Proceedings of the 125th Annual Convention of the School Science and Mathematics Association



The Power of Connection: 125 Years of Bridging Math and Science

Chicago, Illinois
November 5–7, 2026

PROCEEDINGS OF THE 125th ANNUAL CONVENTION OF THE SCHOOL SCIENCE AND MATHEMATICS ASSOCIATION

November 5–7, 2026
Chicago, IL

Editors: Katie Anderson-Pence and Ryann N. Shelton

Cite Proceedings Volume as:
Anderson-Pence, K., & Shelton, R. N. (Eds.). (2026). *Proceedings of the 125th annual convention of the School Science and Mathematics Association* (Vol. 13). Chicago, IL: SSMA.

Cite Proceedings Chapter as:
Author. (2026). Title of chapter. In K. Anderson-Pence & R. N. Shelton (Eds.), *Proceedings of the 125th annual convention of the School Science and Mathematics Association* (Vol. 13). Chicago, IL: SSMA.

Website: www.ssma.org
The proceedings are also available via http://www.ssma.org/convention-proceedings

Copyrights © 2026 left to the authors

Articles published in the Proceedings are copyrighted by the author. Permission to reproduce an article or portions from an article must be obtained from the author.

All rights reserved

ISBN: 978-1-7343159-7-4

SSMA 2026 ANNUAL CONVENTION: CHICAGO, IL

2026 Convention Coordination

Convention Director
Christa Jackson, Saint Louis University

Convention Chair
Rebekah Hammack, Purdue University

SSMA Leadership

President, 2024–2026
Sandi Cooper
Baylor University

President-Elect, 2025–2026
Megan Che
Clemson University

Directors-at-Large, 2023–2026
Amy Corp
East Texas A&M University

Casey Wells
Appalachian State University

Directors-at-Large, 2024–2027
Katie Anderson-Pence
University of Wisconsin-Stevens Point

Bridget Miller
University of South Carolina

Directors-at-Large, 2025–2028
Kristin Cook
Bellarmine University

Merryn Cole
University of Nevada, Las Vegas

Co-Executive Directors, 2024–2028
Christie Martin
University of South Carolina

Bridget Miller
University of South Carolina

Social Media Coordinator
Ashley Schmidt
University of Wisconsin-Milwaukee

Journal Co-Editors, 2021–2026
Christie Martin
University of South Carolina

Bridget Miller
University of South Carolina

Proceedings Editors
Katie Anderson-Pence
University of Wisconsin-Stevens Point

Ryann N. Shelton
Baylor University

Special thanks to the Publications Committee and SSMA members for reviewing and providing valuable feedback on the submitted manuscripts.

Anderson-Pence, K., & Shelton, R. N. (Eds.). (2026). *Proceedings of the 125th annual convention of the School Science and Mathematics Association* (Vol. 13). Chicago, IL: SSMA.

PRESIDENT'S FORWARD

125th Anniversary of the School Science and Mathematics Association

This year marks an extraordinary milestone in our organization's history—125 years of advancing excellence in science and mathematics education through scholarship, collaboration, and service. Since its founding in 1901, SSMA has brought together educators, researchers, and leaders who share a commitment to strengthening teaching and learning and inspiring future generations. As we celebrate this remarkable anniversary, we honor the vision of those who established our association, the dedication of those who have sustained it, and the collective contributions that continue to shape its future.

An anniversary of this significance invites us to reflect on both how far we have come and the possibilities that lie ahead. Over the past 125 years, science and mathematics education has experienced profound transformations in curriculum, instructional practices, educational research, and technology. Through these changes, SSMA has remained steadfast in its mission to connect research and practice, foster meaningful professional relationships, and promote excellence in science and mathematics education. Our association's enduring strength lies not simply in its longevity, but in the people whose curiosity, dedication, and shared commitment have sustained a vibrant professional community across generations.

As we look toward the future, the importance of our work has never been more evident. Today's educators and researchers face new challenges and opportunities as they prepare learners to think critically, reason mathematically, engage in scientific inquiry, and navigate an increasingly complex and interconnected world. The continued advancement of science and mathematics education requires us to embrace innovation while remaining grounded in sound research, effective teaching, and a commitment to providing meaningful learning opportunities for all students. SSMA continues to serve as an important space for these conversations, encouraging collaboration across disciplines and connecting ideas that have the potential to transform educational practice.

The presentations, research papers, and professional conversations represented in these proceedings reflect the vitality of our association and the dedication of its members. They demonstrate the power of bringing together diverse perspectives, innovative approaches, and a shared passion for improving science and mathematics education. Each contribution adds to the rich legacy of scholarship and professional exchange that has defined SSMA for 125 years.

As we celebrate this historic milestone, I am reminded that the true legacy of SSMA is found not only in its accomplishments, but also in the relationships it has fostered, the educators it has supported, and the students whose learning has been enriched through the work of its members. We are both beneficiaries and stewards of this legacy. It is now our opportunity and responsibility to build upon the foundation established by those who came before us and ensure that SSMA remains a thriving, relevant, and welcoming professional community for generations to come.

Thank you to our members, presenters, researchers, educators, and organizational leaders—past and present—whose contributions have made this anniversary possible. Your commitment, collegiality, and passion continue to give SSMA its purpose and strength.

As we commemorate **125 years of connecting research, practice, and people**, may we celebrate our shared history, find inspiration in one another's work, and look forward with renewed enthusiasm to the possibilities of the next 125 years.

