco offered, "I would attribute [the PST's progress to the] … cycles of teacher moments and Avatar-Based Simulation; I think that's what helped them develop [their] skills."

Others noted specific components of OPS or its implementation as drivers of improvement. For example, Dawn made specific reference to preparation, enactment, and debrief: "Just all of the preparation, the actual implementation, the debrief really helped them," Christine attributed progress only to the preparation component, and Erin attributed it to the enactment component. Of note, while all TEs attributed improvement to the OPS, they also mentioned other factors. For example, Christine suggested that one reason the PSTs' progress was notable was because they started with a great deal of room to grow, stating that "part of their progress comes from starting … at such a low level of preparation to facilitate an argument-based discussion."

During the interview, TEs were also asked the following question: "How does your PSTs' progress this semester compare to their progress in other semesters when you've taught this course?" All of the TEs responded to the question by suggesting that their OPS-semester PSTs made more progress than past semesters, four of the TEs suggesting that it was a lot more progress than in prior semesters. Erin, for example, compared talking about discussions with PSTs in past semesters with actually preparing for and facilitating discussions this semester, saying, "But just … doing it on your own, having to plan, having to be on the spot and go through with the experience, I think it pushed them to see the value in it because they had the experience," helping them to be "so much further along and in moving towards discussion and argumentation than students I've taught in the past."

Dawn said that the OPS PSTs created better lesson plans, in general, that included better questions than past groups of PSTs. Overall, the TEs suggested that the OPS and the multiple facets of the experience — including not only planning for but facilitating discussions — enabled PSTs to become better prepared for facilitating argument-based discussions.

Discussion

RQ1: PST Learning and Perception of Value

PSTs and TEs were aligned in reporting what they believed the PSTs had learned from each simulation. PST responses most frequently reflected learning that we categorized as beliefs, although their descriptions of how the simulations worked to support them gravitated toward discussion of how the opportunity to practice was useful and that they practiced improving skills. Although there was some variability and a small number of dissenting responses, PSTs, and TEs overwhelmingly perceived the simulations to be appropriate inclusions in methods courses and recommended their future use. This is consistent with the literature indicating that approximations can support PST learning and are generally welcomed within programs of teacher preparation (Ghousseini, 2017; Lee & Lee, 2023; Mikeska et al., 2023).

Of note, we found that takeaways were most often ones we would consider to be shifts in beliefs or in skill, with little direct report of shifts in understanding of argumentation. There are limitations to this analysis, of course. Identifying a single main takeaway does not preclude there having been other learnings; it simply captures the one PSTs considered to have been the strongest. It is notable, however, that so many PSTs reported shifts in belief as a main takeaway, given that the literature generally conceptualizes simulations as focused on skill improvement. This suggests that calls to attend to beliefs when investigating interventions on skill (e.g. Cross Francis et al., 2023; Pajares,1992) still ring true, even in contexts such as digital simulations of practice.

PSTs articulated an understanding of how the Teacher Moments and the Avatar-Based Simulation activities were supporting their role as a teacher that was consistent with the project's understanding of how the activities would do so, as indicated by their focus on practicing discussion and specific subcomponents of it, and on argumentation-focused discussion, in particular. They also pointed out that the practice helped prepare them for real teaching and to understand how they were doing as teachers, which is consistent with the intended design of the OPS and with prior research indicating that learning requires productive beliefs about the innovation supporting learning (Charalambous, 2015). Concluding that learning is taking place does not, of course, logically require that the learner understand how and why the learning is taking place, but this consistency between self-reported mechanisms of learning and intended mechanisms triangulates the self-report from PSTs and TEs, supporting the claim that PSTs learned from these experiences.

RQ2: PST Understanding of Argumentation

There was no strong evidence of improvement in understanding of argumentation between pre- and postmeasurement time point. This may be because PSTs' understandings of argumentation were already relatively strong, as the coding suggests that PSTs reported characteristics at both time points that were generally consistent with the project definition of argumentation, such as providing rationales, evidence, and reasoning to justify their own or others' strategies, ideas, or claims.

Just over half of the PSTs did self-report change in understanding of argumentation, but our coding of the nature of that change suggests that what PSTs categorized as shifts in understanding argumentation might be better described as better understanding the importance of argumentation to support student learning. This may suggest, using Chen et al.'s (2019) language, that improvement in understanding of argumentation was more focused on the *why* than on the *what*, a distinction that our survey questions did not allow us to disentangle.

This is consistent with the findings from RQ1 that ideas we categorized as beliefs were most frequently reported as the main learning or takeaway for PSTs and with the relatively stronger evidence of shifts in beliefs reported under RQ3. As was the case for RQ1, where PSTs reported change, the activities to which they attributed that change were largely associated with the OPS and not with unrelated coursework or experiences, providing some evidence that this deepened understanding of the importance of argumentation was associated with the OPS.

RQ3: PST Change in Beliefs: Content Instruction and Teaching Preparedness

The evidence of shifts in PSTs' beliefs is the strongest of the three constructs investigated, corroborating the finding under RQ1 that the main takeaways PSTs reported were most often beliefs. Most scales on the quantitative beliefs measures showed statistically significant shifts in the anticipated direction despite the smaller sample sizes on the separate science and mathematics measures, and all showed directionally consistent change. PSTs' responses indicate that they largely believed argumentation was important at the start of the semester, but their responses also show an increase from pre- to postmeasurement timepoint. This is consistent with the findings of RQ2 that showed self-reported shifts in PST understanding to be more consistent with shifts we would classify as a deeper understanding of its importance.

Prior literature on changing teachers' beliefs suggests immersion in environments that challenge preexisting beliefs can stimulate change (Cross, 2009). It could be that the suite's focus on argumentation provided a focus and a context in which the PSTs were considering their beliefs about its importance. The reasons given for argumentation's importance were consistent with the project's definition of argumentation, including justification, supporting the development of content understanding and argumentation skills, and considering multiple perspectives. Just under half of PSTs reported change in their perception of the understanding of argumentation, with the majority attributing that change to the OPS.