With appreciation,

Sandi Cooper

Sandi Cooper
President, School Science and Mathematics Association (2024–2026)



PREFACE

These proceedings are a written record of some of the research and instructional innovations presented at the 125th Annual Convention of the School Science and Mathematics Association held November 5–7, Chicago, IL. The blinded, peer-reviewed proceedings include twenty-five papers regarding instructional innovations and research. The acceptance rate for the proceedings was 92.5%. We are pleased to present these Proceedings as an important resource for the mathematics, science, and STEM education community.

Katie Anderson-Pence & Ryann N. Shelton
Editors

Anderson-Pence, K., & Shelton, R. N. (Eds.). (2026). *Proceedings of the 125th annual convention of the School Science and Mathematics Association* (Vol. 13). Chicago, IL: SSMA.

TABLE OF CONTENTS
VOLUME 13

SECTION I: SCIENCE

CURRICULUM FIDELITY AND A REFORM-ORIENTED SCIENCE TEACHING IDENTITY: AN INTERSECTION OF PEDAGOGY AND SELF-UNDERSTANDING **9**
Richard H. Kozoll & Peter S. Ower

INTEGRATED STEM DESIGN: MAINTAINING CONCEPTIONAL FOCUS IN SCIENCE AND CODING **17**
Webster Wong, Anna Bloodworth, Aida Alibek, Barbara A. Crawford, & AnnaMarie Conner

TASK PERCEPTION AS A DIMENSION OF STEM TEACHER IDENTITY: INSIGHTS FROM INSTRUCTIONAL ARTIFACTS IN A PHYSICS-TO-CLASSROOM CAREER EXPLORATION PATHWAY **26**
Terrie M. Galanti, W. Brian Lane, & J. Caleb Speirs

SECTION II: MATHEMATICS

A FRAMEWORK TO ASSESS PROSPECTIVE TEACHERS' ASSESSMENT OF STUDENT THINKING **36**
Nama Namakshi & Hiroko Warshauer

CATEGORIZING NON-ROUTINE ALGEBRA DIFFICULTIES TO SUPPORT STEM LEARNERS **45**
Eunmi Joung & Kuo-Liang Chang

ETHNOMODELING: A VEHICLE FOR CATALYZING CHANGE IN MATHEMATICS TEACHER PREPARATION **55**
Siddhi Soni, Alyssa Harbin, & Trena Wilkerson

LONGITUDINAL DEVELOPMENT OF ELEMENTARY STUDENTS' WRITTEN MATHEMATICAL ARGUMENTS **64**
Madelyn Colonnese, Fabiana Cardetti, & Christie Martin

ELICITING MATHEMATICAL THINKING: TEACHERS' QUESTIONING DURING READ-ALOUDS **74**
Mahtob Aqazade, Laura Bofferding, Beatriz Ruiz Ordoñez, Lisa Hawley, & Jelena Byers

THE USE OF AI IN DESIGNING MATHEMATICS INTERVENTIONS FOR STUDENTS WITH DISABILITIES **82**
Mahtob Aqazade, Sebnem Tekin, & Matt Mauntel

SENSING INTERVENTION OPPORTUNITIES IN SMALL-GROUP MATHEMATICAL ARGUMENTATION **91**
Jennifer Kleiman, Ngutor Tembe, & AnnaMarie Conner

PRESERVICE TEACHERS' MATHEMATICAL KNOWLEDGE FOR TEACHING: CONTEXTUAL FRACTION MULTIPLICATION **100**
Tegan W. Nusser, Brooke Krejci, Bima Sapkota, & Bona Kang

EXPLORING FACTORS THAT SUPPORT STUDENTS' LEARNING IN A SOCIAL JUSTICE MATHEMATICS CLASSROOM **110**
Waleed A. Raja, Nicole L. Fonger, & Anis Munfarikhatin

REPRESENTATIONS OF MATHEMATICAL PROOF ACROSS MAJOR U.S. STANDARDS DOCUMENTS **119**
Seyedehkhadijeh Azimi Asmaroud & Somya N. Thakkar

CHILDREN'S MAP DRAWING: ANALYZING VISUOSPATIAL REASONING ELEMENTS IN A SUMMER LITERACY PROGRAM **127**
Maria Eloisa Nuguid, Comfort Aje, Abigail Erskine, Guili Zhang, Breanya Hogue, & Signe Kastberg

CENTERING MATHEMATICS: A MATH-FOCUSED INTEGRATED STEM APPROACH THROUGH DIGITAL FLOWERS GENETICS **137**
Hanan Alyami & Rebecca Palu

INTRODUCING AN ONLINE FLIPPED MODEL (OFM) IN UNIVERSITY PRECALCULUS COURSES: A PRACTITIONER'S GUIDE **146**
Nelson L. Carter & Sue Brown

OPERATIONALIZATION OF PARTNERSHIP IN MATHEMATICS EDUCATION RESEARCH: A SCOPING REVIEW **154**
S. Megan Che, Trena Wilkerson, Colleen Eddy, Karen Enderle, Amy James, & Lianne Jones

SECTION III: INTEGRATED STEM

LEARNING SCIENCE IN ACTION: USING EXAMPLES AND NON-EXAMPLES TO DEEPEN MATH AND SCIENCE UNDERSTANDING **165**
Andrea Foster & Kristi Martin

INCREDIBLE DECISIONS: INVESTIGATING A TEACHER'S CHANGES TO THEIR LESSON PLAN IN ENACTMENT **174**
Alissa H. Chung, Jenna Gist, & Rebekah Hammack

ENHANCING THE VALIDITY OF PROBABILITY PROMPTS USING FOUR GENERATIVE ARTIFICIAL INTELLIGENCE TOOLS **183**
Ayse Fethiye Gemci, Amanda Gantt Sawyer, Haley Smiley, Richard Busi, & Juliana Galano

FOSTERING CONNECTION: USING AI TO DEVELOP TEACHER CANDIDATES' DISCERNMENT AND REFLECTIVE PRACTICE **192**
Thomas Roberts, Holly M. Plank, Amber Lawson, Kate Brodeur, Karen Wilhelm, & Stefanie D. Livers

STEM CAREER CHANGERS AND CRITICAL REFLECTIONS: IDENTITY, PRACTICE, AND EQUITY **201**
Jennifer R. Simons

INVESTIGATING FACTORS IMPACTING PRESERVICE ELEMENTARY TEACHERS' iSTEM IDENTITY DEVELOPMENT **209**
Azadeh Hassani, Ursula Nguyen, Fatemeh Ashrafabadi, Deepika Menon, & Amanda Henson

NOT TOO SMALL FOR EMERGING SCIENCE: NAVIGATING SCALE, COMPLEXITY, AND QUANTUM CONCEPTS IN ELEMENTARY CLASSROOMS **218**
Nancy Holincheck, Jennifer Simons, Tiffany Butler, Jessica Rosenberg, Benjamin W. Dreyfus, Stephanie Dodman, Christina Cameron, & Laura Akesson

NAVIGATING DISCIPLINARY BOUNDARIES: A QUALITATIVE CASE STUDY OF CROSS-DISCIPLINARY UNDERGRADUATE RESEARCH IN TEACHER EDUCATION **227**
Lacey Fitts

SECTION I: SCIENCE

CURRICULUM FIDELITY AND A REFORM-ORIENTED SCIENCE TEACHING IDENTITY: AN INTERSECTION OF PEDAGOGY AND SELF-UNDERSTANDING

Richard H. Kozoll
DePaul University
rkozoll@depaul.edu

Peter S. Ower
Wilmette Public Schools
peteower@gmail.com

Abstract

The adoption of reformed science curricula has become a common strategy for schools to support science teaching best practices. Research on teachers' use of this curriculum has documented their rejections and adaptations, but we lack accounts describing teachers' experiences and reflections concerning curricular fidelity. Here, we offer a phenomenological case study of a sixth-grade science teacher's implementation of such a curriculum as intended. Her accounts constructed a science teaching identity committed to reform, rooted in intersecting pedagogical values and professional self-understandings. Thus, we recommend professional development on the use of curricula to support a teacher's reform-oriented science teaching identity.

Keywords: science, curriculum, reform, identity, phenomenology

Introduction

Currently, reformed science curricula depend on storylines to contextualize inquiry learning around phenomena. Within these stories, students use science and engineering practices to explore a main question or solve a fundamental problem. These phenomena-based curricula reflect the Next Generation Science Standards' goals (NGSS Lead States, 2013). They aim to further students' science interests and prepare them for college, careers, and social responsibility. Whether commercially produced or district-developed, schools rely upon these science curricula to realize science education reform and support best practices. Therefore, studies have considered how teachers respond to science education reform and employ its associated curricula. This research has documented teachers' rejection and adaptations of such curricula, but studies detailing their experiences with curriculum fidelity are needed (Ryder, 2015).

We can understand a teacher's curricular decisions by focusing on their voice where lived experience brings their science teaching identity to the foreground (Ritchie, 2009). While context-bound, prior research on an individual teacher's science teaching identity has offered deep insight

Anderson-Pence, K., & Shelton, R. N. (Eds.). (2026). *Proceedings of the 125th annual convention of the School Science and Mathematics Association* (Vol. 13). Chicago, IL: SSMA.

into their instructional practices. By examining the construction of a teacher's science teaching identity following the implementation of a phenomena-based science curriculum, this study further connects this type of research to broader reform initiatives (Avraamidou, 2014). These findings can inform professional development dedicated to teachers' use of reform-based curricula.

Objectives of the Study

Our broader research investigated how a suburban middle school's teachers experienced transitioning to a phenomena-based science program. Here, we offer a phenomenological case study of Andrea's (a pseudonym) faithful use of the new sixth-grade science curriculum. To appreciate her curricular adherence, we focused on Andrea's lived experience, recognizing her distinct insights and voice as a science teacher (Brown & Weinburgh, 2024). Andrea's accounts offer experiences of curricular fidelity missing from research on how teachers respond to science curriculum reform (Ryder, 2015). Moreover, her reflections address the need to situate science teaching identity research within current reform initiatives (Avraamidou, 2014). Andrea's depictions suggest that professional learning for reformed curricula could better support a reform-oriented science teaching identity.

Theoretical Framework

Current science education reform supports learning science through approaches that engage students in practices that progressively advance their understanding (NGSS Lead States, 2013). Reform-based curricula are one way to support teachers' application of those teaching strategies. Research has found that teachers' responses to reform efforts and their curricula can reflect their pedagogical values (Irez & Han, 2011). These pedagogical beliefs, as they inform classroom practice, establish curriculum implementation as a manifestation of science teaching identity (Zhai et al., 2024). However, the ideals that shape teachers' practices are further linked to who they are in the classroom (Taylor, 1991). Thus, the pedagogical aspect of a science teaching identity shapes how an individual understands their professional role. This overlap influences their reliance on reform-based curricula and their practices (Avraamidou, 2016). Andrea's descriptions of curricular fidelity revealed certain pedagogical values. These values crossed with her professional self-understandings, holistically framing the construction of her reform-oriented science teaching identity (Beijaard et al., 2004).