RQ4: PST Improvement Facilitating Argument-Focused Discussions

The results for RQ4 are mixed and, therefore, more challenging to interpret. Our effort in this study, marking a divergence from the approach taken in the prior study with elementary teachers (Mikeska et al., 2025), was to design a lighter lift pre/post assessment that would still reveal evidence of growth, an effort that we judged to have sufficient potential payoff in simplicity and scalability to outweigh the known risk of designing a new instrument without established validity evidence. We did not, however, observe clear evidence to support the conclusion that PSTs improved in the target skill between pre- and postmeasurement timepoints.

Self-report from PSTs and TEs provides stronger evidence of improvement in PST skill. PST survey responses indicate that most PSTs felt more prepared by the semester's end to facilitate argumentation-focused discussions, with the sources of that improvement clearly associated with the OPS in just over half of responses. TE reports corroborate this finding, with most responses indicating that PSTs made some or a lot of progress, and their attributions of this progress were nearly entirely related to the OPS.

Interview responses further supported the claim that PST learning was associated with the OPS by comparing it to similar classes taught in past semesters, with all of the TEs reporting their perception that PSTs made more progress in developing their skills to lead argumentation-focused discussion this semester than in prior semesters. As noted under RQ1, PSTs' self-identified takeaways were associated with improvement in skill in 25% of responses, and their descriptions of how OPS supported their role as a teacher were associated with skill improvement in 38% of responses. While our quantitative measure did not support the conclusion, PSTs and TEs clearly believed that the PSTs' skills had improved.

Given the mixed evidence, it is worth digging into the evidence we have, even if weak, from the quantitative measure and what it suggests for future refinement or design. The sample size was small, but the lack of consistent directionality suggests that the instrument did not function as designed. PSTs and TEs reported self-improvement consistent with reported improvement on associated studies that found a strong pre/post change in skill (Mikeska et al., 2023), but our pre/post measure did not confirm that improvement.

The fact that some indicators showed predicted growth suggests some potential, perhaps as of yet unrealized, for such an instrument to function well. Given the data we have, it is impossible to draw strong conclusions about whether the instrument accurately demonstrates a lack of growth or is inadequate to pick up growth, but we suspect, based on the relative strength of the self-reports data, that either the instrument, the scoring process, or the format may be at issue.

For example, one might hypothesize that the written response format is too far from the enactment of argumentation to demonstrate a true measure of skill and that an interactive assessment like the one used in Mikeska et al. (2025) is needed. Put another way, and perhaps just as scholars theorize that teachers best learn through approximating practice (Ghousseini, 2017; Grossman et al., 2009; Kavanagh, Conrad et al., 2020), we might need to measure change in practice through a closer approximation of the practice itself. Alternately, one might hypothesize that our scoring, which attended both to the quality of response and to the consistent presence of markers across four different open field entries, might be out of alignment with the ways teachers respond across multiple response boxes and what assumptions they make about the appropriate length of each individual response, and that a different scoring process might have yielded more useful information. All of these are hypotheses worth exploring in future work, as the goal of measuring a complex construct in a relatively straightforward way is still worth pursuing even if our initial exploratory efforts did not accomplish it.

Limitations

Limitations have been noted throughout but are worth reiterating here. Most critically, our sample sizes are generally small, particularly where mathematics and science instrumentation are different. This limits our quantitative findings, although it also points to the strength of the findings that are statistically significant despite this limitation. Given the smaller sample sizes and the exploratory nature of both the approach and the tools available to measure its success, we conceptualized the study more as an initial effort to begin to build evidence to support and understand the approach in this context. To strengthen the evidence base, in future research, it would be important to address similar questions with larger and more diverse samples, including preservice and in-service teachers across varied teacher education programs and school districts.

As discussed previously, a significant known risk of the study was the use of a newly enacted measure using the Teacher Moments platform, which was untested prior to the study. Our evidence also relies on self-report data, which has inherent weaknesses, including the possibility of priming PSTs to see OPS in a more positive light because we asked them about it frequently. We did, however, make substantial efforts to address these by drawing on multiple data types, quantitative and qualitative, and across both the PST and TEs groups to triangulate findings more generally. We did not see evidence of findings reliant on self-report data contradicting other findings, except the mixed results for the fourth research question. In future research, it would be useful to further refine the assessment tools, especially the newly enacted measure used in this study, and examine their use with larger and more diverse teacher populations.

Another inherent limitation of the study design is that the OPS was implemented as a whole, making it impossible to disentangle the impact of various components. For example, even in cases where the evidence supports the OPS having a strong impact on PST thinking and PSTs and TEs cite the simulations and practice they provided as the source of that impact, we cannot dismiss competing hypotheses, such as the impact being as much a result of the sustained focus on argumentation than on the specific configuration of tools and approaches that made up the OPS. Likewise, we cannot dismiss the possibility that activities and experiences unrelated to the implementation of the OPS contributed to or drove impact for individuals, as PSTs are not just enrolled in a single course but are involved in larger collections of experiences intended to support their development as teachers.

Conclusion

This study was exploratory in nature, designed to begin building an evidentiary basis for the claim that approaches like the OPS, focused on the use of a scaffolded set of simulations related to a single teaching practice, can over time support secondary mathematics and science PSTs' learning of that practice. Our results were mixed, with strong evidence that PSTs believed they benefited from the experience and that both TEs and PSTs saw value in the approach. The strongest evidence, however, accumulated around just one of the main constructs investigated: PST beliefs.

This finding is particularly intriguing because most work on simulation design has focused on the development of skills, not beliefs, with the idea that "practice" quite literally helps one become better at doing the thing being practiced (Grossman et al., 2009). It is also intriguing because beliefs are generally seen as influencing instructional decisions (e.g., Luft et al., 2011) rather than instructional moves influencing beliefs.

Our results suggest either a different directionality or a codevelopment that is more complex. PSTs in our study appear to have had a reasonable understanding of argumentation, to begin with, and they may have become better at facilitating argumentation-based discussions (the evidence is inconclusive), but they clearly did develop a deeper appreciation for argumentation and its importance. This is especially notable given that beliefs are notoriously difficult to change; the use of an experiential endeavor like simulations is not a typical approach to changing beliefs, but the evidence suggests it might be one worth considering. We hope that future research will continue to build on the emergent evidence base, continue to press for more efficient instrumentation while learning from the limitations reported, and continue to explore the relationship between skills and beliefs in PST learning.

Author Note

This study was supported by a grant from the National Science Foundation (Grant number 2037983). The opinions expressed herein are those of the authors and not the funding agency.

References

Arias, A. M., & Davis, E. A. (2017). Supporting children to construct evidence-based claims in science: Individual learning trajectories in a practice-based program. *Teaching and Teacher Education*, *66*, 204–218. https://doi.org/10.1016/j.tate.2017.04.011

Arias, A. M., Smith, P. S., Davis, E. A., Marino, J. C., & Palincsar, A. S. (2017). Justifying predictions: Connecting use of educative curriculum materials to students' engagement in science argumentation. *Journal of Science Teacher Education*, *28*(1), 11–35. https://doi.org/10.1080/1046560X.2016.1277597

Ayan, R., & Isiksal-Bostan, M. (2019). Middle school students' proportional reasoning in real real-life contexts in the domain of geometry and measurement. *International Journal of Mathematical Education in Science and Technology*, *50*(1), 65–81. https://doi.org/10.1080/0020739X.2018.1468042

Ball, D. L., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, *60*(5), 497–511. https://doi.org/10.1177/0022487109348479

Ball, D. L., & Forzani, F. M. (2011). Building a common core for learning to teach: And connecting professional learning to practice. *American Educator*, *35*(2), 17–21.

Banilower, E. R., Smith, P. S., Malzahn, K. A., Plumley, C. L., Gordon, E. M., & Hayes, M. L. (2018). *Report of the 2018 NSSME+*. Horizon Research, Inc.

Batanero, C., Arteaga, P., Serrano, L., & Ruiz, B. (2014). Prospective primary school teachers' perception of randomness. In E. Chernoff & B. Sriraman (Eds.), *Probabilistic thinking: Presenting plural perspectives* (pp. 345–366). Springer. https://doi.org/10.1007/978-94-007-7155-0_19

Begolli, K. N., Dai, T., McGinn, K. M., & Booth, J. L. (2021). Could probability be out of proportion?. Self-explanation and example-based practice help students with lower proportional reasoning skills learn probability. *Instructional Science*, *49*, 441–473. https://doi.org/10.1007/s11251-021-09550-9

Benedict-Chambers, A., Fick, S. J., & Arias, A. M. (2020). Preservice teachers' noticing of instances for revision during rehearsals: A comparison across three university contexts. *Journal of Science Teacher Education*, *31*(4), 435–459. https://doi.org/10.1080/1046560X.2020.1715554

Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education*, *93*(1), 26–55. https://doi.org/10.1002/sce.20286

Bogar, Y. (2019). Synthesis study on argumentation in science education. *International Education Studies*, *12*(9), 1–14. https://doi.org/10.5539/ies.v12n9p1

Bondie, R., Mancenido, Z., & Dede, C. (2021). Interaction principles for digital puppeteering to promote teacher learning. *Journal of Research on Technology in Education*, *53*(1), 107–123. https://doi.org/10.1080/15391523.2020.1823284

Brown, P. (2011). Teaching about heat and temperature using an investigative demonstration. *Science Scope*, *35*(4), 31–35.

Brownell, M. T., Benedict, A. E., Leko, M. M., Peyton, D., Pua, D., & Richards-Tutor, C. (2019). A continuum of pedagogies for preparing teachers to use high-leverage practices. *Remedial and Special Education*, *40*(6), 338–355. https://doi.org/10.1177/0741932518824990

Campbell, M. P., & Elliott, R. (2015). Designing approximations of practice and conceptualising responsive and practice-focused secondary mathematics teacher education. *Mathematics Teacher Education and Development*, *17*(2), 146–164.

Campbell, M. P., Baldinger, E. E., & Graif, F. (2020). Representing student voice in an approximation of practice: Using planted errors in coached rehearsals to support teacher candidate learning. *Mathematics Teacher Educator*, *9*(1), 23–49. https://doi.org/10.5951/MTE.2020.0005

Charalambous, C. Y. (2015). Working at the intersection of teacher knowledge, teacher beliefs, and teaching practice: A multiple-case study. *Journal of Mathematics Teacher Education*, *18*, 427–445. https://doi.org/10.1007/s10857-015-9318-7

Chen, Y. C., Benus, M. J., & Hernandez, J. (2019). Managing uncertainty in scientific argumentation. *Science Education*, *103*(5), 1235–1276. https://doi.org/10.1002/sce.21527

Choi, A., Klein, V., & Hershberger, S. (2015). Success, difficulty, and instructional strategy to enact an argument-based inquiry approach: experiences of elementary teachers. *International Journal of Science and*

Mathematics Education, *13*, 991–1011 (2015). https://doi.org/10.1007/s10763-014-9525-1

Clary, R. (2017). Beat the Heat!. *Science Scope*, *40*(8), 34–41.

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

Conner, A., Singletary, L. M., Smith, R. C., Wagner, P. A., & Francisco, R. T. (2014). Teacher support for collective argumentation: A framework for examining how teachers support students' engagement in mathematical activities. *Educational Studies in Mathematics*, *86*(3), 401–429. https://doi.org/10.1007/s10649-014-9532-8

Cross, D. (2009). Creating optimal mathematics learning environments: Combining argumentation and writing to enhance achievement. *International Journal of Science and Mathematics Education*, *7*(5), 905–921. https://doi.org/10.1007/s10763-008-9144-9

Cross-Francis, D., Kaur Bharaj, P., Habib., K., Hinden, A., & Hong, J. (2023, April 13-16). *"Library to Courtroom": Influence of Problem Solving and Refutation Texts on Pre-Service Teachers' Math-related Beliefs* [Paper presentation]. Annual Conference of the American Educational Research Association, Chicago, IL, United States.