Methodology

We drew Andrea's phenomenological case study from research investigating teachers' experiences transitioning to a phenomena-based science curriculum at a suburban middle school. Such research examines student and teacher perceptions to understand the teaching and learning of science (Østergaard et al., 2008). Accordingly, we investigated teachers' depictions of their teaching activities to uncover insights influencing their curriculum use. A phenomenological approach offered us the teachers' unique lived experience of the curriculum.

Data collection included five semi-structured, audio-recorded interviews with participating teachers. The interviews took place the summer after the middle school launched the curriculum and lasted approximately two hours. Our interview questions encouraged participants to share their lived experience of curriculum implementation. These accounts situated their broader explanations and generalizations (van Manen, 2023).

Ongoing analysis of the interview recordings and transcripts uncovered curricular fidelity as the primary theme in Andrea's conversations. Post field analysis moved from a holistic to a detailed examination of interview data. This progression provided the interpretive depth necessary for rigorous phenomenological research (van Manen, 2023). We began with an open examination of transcribed interviews. In Andrea's case, we identified narratives that exemplified her curricular adherence and verified that theme (Creswell & Poth, 2024). We further explored those descriptions and categorized them into groups that included "phenomena," "evidence," and "ambiguity" (Tavakol & Sanders, 2025). Finally, coding that dialogue (Brinkman & Kvale, 2015) uncovered a consistency. Across the categories, Andrea framed her curricular adherence in terms of both her pedagogy and her self-understandings as a science teacher. Through our coding, we unified her discussion under pedagogy and professional self-understanding. Thus, Andrea's science teaching identity emerged through her accounts of curricular fidelity (Volkmann & Zgagacz, 2004).

Results and Discussion

Phenomena

Reflecting on her experiences using the sixth-grade science curriculum, Andrea noted its reliance on various anchoring phenomena. When discussing the events that drove the curriculum, Andrea specifically recalled the life science unit. In that storyline her class acted as medical students investigating a fictional patient's fatigue:

[The patient] described not having a lot of energy and not being able to participate in sports activities. [My students] brainstormed some things that could possibly be going on. One of their first ideas is that she's not getting enough sleep or she's not eating healthy. And they did have an opportunity to look at a few pieces of evidence right away that ruled out some of those common ones that they initially brought up. They present some evidence that shows the amount of sleep that she's getting compared to typical peers. And the kinds of foods that she's eating compared to her peers. So, they ruled out those two possibilities right away. The last possibility left is that there's something happening with one of her body systems that is making it not function correctly.

Pedagogically, Andrea concentrated on how the curriculum's assorted phenomena related to everyday circumstances. "That phenomena basis, framing it in a real-world context, I think is aligned with what I think science education should be." Andrea also noted that the units' contextualized phenomena were "super engaging" for her students and made it "a little bit easier for them to connect the different lessons." Andrea's pedagogical alignment with the curriculum's phenomena also revealed her professional self-understanding. "I'm there to challenge them and [the phenomenon] has them think things through it a little bit more. Defend why they did things a certain way and explain their answers rather than going through it in a formulaic way." Thus, Andrea's pedagogical beliefs coincided with this professional self-understanding. Together, they developed a science teaching identity that became evident through her commitment to the curriculum.

Evidence

Andrea also recognized that the curriculum emphasized the evaluation of evidence. Recalling the life science unit, she noted that her students used a simulation to collect diagnostic evidence from the same fatigued patient:

So, it started off showing how a typical human body functions. But then [my students] learned about four different possible conditions that could be causing their patient to feel tired. And in groups they read a little bit about each of those different conditions and explained it to the other people in their group. Then they predicted what would look different in that body system model if one of those conditions was in place. There are different settings that you can put in the simulation. So, then they actually modeled it in the simulation and saw if it contradicted their prediction. Then they were presented with additional evidence about their patient from tests that were previously run

on each of those body systems. So, they used all of that evidence to try to decide which one of those four diagnoses makes the most sense.

More broadly, Andrea claimed that the curriculum's emphasis on students' use of evidence better suited her teaching style. "I think the value in learning science is not necessarily just being able to recall specific facts but learning how to think about things and identify a question and evaluate evidence to answer that question." Engaging her students in this scientific process reflected how Andrea "wanted to teach anyway." Andrea's pedagogy supported her students' use and evaluation of evidence. Her discussion of this process exposed an additional understanding of her classroom role. "I want to give them a more complete picture of what the process of science really is and how it works … You do research and you collect evidence to try to solve the problem." Through continued pedagogical discussion, Andrea again recognized her professional self. These coinciding values and professional self-understandings further articulated the science teaching identity around her curricular fidelity.

Ambiguity

Alongside engaging students with the curriculum's anchoring phenomena and evidence collection, Andrea also welcomed the units' ambiguous resolutions. While addressing this aspect of the curriculum, Andrea discussed a transition to the life science unit's storyline. This shift directed students' attention to a competitive cyclist whose dramatic improvement through two races suggested blood doping:

In that same unit they learned about an athlete and the question was if he succeeded in a bike race by blood doping? … So, there's a science seminar activity and they had a chance to argue. And with evidence, groups came to different conclusions based on either how they interpreted it or which pieces of evidence they thought were higher quality or which pieces they thought were the most important. They talked about whether or not the things that they were using were relevant, the way they prioritized different pieces of evidence, and the way they choose to interpret different pieces of evidence. At the and they asked what's the answer. I told them, "The answer that you decided is the answer. You used the evidence, you came to a conclusion, and you explained why that was your conclusion. That is science.

Just as she did in her discussions of phenomena and evidence, Andrea continued to align her pedagogy with the curriculum's practices. "I like that it gives them a chance to experience different conclusions based on different evidence, based on their own background or beliefs or values. The

curriculum is introducing the fact that there's often a lot of ambiguity." With this final curricular example, Andrea expanded on her instructional role in the classroom. "I should let them see it depends on what data you focus on, or what they're interpreting it through. The result is not necessarily that anyone who does the study is going to come to the same conclusion." Andrea's account of this uncertain curricular outcome demonstrated a final integration of her pedagogy and professional self-understanding. Consequently, this recollection sustained a science teaching identity embodying both this convergence and Andrea's commitment to reform.

Implications

Andrea's lived experience of the curriculum highlighted certain instructional elements. These included the anchoring phenomena, the emphasis on evidence, and ambiguous outcomes. Her focus on these three instructional components revealed aspects of her pedagogy. Pedagogically, Andrea valued situating conceptual connections in everyday life while engaging her students in scientific practices. Andrea's reflections on these curricular components further illustrated her self-understandings as a science teacher. She believed she should challenge students to engage with evidence and scientific interpretation. Through her lived experience of curricular fidelity, Andrea constructed a distinct reform-oriented science teaching identity, which intersected her pedagogical values with her professional self-understandings (Klechtermans, 2009). While such science teaching identities can conflict with instructional reform (Guzey & Ring-Whalen, 2018), Andrea's emerged in alignment with the phenomenon-based curriculum and its intended teaching practices.

Before implementing the newly adopted phenomena-based curricula, Andrea and her fellow middle school science teachers completed professional development. According to Andrea, the in-services mirrored conclusions on how to effectively support curriculum use (Jones & Eick, 2007). This professional development familiarized Andrea with the curriculum's overall approach and characteristics. She and the other middle school science teachers also participated in certain program activities to prepare for classroom implementation.

However, in her discussion of that professional development, Andrea did not include expressing her pedagogical ideals or their relationship to her role as a science teacher. Yet, these values and positional understandings were central to Andrea's science teaching identity and her implementation of the curriculum as intended. Such discussions move these in-services beyond simple familiarity with the curriculum. They provide professional development for reform-based curriculum with a reflective element – a praxis that nurtures professional growth when teachers

implement a new curriculum (Lesnefsky et al., 2026). Perhaps such professional development would have validated Andrea's lived experience while supporting the ongoing development of her reform-oriented science teaching identity (Chen & Mensah, 2022).

References

Avraamidou, L. (2014). Studying science teacher identity: Current insights and future research directions. *Studies in Science Education*, *50*(2), 145–179. https://doi.org/10.1080/03057267.2014.937171

Avraamidou, L. (Ed.). (2016). *Studying science teacher identity: Theoretical, methodological, and empirical explorations*. Sense Publishers. https://doi.org/10.1007/978-94-6300-528-9

Beijaard, D., Meijer, P. C., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education, 20*(2), 107–128. https://doi.org/10.1016/j.tate.2003.07.001

Brinkmann, S., & Kvale, S. (2018). *Doing interviews* (2nd ed.). SAGE. https://doi.org/10.4135/9781529716665

Brown, K. A., & Weinburgh, M. (2024). Don't tell me what I need: Exploring science teacher professional development with narrative research. In B. Cory & A. Ray (Eds.), *Proceedings of the 123rd annual convention of the School Science and Mathematics Association* (Vol. 11). Knoxville, TN: SSMA.

Chen, J. L., & Mensah, F. M. (2022). Toward socially just science teaching through professional development: The science teacher identity development and agency of two elementary teachers of color. *Science Education, 106*(2), 385–411. https://doi.org/10.1002/sce.21699

Creswell, J. W., & Poth, C. N. (2024). Qualitative inquiry & research design: Choosing among five approaches (5th ed.). SAGE.

Guzey, S. S., & Ring-Whalen, E. A. (2018). Negotiating science and engineering: An exploratory case study of a reform-minded science teacher. *International Journal of Science Education*, *40*(7), 723–741. https://doi.org/10.1080/09500693.2018.1445310

Irez, S., & Han, C. (2011). Educational reforms as paradigm shifts: Utilizing Kuhnian lenses for a better understanding of the meaning of, and resistance to, educational change. *International Journal of Environmental and Science Education, 6*(3), 251–266.