Dalinger, T., Thomas, K. B., Stansberry, S., & Xiu, Y. (2020). A mixed reality simulation offers strategic practice for pre-service teachers. *Computers & Education*, *144*, 103696. https://doi.org/10.1016/j.compedu.2019.103696

Dieker, L. A., Rodriguez, J. A., Lignugaris/Kraft, B., Hynes, M. C., & Hughes, C. E. (2014). 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

Dieker, L. A., Hughes, C. E., Hynes, M. C., & Straub, C. (2017). Using simulated virtual environments to improve teacher performance. *School-University Partnerships*, *10*(3), 62–81.

Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education*, *32*(1), 268–291. https://doi.org/10.3102/0091732X07309371

Dutro, E., & Cartun, A. (2016). Cut to the core practices: Toward visceral disruptions of binaries in PRACTICE-based teacher education. *Teaching and Teacher Education*, *58*, 119–128. https://doi.org/10.1016/j.tate.2016.05.001

Enochs, L. G., & Riggs, I. M. (1990, April 8-11). *Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale* [Paper presentation]. Annual Meeting of the National Association for Research in Science Education, Atlanta, GA, United States.

Faize, F. A., Husain, W., & Nisar, F. (2017). A critical review of scientific argumentation in science education. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(1), 475–483. https://doi.org/10.12973/ejmste/80353

Festiger, L. (1957). *A theory of cognitive dissonance*. Row, Peterson, New York.

Fielding-Wells, J., Dole, S., & Makar, K. (2014). Inquiry pedagogy to promote emerging proportional reasoning in primary students. *Mathematics Education Research Journal*, *26*, 47–77. https://doi.org/10.1007/s13394-013-0111-6

Fishman, E. J., Borko, H., Osborne, J., Gomez, F., Rafanelli, S., Reigh, E., Tseng, A., Million, S., & Berson, E. (2017). A practice-based professional development program to support scientific argumentation from evidence in the elementary classroom. *Journal of Science Teacher Education*, *28*(3), 222–249. https://doi.org/10.1080/ 1046560X.2017.1302727

Ford, M. (2008). Disciplinary authority and accountability in scientific practice and learning. *Science Education*, *92*(3), 404–423. https://doi.org/10.1002/sce.20263

Ford, M. J. (2012). A dialogic account of sense-making in scientific argumentation and reasoning. *Cognition and Instruction*, *30*(3), 207–245. https://doi.org/10.1080/07370008.2012.689383

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

Ghousseini, H. (2017). Rehearsals of teaching and opportunities to learn mathematical knowledge for teaching. *Cognition and Instruction*, *35*(3), 188–211. http://dx.doi.org/10.1080/07370008.2017.1323903

GO Discuss Project (2021). *Scoring*. Qualitative Data Repository. https://doi.org/10.5064/F6NJU10I

González-Howard, M., McNeill, K. L., Marco-Bujosa, L. M., & Proctor, C. P. (2017). ‘Does it answer the question or is it French fries?’: An exploration of language supports for scientific argumentation. *International Journal of Science Education*, *39*(5), 528-547. https://doi.org/10.1080/09500693.2017.1294785

González-Howard, M., & McNeill, K. L. (2019). Teachers' framing of argumentation goals: Working together to develop individual versus communal understanding. *Journal of Research in Science Teaching*, *56*(6), 821–844. https://doi.org/10.1002/tea.21530

Grooms, J., Sampson, V., & Enderle, P. (2018). How concept familiarity and experience with scientific argumentation are related to the way groups participate in an episode of argumentation. *Journal of Research in Science Teaching*, *55*(9), 1264–1286. https://doi.org/10.1002/tea.21451

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, *111*(9), 2055–2100. https://doi.org/10.1177/016146810911100905

Gundel, E., Piro, J. S., Straub, C., & Smith, K. (2019). Self-efficacy in mixed reality simulations: Implications for preservice teacher education. *The Teacher Educator*, *54*(3), 244–269. https://doi.org/10.1080/08878730.2019.1591560

Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics*, *44*, 5–23. https://doi.org/10.1023/A:1012737223465

Hatzikraniotis, E., Kallery, M., Molohidis, A., & Psillos, D. (2010). Students' design of experiments: An inquiry module on the conduction of heat. *Physics Education*, *45*(4), 335–344. https://doi.org/10.1088/0031-9120/45/4/002

Henderson, J. B., McNeill, K. L., González-Howard, M., Close, K., & Evans, M. (2018). Key challenges and future directions for educational research on scientific argumentation. *Journal of Research in Science Teaching*, *55*(1), 5–18. https://doi.org/10.1002/tea.21412

Herrmann-Abell, C. F., & DeBoer, G. E. (2018). Investigating a learning progression for energy ideas from upper elementary through high school. *Journal of Research in Science Teaching*, *55*(1), 68–93. https://doi.org/10.1002/tea.21411

Hillaire, G., Waldron, R., Littenberg-Tobias, J., Thompson, M., O'Brien, S., Marvez, G. R., & Reich, J. (2022). Digital clinical simulation suite: Specifications and architecture for simulation-based pedagogy at scale. In *Proceedings of the Ninth ACM Conference on Learning@ Scale* (pp. 212–221).

Hilton, A., & Hilton, G. (2016). Proportional reasoning: An essential component of scientific understanding. *Teaching Science*, *62*(4), 32–42. https://search.informit.org/doi/10.3316/informit.475176477252213

Hilton, A., Hilton, G., Dole, S., & Goos, M. (2016). Promoting middle school students' proportional reasoning skills through an ongoing professional development programme for teachers. *Educational Studies in Mathematics*, *92*, 193–219. https://doi.org/10.1007/s10649-016-9694-7

Hoffman, B. L., Breyfogle, M. L., & Dressler, J. A. (2009). The power of incorrect answers. *Mathematics Teaching in the Middle School*, *15*(4), 232–238. https://doi.org/10.5951/MTMS.15.4.0232

Howe, C., Nunes, T., & Bryant, P. (2011). Rational number and proportional reasoning: Using intensive quantities to promote achievement in mathematics and science. *International Journal of Science and Mathematics Education*, *9*, 391–417. https://doi.org/10.1007/s10763-010-9249-9

Howell H., & Mikeska J. N. (2021). Approximations of practice as a framework for understanding the role of authenticity in simulations of teaching. *The Journal of Research on Technology in Education*, 53(1), 8–20. https://doi.org/10.1080/15391523.2020.1809033

Hunt, J., Duarte, A., Miller, B., Bentley, B., Albrecht, L., & Kruse, L. (2023). Teacher beliefs and perspectives of practice: Impacts of online professional learning. *Education Sciences*, *13*(1), 68–92. https://doi.org/10.3390/educsci13010068

Juniati, D., & Siswono, T. Y. E. (2019). Examining prospective teachers' belief and pedagogical content knowledge towards teaching practice in mathematics class: A case study. *Journal on Mathematics Education*, *10*(2), 185-202.

Kalogiannakis, M., Papadakis, S., & Zourmpakis, A.-I. (2021). Gamification in science education. A systematic review of the literature. *Education Sciences*, *11*(22), 1–36. https://doi.org/10.3390/educsci11010022

Kavanagh, S. S., Conrad, J., & Dagogo-Jack, S. (2020). From rote to reasoned: Examining the role of pedagogical reasoning in practice-based teacher education. *Teaching and Teacher Education*, *89,* 102991. https://doi.org/10.1016/j.tate.2019.102991

Kavanagh, S. S., Metz, M., Hauser, M., Fogo, B., Taylor, M. W., & Carlson, J. (2020). Practicing responsiveness: Using approximations of teaching to develop teachers' responsiveness to students' ideas. *Journal of Teacher Education*, *71*(1), 94–107. https://doi.org/10.1177/0022487119841884

Kazemi, E., & Wæge, K. (2015). Learning to teach within practice-based methods courses. *Mathematics Teacher Education and Development*, *17*(2), 125–145.

Keeley, P. (2012). Misunderstanding misconceptions. *Science Scope*, *35*(8), 12–15.

Khader, F. R. (2012). Teachers' pedagogical beliefs and actual classroom practices in social studies instruction. *American International Journal of Contemporary Research*, *2*(1), 73–92.

Lamon, S. J. (2007). Rational numbers and proportional reasoning: Toward a theoretical framework for research. *Second Handbook of Research on Mathematics Teaching and Learning*, *1*, 629–667.

Lampert M., & Cobb P. (2003). Communication and learning in the mathematics classroom. In J. Kilpatrick & D. Shifter (Eds.), *Research companion to the NCTM standards*. (pp. 237–249). National Council of Teachers of Mathematics.

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, 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://dx.doi.org/10.1177/0022487112473837

Leder, G. C., Pehkonen, E., & Törner, G. (Eds.). (2002). *Beliefs: A hidden variable in mathematics education?* (Vol. 31). Springer Science & Business Media.

Lee, C. K. (2014). A conceptual change model for teaching heat energy, heat transfer and insulation. *Science Education International*, *25*(4), 417–437.

Lee, J. E., & Lee, M. Y. (2023). How elementary prospective teachers use three fraction models: Their perceptions and difficulties. *Journal of Mathematics Teacher Education*, *26*(4), 455–480. https://doi.org/10.1007/s10857-022-09537-4

Lesh, R., Post, T. R., & Behr, M. (1988). Proportional reasoning. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades* (pp. 93–118). National Council of Teachers of Mathematics, Lawrence Erlbaum Associates.

Li, X., Song, N., Hwang, S., & Cai, J. (2020). Learning to teach mathematics through problem posing: Teachers' beliefs and performance on problem posing. *Educational Studies in Mathematics*, *105*, 325–347. https://doi.org/10.1007/s10649-020-09981-0

Lottero-Perdue, P. S., Howell, H., Kaur Bharaj, P., Park Rogers, M., Cross Francis, D., Maltese, A., & Shekell, C. Mikeska, J. N. (2024). Integrating simulated approximations of practice into methods courses: Teacher educators as designers. *Contemporary Issues in Technology and Teacher Education, 24*(4). https://citejournal.org/volume-24/issue-4-24/general/teacher-educators-as-designers-integrating-digital-simulations-into-methods-courses

Lubben, F., Netshisaulu, T., & Campbell, B. (1999). Students' use of cultural metaphors and their scientific understandings related to heating. *Science Education*, *83*(6), 761–774. https://doi.org/10.1002/(SICI)1098-237X(199911)83:6%3C761::AID-SCE7%3E3.0.CO;2-O

Luft, J. A., Firestone, J. B., Wong, S. S., Ortega, I., Adams, K., & Bang, E. (2011). Beginning secondary science teacher induction: A two-year mixed methods study. *Journal of Research in Science Teaching*, *48*(10), 1199–1224. https://doi.org/10.1002/tea.20444

Matsumoto-Royo, K., & Ramírez-Montoya, M. S. (2021). Core practices in practice-based teacher education: A systematic literature review of its teaching and assessment process. *Studies in Educational Evaluation*, *70*, 101047. https://doi.org/10.1016/j.stueduc.2021.101047

Makar, K., Bakker, A., & Ben-Zvi, D. (2015). Scaffolding norms of argumentation-based inquiry in a primary mathematics classroom. *ZDM*, *47*, 1107–1120. https://doi.org/10.1007/s11858-015-0732-1

Mavhunga, E., & Rollnick, M. (2016). Teacher-or learner-centred? Science teacher beliefs related to topic specific pedagogical content knowledge: A South African case study. *Research in Science Education*, *46*, 831–855. https://doi.org/10.1007/s11165-015-9483-9

McDonald, S. P., & Kelly, G. J. (2012). Beyond argumentation: Sense-making discourse in the science classroom. In M. Khine (Ed.), *Perspectives on scientific argumentation* (pp. 265–281). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-2470-9_13

McNeill, K. L., & Krajcik, J. (2007). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. In M. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 233–265). Erlbaum. https://doi.org/10.4324/9780203810057