Jones, M. T., & Eick, C. J. (2007). Implementing inquiry kit curriculum: Obstacles, adaptations, and practical knowledge development in two middle school science teachers. *Science Education, 91*(3), 492–513. https://doi.org/10.1002/sce.20197

Kelchtermans, G. (2009). Who I am in how I teach is the message: Self-understanding, vulnerability and reflection. *Teachers and Teaching*, *15*(2), 257–272. https://doi.org/10.1080/13540600902875332

Lesnefsky, R. R., Sadler, T. D., & Fortus, D. (2026). Beyond implementation: How teachers reflect, adapt, and grow with an innovative science curriculum. *Journal of Science Teacher Education*, *37*(3), 343–367. https://doi.org/10.1080/1046560X.2025.2551405

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

Østergaard, E., Dahlin, B., & Hugo, A. (2008). Doing phenomenology in science education: A research review. *Studies in Science Education*, *44*(2), 93–121. https://doi.org/10.1080/03057260802264081

Ritchie, S. M. (2009). Looking for Daisy: Constructing teacher identities. *Cultural Studies of Science Education, 4*(3), 595–599. https://doi.org/10.1007/s11422-008-9172-3

Ryder, J. (2015). Being professional: Accountability and authority in teachers' responses to science curriculum reform. *Studies in Science Education*, *51*(1), 87–120. https://doi.org/10.1080/03057267.2014.1001629

Tavakol, M., & Sandars, J. (2025). Twelve tips for using phenomenology as a qualitative research approach in health professions education. *Medical Teacher*, *47*(9), 1441–1446. https://doi.org/10.1080/0142159X.2025.2478871

Taylor, C. (1991). The dialogical self. In D. R. Hiley, J. F. Bohman, & R. Shusterman (Eds.), *The interpretive turn: Philosophy, science, culture* (pp. 304–314). Cornell University Press. https://doi.org/10.7591/9781501735028-017

van Manen, M. (2016). *Researching lived experience: Human science for an action sensitive pedagogy* (2nd ed.). Routledge. https://doi.org/10.4324/9781315421056

van Manen, M. (2023). *Phenomenology of practice: Meaning-giving methods in phenomenological research and writing* (2nd ed.). Routledge. https://doi.org/10.4324/9781003228073

Volkmann, M. J., & Zgagacz, M. (2004). Learning to teach physics through inquiry: The lived experience of a graduate teaching assistant. *Journal of Research in Science Teaching, 41*(6), 584–602. https://doi.org/10.1002/tea.20017

Zhai, Y., Tripp, J., & Liu, X. (2024). Science teacher identity research: A scoping literature review. *International Journal of STEM Education, 11*(20), Article 20. https://doi.org/10.1186/s40594-024-00481-8

INTEGRATED STEM DESIGN: MAINTAINING CONCEPTUAL FOCUS IN SCIENCE AND CODING

Webster Wong
University of Georgia
whw99124@uga.edu

Anna Bloodworth
University of Georgia
anna.bloodworth@uga.edu

Aida Alibek
University of Georgia
aida.alibek@uga.edu

Barbara A. Crawford
University of Georgia
barbarac@uga.edu

AnnaMarie Conner
University of Georgia
aconner@uga.edu

Abstract

We present a fourth-grade science unit integrating robotics to consider aspects of high-quality integrated curriculum design. An interdisciplinary team of teachers, teacher leaders, and researchers developed the unit, emphasizing argumentation, integration, and level of cognitive demand. This unit provides an example of meaningful integration across disciplines that maintains conceptual focus.

Keywords: STEM integration, Elementary, Science, Coding, Curriculum

Introduction

More than 20 years ago, Venville et al. (2002) reflected on curricular integration in the context of science education, claiming that "[l]earning gains in one area… are often offset by losses in another" (p. 61). This is a valid concern: integration is not worthwhile if teachers cannot achieve what they need to in terms of disciplinary content. As educators embrace STEM integration (NRC, 2013), it is important to consider whether the quality of the instructional design is maintained across the different disciplines. One way of integrating STEM is to introduce educational robots within lessons. However, when robotics is used with science or mathematics, it can be difficult to maintain a conceptual focus to support students' learning of the grade-level standards (Estapa & Tank, 2017).

Objective of Study

We investigated a unit of elementary science instruction that integrated coding and robotics with attention to high-level thinking and focus on disciplinary concepts. We use this unit as a case to illustrate quality integrated units designed to support high-level learning across disciplines.

Related Literature and Theoretical Framework

Recent years have seen a great deal of development in research on STEM integration. Moore et al. (2020) provide a thorough overview of definitions of STEM integration used in research, noting a wide spectrum of frameworks that differ on a variety of points (e.g., number of disciplines, levels of integration, implementation strategies). We view STEM integration as the intentional inclusion of two or more of science, technology, engineering, or mathematics in a lesson or unit.

One pertinent issue in the area is disciplinary content learning and alignment with subject-specific standards. In the Moore et al. (2014) framework for assessing and developing integrated STEM curricula, the fourth tenet states that "in order for the school learning to be meaningful and worth the time it takes to implement, there must be mathematics and/or science content as main objectives of the activity" (p. 43). Guzey et al. (2016) developed the STEM Integration Curriculum Assessment (STEM-ICA) with two items on the integration of science and mathematics content rated on a 5-point scale (0-not present through 4-excellent). However, in their study of 20 integrated units, none scored a 4 in science or math content, with only a fraction scoring a 3 in either subject. Estapa & Tank (2017) have noted similar concerns in a study on upper elementary STEM units during and after a STEM professional development experience: despite teachers incorporating science and math concepts in their engineering challenge, few lessons explicitly identified science or math content learning goals. When applying the STEM-ICA (Guzey et al., 2016) to the curricula, the authors found that even though the engineering design score average was 2, the average scores for the integration of mathematics and science content was 0.3 and 0.2 respectively.