McNeill, K. L., Katsh-Singer, R., González-Howard, M., & Loper, S. (2016). Factors impacting teachers' argumentation instruction in their science classrooms. *International Journal of Science Education*, *38*(12), 2026–2046. https://doi.org/10.1080/09500693.2016.1221547

Mikeska, J.N., Cross Francis, D. I., Lottero-Perdue, P. S., Park Rogers, M. A., Shekell, C., Bharaj, P. K., Howell, H., Maltese, A., Thompson, M. & Reich, J. (2025). Promoting preservice teachers' facilitation of argumentation in mathematics and science through digital simulations. *Teaching and Teacher Education*, 154. https://doi.org/10.1016/j.tate.2024.104858

Mikeska, J. N., & Howell, H. (2020). Simulations as practice-based spaces to support elementary teachers in learning how to facilitate argumentation-focused science discussions. *Journal of Research in Science Teaching*, *57*(9), 1356–1399. https://doi.org/10.1002/tea.21659

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/10.1177/00224871221142

Mikeska, J. N., 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. https://citejournal.org/volume-21/issue-3-21/general/understanding-the-role-of-simulations-in-k-12-mathematics-and-science-teacher-education-outcomes-from-a-teacher-education-simulation-conference/

Mikeska, J. N., Webb, J., Bondurant, L., Kwon, M., Imasiku, L., Domjan, H. N., & Howell, H. (2022). Using and adapting simulated teaching experiences to support preservice teacher learning. In P. H. Bull & G. C. Patterson (Eds.), *Redefining teacher education and teacher preparation programs in the post-COVID-19 era* (pp. 46-78). IGI Global.

Nation, M. T., & Feldman, A. (2022). Climate change and political controversy in the science classroom: How teachers' beliefs influence instruction. *Science & Education*, *31*(6), 1567–1583. https://doi.org/10.1007/s11191-022-00330-6

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. https://www.thecorestandards.org/Math/

Nespor, J. (1987). The role of beliefs in the practice of teaching. Journal of Curriculum Studies, 19, 317–328. https://doi.org/10.1080/0022027870190403

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. The National Academies Press.

Nguyen, D. T., & Tran, D. (2023). High school mathematics teachers' changes in beliefs and knowledge during lesson study. *Journal of Mathematics Teacher Education*, *26,* 809–824 https://doi.org/10.1007/s10857-022-09547-2

Norris, S., Philips, L., & Osborne, J. (2007). Scientific inquiry: the place of interpretation and argumentation. In J. Luft, R. Bell & J. Gess-Newsome (Eds.), *Science as inquiry in the secondary setting* (pp. 87-98). NSTA Press.

Noviyanti, N. I., Mukti, W. R., Yuliskurniawati, I. D., Mahanal, S., & Zubaidah, S. (2019). Students' scientific argumentation skills based on differences in academic ability. In *Journal of physics: Conference series* (Vol. 1241, No. 1, p. 012034). IOP Publishing. https://doi.org/10.1088/1742-6596/1241/1/012034

Osborne, J. F., & Patterson, A. (2011). Scientific argument and explanation: A necessary distinction? *Science Education*, *95*(4), 627–638. https://doi.org/10.1002/sce.20438

Osborne, J., Simon, S., Christodoulou, A., Howell-Richardson, C., & Richardson, K. (2013). Learning to argue: A study of four schools and their attempt to develop the use of argumentation as a common instructional practice and its impact on students. *Journal of Research in Science Teaching*, *50*(3), 315–347. https://doi.org/10.1002/tea.21073

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*(3), 307–332. https://doi.org/10.3102/00346543062003307

Peretz, D. (2006). Enhancing reasoning attitudes of prospective elementary school mathematics teachers. *Journal of Mathematics Teacher Education*, *9*, 381–400. https://doi.org/10.1007/s10857-006-9013-9

Purnama, A. N. H., Winarno, N., Prima, E. C., & Ahmad, N. J. (2023). Fostering students' concept mastery through stem-engineering design process in thermal energy and heat transfer topic. *Integrative Science Education and Teaching Activity Journal*, *4*(2), 209–230. https://doi.org/10.21154/insecta.v4i2.7240

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. https://doi.org/10.70725/023707spaywm

Rudsberg, K., Öhman, J., & Östman, L. (2013). Analyzing students' learning in classroom discussions about socioscientific issues. *Science Education*, *97*(4), 594–620. https://doi.org/10.1002/sce.21065

Rumsey, C., & Langrall, C. W. (2016). Promoting mathematical argumentation. *Teaching Children Mathematics*, *22*(7), 412–419. https://doi.org/10.5951/teacchilmathematics.22.7.0412

Sampson, V., & Blanchard, M. R. (2012). Science teachers and scientific argumentation: Trends in views and practice. *Journal of Research in Science Teaching*, *49*(9), 1122–1148. https://doi.org/10.1002/tea.21037

Sampson, V., Enderle, P., & Grooms, J. (2013). Argumentation in science education. *The Science Teacher*, *80*(5), 30–33. https://doi.org/10.2505/4/tst13_080_05_30

Schoenfeld, A. H. (1998). Toward a theory of teaching-in-context. *Issues in Education*, *4*(1), 1–94. https://doi.org/10.1016/S1080-9724(99)80076-7

Schwaighofer, M., Vogel, F., Kollar, I., Ufer, S., Strohmaier, A., Terwedow, I., Ottinger, S., Reiss, K., & Fischer, F. (2017). How to combine collaboration scripts and heuristic worked examples to foster mathematical argumentation – when working memory matters. *International Journal of Computer-Supported Collaborative Learning*, *12*(3), 281–305. https://doi.org/10.1007/s11412-017-9260-z

Shamim, M. R. H., Al Mamun, M. A., & Raihan, M. A. (2022). Mapping the research of technical teachers' pedagogical beliefs about science technology engineering and mathematics (STEM) Education. *International Journal of Instruction*, *15*(4), 797–818. https://doi.org/10.29333/iji.2022.15443a