In addition to content alignment, tasks used for instruction also determine the kinds of reasoning that students use as they learn. We follow Tekkumru-Kisa et al.'s (2020) definition of a task, which is "a segment of a classroom activity devoted to the development and assessment of a disciplinary idea and/or a practice" (p. 607). Researchers have used cognitive demand to describe the level or intensity of intellectual work required to complete a task (e.g., Bloodworth et al., 2023, Tekkumru-Kisa et al., 2015). Tasks that require higher levels of thinking and reasoning have a higher cognitive demand compared to tasks that require little thinking or reasoning such as tasks requiring students to only recall memorized facts. The Task Analysis Guide in Science (TAGS; Tekkumru-Kisa et al., 2015) and the Task Analysis Guide for Coding (TAG-C; Bloodworth et al., 2023) provide classification frameworks for characterizing the cognitive demand of tasks.

The TAGS framework has five levels of cognitive demand across four descriptive categories: Memorization Tasks (level 1), Tasks Involving Scripts (level 2), Tasks Involving Guidance for

Understanding (level 3 or 4), and Doing Science Tasks (level 5). Memorization Tasks require the reproduction of definitions or explanations of either scientific practices or content. Tasks Involving Scripts require following a script to work on some practice or content in a more integrated setting. Tasks Involving Guidance for Understanding are tasks that involve guidance for understanding practices, for understanding content, or for working with practices tied to specific science content. In terms of levels of cognitive demand, Tasks Involving Guidance for Understanding are level 3 if only practices or content is involved and are level 4 if the practices and content are robustly integrated. Doing Science Tasks are unstructured without clearly known procedures, requiring students to access knowledge and use scientific practices as they regulate, monitor, and adjust their approaches (Tekkumru-Kisa et al., 2015, p. 666).

In coding, Bloodworth et al. (2023) created the TAG-C framework with four levels of cognitive demand. Tasks with a low level of cognitive demand do not require students to think more deeply about coding to respond; in these tasks students may reproduce memorized code or copy a given code. For tasks with a medium-low level of cognitive demand, students must make decisions or put some lines of code together, but without thinking about the larger code structure or an entire scenario. The lines of code can be completed relatively independently with a clear script for accomplishing each part. Tasks with a medium-high level of cognitive demand require some aspect of thinking through the situation prior to writing the code or more thinking and reflection as they put the code together to make sure all problem constraints are met. These often include tasks that involve analyzing, comparing, or modifying code. Tasks that are a high level of cognitive demand go a step further in the required thinking; in these tasks students must create a program to solve a larger problem, and there are not known, clear-cut procedures to accomplish the task.

Lastly, research has demonstrated the value of argumentation in student learning (Cartier et al., 2013; Choresh et al., 2009) and teachers' facilitation of argumentation (McNeill et al., 2017). However, given how widely curriculum materials have been theorized as an important instrument for change in classroom activities, there are relatively few investigations done around opportunities for argumentation provided in curriculum materials. Stylianides (2009) studied the opportunities in a textbook where students can engage with various types of proofs or non-proof arguments. Bieda et al. (2014) built on Stylianides' (2009) work and proposed categories for the purpose of a curricular task related to argumentation, i.e., for students to make or justify a claim. Ayalon and Hershkowitz (2018) used these types of purposes (Bieda et al., 2014) to inform their study on teachers' perception of tasks that have potential for classroom argumentation.

Methodology

We partnered with a school district in rural Georgia to form interdisciplinary teams with third through fifth grade teachers to design integrated units for their students. The goal of the project was to create units that integrate coding into mathematics and science in a way that students can learn coding and either science or mathematics in the same unit but without an expectation that students would code in every lesson. The classroom teachers led the creation of the units and were supported by district specialists and researchers in mathematics education, science education, coding, and engineering. Our partner teachers implemented these units in the 2024–2025 school year and revised the units in the summer of 2025. The analysis presented here was conducted after the first revision, and before the second implementation.

The data analyzed included only the curricular units produced for the integrated mathematics and science units for grades 3, 4, and 5: one per discipline per grade level for a total of six units. (The six units are available at https://coe.uga.edu/research/c4oms-project.) Our partner teachers created these units around specific grade-level topics to align with their curriculum. Using the definition of task above, we identified the tasks across all six units and determined where to create divisions between tasks by evaluating (a) if the disciplinary idea or practice changes, (b) if the activity changes, and if unclear (c) based on how the teachers separated sections of the lesson within the curricular materials. We focus on the fourth-grade science unit as a sample case from the project. We examined each task in this unit to identify the elements of the Georgia fourth-grade science standards (SuitCASE, 2025) that were addressed. The Georgia standards (SuitCASE, 2023) include only one coding standard across grades three through five. Therefore, we chose to use the 10 standards published by the Computer Science Teacher Association (CSTA) classified under the Algorithms & Programming concept in the 3–5 grade band (CSTA K12 Standards, 2017). We examined the presence or absence of element(s) and/or standard(s) in each task.

For our analysis of the levels of science and coding cognitive demand, we used the frameworks from Tekkumru-Kisa et al. (2015) and Bloodworth et al. (2023). For science, we assigned codes 1 to 5 in the same fashion as Tekkumru-Kisa et al. for levels of cognitive demand from memorization to doing science. For coding, we used codes 1 to 4 to represent Bloodworth et al.'s levels of cognitive demand low, medium-low, medium-high, and high respectively. We used 0 to indicate a task that did not include coding. Through discussions, we determined the levels of cognitive demand of science and coding for each task by consensus within our team members.