Shaughnessy, M., Boerst, T. A., & Farmer, S. O. (2019). Complementary assessments of prospective teachers' skill with eliciting student thinking. *Journal of Mathematics Teacher Education*, *22*, 607–638. https://doi.org/10.1007/s10857-018-9402-x

Singletary, L. M., & Conner, A. (2015). Focusing on mathematical arguments. *Mathematics Teacher*, *109*(2), 143–147. https://doi.org/10.5951/mathematicsteacher.109.2.0143

Sleep, L., & Eskelson, S. L. (2012). MKT and curriculum materials are only part of the story: Insights from a lesson on fractions. *Journal of Curriculum Studies*, *44*(4), 537–558. https://doi.org/10.1080/00220272.2012.716977

Smith, M. E., Swars, S. L., Smith, S. Z., Hart, L. C., & Haardörfer, R. (2012). Effects of an additional mathematics content course on elementary teachers' mathematical beliefs and knowledge for teaching. *Action in Teacher Education, 34*(4), 336–348. https://doi.org/10.1080/01626620.2012.712745

Smith, P. S., Smith, A. A., & Banilower, E. R. (2014). Situating beliefs in the theory of planned behavior: The development of the teacher beliefs about effective science teaching questionnaire. In C. M. Czerniak, R. Evans, J. Luft & C. Pea (Eds.), *The role of science teachers' beliefs in international classrooms: From teacher actions to student learning* (pp. 81–102). Sense Publishers.

Solar, H., Ortiz, A., Deulofeu, J., & Ulloa, R. (2021). Teacher support for argumentation and the incorporation of contingencies in mathematics classrooms. *International Journal of Mathematical Education in Science and Technology*, *52*(7), 977–1005. https://doi.org/10.1080/0020739X.2020.1733686

Staples, M., & Newton, J. (2016). Teachers' contextualization of argumentation in the mathematics classroom. *Theory into Practice*, *55*(4), 294–301. https://doi.org/10.1080/00405841.2016.1208070

Straub, C., Dieker, L., Hynes, M., & Hughes, C. (2014). Using virtual rehearsal in TLE TeachLivE™ mixed reality classroom simulator to determine the effects on the performance of mathematics teachers: The 2014 TeachLive National Research Project Year 1 findings. In A. Heyes, S. Hardin, L. Dieker, M. Hynes, & C. Hughes (Eds.), *Proceedings of the 2nd National TLE TeachLivE™ Conference*. University of Central Florida.

Straub, C., Dieker, L., Hynes, M., & Hughes, C. (2015). Using virtual rehearsal in TLE TeachLivE™ mixed reality classroom simulator to determine the effects on the performance of science teachers: A follow-up study (Year 2). In *Proceedings of 3rd National TLE TeachLivE™ Conference*. University of Central Florida.

Straub, C. (2018). *Best in class leadership development: How virtual reality and avatars are changing the learning landscape*. Mursion, Inc. https://www.denasamuels.com/wp-content/uploads/2020/03/How-Virtual-Reality-and-Avatars-are-Changing-the-Learning-Landscape.pdf

Thompson, A. (1992). Teachers' beliefs and conceptions: A synthesis of research. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127-146). Macmillan.

Tizón-Escamilla, N., & Burgos, M. (2023). Creation of problems by prospective teachers to develop proportional and algebraic reasonings in a probabilistic context. *Education Sciences*, *13*(12), 1186. https://doi.org/10.3390/educsci13121186

Tsai, C. Y., Lin, C. N., Shih, W. L., & Wu, P. L. (2015). The effect of online argumentation upon students' pseudoscientific beliefs. *Computers & Education*, *80*, 187–197. https://doi.org/10.1016/j.compedu.2014.08.018

Tschannen-Moran, M., & Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783–805. https://dx.doi.org/10.1016/S0742-051X(01)00036-1

Uluay, G. (2021). Digital simulation experiences of pre-service science teachers: An example of circuits. *International Journal of Progressive Education, 17*(3), 14–30. https://doi.org/10.29329/ijpe.2021.346.2

Wall, C. R. G. (2016). From student to teacher: Changes in preservice teacher educational beliefs throughout the learning-to-teach journey. *Teacher Development*, *20*(3), 364-379.

Wertheim, C., & Leyser, Y. (2002). Efficacy beliefs, background variables, and differentiated instruction of Israeli prospective teachers. *The Journal of Educational Research*, *96*(1), 54–63. https://doi.org/10.1080/00220670209598791

Widodo, A. P. A., Hufad, A., Sunardi, S., & Nandiyanto, A. B. D. (2020). Collaborative teaching in heat transfer for slow learner students. *Journal of Engineering Science and Technology*, *1*, 11–21.

Wilkins, J. L. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education*, *11*(2), 139-164.

Windschitl, M., Thompson, J., & Braaten, M. (2020). *Ambitious science teaching*. Harvard Education Press.

Woolfolk, A. E., Winne, P. H., Perry, N. E., & Shapka, J. (2009). *Educational psychology* (4th Canadian ed.).Pearson Education.

Zourmpakis, A. I., Kalogiannakis, M., & Papadakis, S. (2023). A review of the literature for designing and developing a framework for adaptive gamification in physics education. In M. F. Taşar & P. R. L. Heron (Eds.), *The international handbook of physics education research: Teaching physics*. https://dx.doi.org/10.1063/9780735425712_005

Appendix A
Instruments and Data Type Used to Respond to Research Questions

Research Question	Presemester Instruments		Midsemester Instruments				Postsemester Instruments		
	PST Presurvey	PST Preenacted Measure	PST TM Task Survey	TE TM Task Survey	PST ABS Task Survey	TE ABS Task Survey	PST Post Survey	PST Postenacted Measure	TE Postinterview
RQ1 PST Learning and Perception of Value	-	-	Open Response Likert Yes/No	Open Response Likert	Open Response Likert Yes/No	Open Response Likert	-	-	-
RQ2 PST Understanding of Argument-Focused Discussions	Open Response	-	-	-	-	-	Open Response Likert Yes/No	-	-
RQ3 PST Change in Beliefs: Content Instruction and Teaching Preparedness	Open Response Yes/No	Likert	Open Response Yes/No	-	Open Response Yes/No	-	Open Response Likert Yes/No	-	-
RQ4 PST Improvement Facilitating Argument-Focused Discussions	-	MBI/TBEST Items	-	-	Likert	Likert	Likert Open Response	MBI/TBEST Items	Open Response

Note. PST = preservice teacher; TE = teacher educator; TM = Teacher Moments; ABS = Avatar-Based Simulation; RQ = Research Question; MBI = Mathematics Beliefs Instrument; TBEST = Teacher Beliefs about Effective Science Teaching

Appendix B

Survey Items Analyzed by Research Question

Question (abbreviated)	PST/TE	Survey(s)	Response Type	Codebook	Total Responses	No. of Responses Group Coded	No. of Responses Double Coded	Agreement for Double Coded Responses [a]
RQ1 (Perceptions)								
Describe your main takeaway/learning from the simulation cycle.	PST	TM, ABS	Open	Adapted	77[b]	11 (14%)	16 (21%)	92% (0.734)
What did your PSTs learn from participating in the simulation cycle?	TE	TM, ABS	Open	Adapted	12[d]	0 (0%)	12 (100%)	87% (0.696)
How appropriate do you believe a TM/ABS activity would be as a component of a teacher preparation methods course?	PST, TE	TM, ABS	Likert	n/a	77[b], 12[d]	--	--	--
Would you recommend including TM/ABS discussions in future courses?	PST	TM, ABS	Likert	n/a	77[b]	--	--	--
In your view, in what ways did the simulation experience support you in your role as a math/science teacher?	PST	TM, ABS	Open	New	77[b]	11 (14%)	16 (21%)	93% (0.707)
RQ2 (Understanding)								
In your view, what are the characteristics of high-quality discussions focused on argumentation in a math/science classroom?	PST	Pre, Post	Open	Adapted	77[c]	8 (11%)	16 (21%)	94% (0.864)
Has your view of high-quality discussion focused on argumentation in a math/science classroom changed over the course of the semester?	PST	Post	Likert	n/a	38	--	--	--
Describe how your view of high-quality discussions focused on argumentation has changed over the course of the semester.	PST	Post	Open	New	20[e]	0 (0%)	20 (100%)	97% (0.949)
What activities or experiences, inside or outside of class, do you feel contributed to that change, and how?	PST	Post	Open	New	20	0 (0%)	20 (100%)	98% (0.981)
RQ3 (Beliefs)								
How important do you feel it is for secondary students to have opportunities to engage in argumentation in math/science?	PST	Pre, TM, ABS, Post	Likert	n/a	154[f]	--	--	--
Explain why you think it is or is not important for secondary students to have opportunities to engage in argumentation in science/math.	PST	Pre, TM, ABS, Post	Open	Adapted	154[f]	24 (16%)	26 (17%)	92% (0.780)
Have your perspectives about the importance of argumentation for secondary students changed over the course of the semester?	PST	Post	Likert	n/a	38	--	--	--
Please describe in your own words how your perspectives of the importance of argumentation for secondary students has changed.	PST	Post	Open	New	18[g]	0 (0%)	18 (100%)	94% (0.899)
What experiences inside or outside of class contributed to that change?	PST	Post	Open	New	18[g]	0 (0%)	18 (100%)	87% (0.840)
RQ4 (Skill)								
How did your level of preparedness to implement each element of AFD change over the course of the semester?	PST	Post	Likert	n/a	190[h]	--	--	--
Explain what supports helped you to be more prepared with respect to the following elements of argumentation-focused discussion.	PST	Post	Open	New	167	17 (10%)	34 (20%)	90% (0.879)

Question (abbreviated)	PST/TE	Survey(s)	Response Type	Codebook	Total Responses	No. of Responses Group Coded	No. of Responses Double Coded	Agreement for Double Coded Responses [a]
To what extent have your PSTs shown progress in developing their skills in elements of AFD from the beginning to the end of the semester?	TE	ABS	Likert	n/a	30[i]	--	--	--

[a] Percent Agreement (One-way random Average Measures ICC)
[b] This question was asked on both the TM survey and the ABS survey. All 39 PSTs responded to the ABS survey; however, one PST did not complete the TM survey, resulting in 38 responses for the TM survey, and a total of 77 responses overall.
[c] This question was asked on both the presurvey and postsurvey. All 39 PSTs responded to the presurvey. However, one PST did not complete the postsurvey, resulting in 38 responses for the postsurvey and a total of 77 responses overall.
[d] This question was asked on both the presurvey and postsurvey. All six TEs responded to both the presurvey and postsurvey, resulting in a total of 12 responses overall.
[e] Only 20 of the 38 PSTs who completed the postsurvey responded "yes" to the question "Has your view of the high-quality discussion focused on argumentation in a math/science classroom changed over the course of the semester?" on the postsurvey. Therefore, only 20 PSTs were given the follow-up prompt.
[f] This question was asked on the presurvey, TM survey, ABS survey, and postsurvey. All 39 PSTs responded to the presurvey and ABS survey; however, one PST did not complete the TM survey and postsurvey, resulting in 38 responses for the TM survey and postsurvey and a total of 154 responses.
[g] Only 18 of the 38 PSTs who completed the postsurvey responded "yes" to the question "Have your perspectives about the importance of argumentation for secondary students changed over the course of the semester?" on the postsurvey; therefore, only 18 PSTs were given the follow up prompt to describe how their perspectives changed.
[h] This question was asked on the postsurvey. However, only 38 PSTs completed the postsurvey. There were a total of 190 data points for this question, as the 38 PSTs responded across each of the five elements of high-quality discussion focused on argumentation.
[i] This question was asked on the postsurvey. which was completed by all six TEs. There were a total of 30 data points for this question, as the 6 TEs responded across each of the five elements of high-quality discussion focused on argumentation.
Note. PST = preservice teacher; TE = teacher educator; TM = Teacher Moments; ABS = Avatar-Based Simulation; RQ = Research Question