We analyzed the opportunities for argumentation in the unit by adapting Bieda et al.'s (2014) framework. Bieda et al.'s framework was developed to analyze opportunities for reasoning and proof presented in student-facing curriculum materials while consulting other supporting materials. Because our units were written as teacher-facing materials along with student-facing materials, we considered both in our analysis. In other words, we considered both discussion prompts suggested to be given verbally to the students and prompts given directly to students in the materials, with the desired outcomes of those prompts. Our adaptation of Bieda et al.'s (2014) framework focused on the following two purposes of a task: (a) making claims and (b) making and justifying claims.

Results and Discussion

In this section, we first provide an overview of the fourth-grade science unit that contains 16 tasks in terms of standards alignment, cognitive demand, and opportunities for argumentation. Then, we briefly describe one integrated task in this unit in more detail to illustrate the way our partner teachers (together with the district specialists and researchers) designed these integrated tasks while maintaining cognitive demand and providing opportunities for argumentation.

The fourth-grade unit we analyzed addresses the life science standard "Obtain, evaluate, and communicate information about the roles of organisms and the flow of energy within an ecosystem" (SuitCASE, 2025, S4L1). This standard and all its four overlapping elements were addressed in this unit. Each task, except for 4S2 and 4S6, addresses at least one element. In 4S2, before students are introduced to the roles of organisms, students are asked to sort cards of organisms however they choose. The purpose of this task is to group organisms as a precursor to learning about scientific classification and thus does not directly address any fourth-grade standards. Task 4S6 is about living and non-living things, which is not a fourth-grade standard. The science content of the task was intentionally simplified to introduce students to a new coding concept.

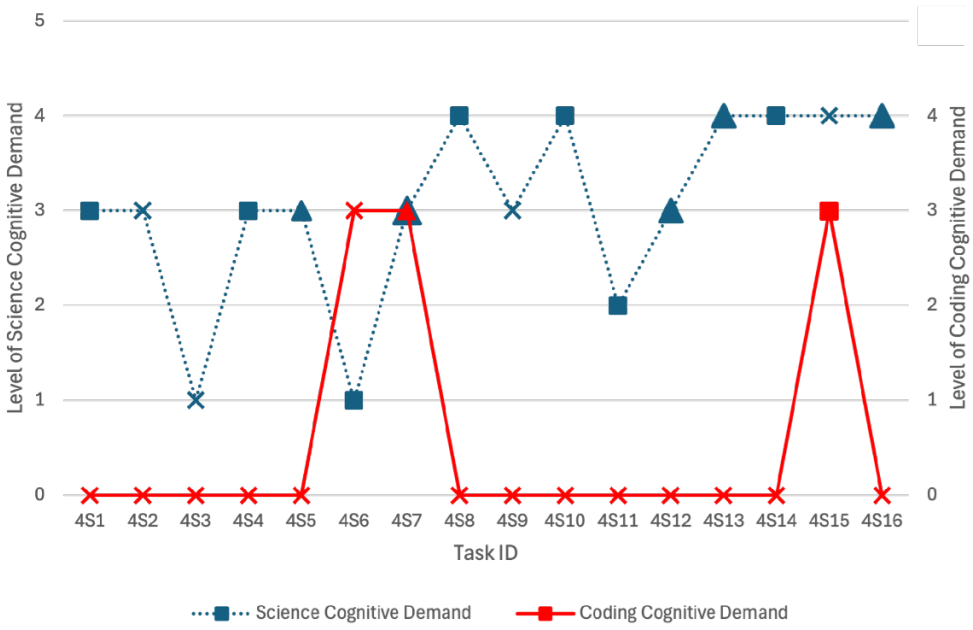
Three tasks in this unit integrated coding and science (4S6, 4S7, 4S15). These tasks addressed 6 out of 10 CSTA standards: create programs that use variables to store and modify data (1B-AP-09), create programs that include sequences, events, loops, and conditionals (1B-AP-10), decompose (break down) problems into smaller, manageable subproblems to facilitate the program development process (1B-AP-11), test and debug a program or algorithm to ensure it runs as intended (1B-AP-15), take on varying roles, with teacher guidance, when collaborating with peers during the design, implementation, and review stages of program development (1B-AP-16), and describe choices made during program development using code comments, presentations, and demonstrations (1B-AP-17).

All levels of science cognitive demand were present in this unit except for Doing Science. There are two Memorization Tasks, one Task Involving Scripts, seven Tasks Involving Guidance for Understanding in either practice or content, and six Tasks Involving Guidance for Understanding in practice and content (Figure 1). Three tasks integrated coding at a medium-high level. We see a variety of levels of cognitive demand in this unit, and these levels are generally on the higher end.

Of the 16 tasks in this unit, only three did not provide any opportunities for argumentation. Eight of the sixteen provided opportunities for making claims. Five tasks provided opportunities for making and justifying claims. Moreover, of the three tasks that are integrated, one provided opportunities for making claims in science, one provided opportunities for making claims in coding, and one provided opportunities for making and justifying claims in science and in coding.

Figure 1

Overview of Cognitive Demand and Opportunities for Argumentation



Note. A cross marker indicates no opportunities for argumentation. A square marker indicates opportunities for making claims. A triangle marker indicates opportunities for making and justifying claims. A 0 indicates coding was not present.

In this science unit, tasks range from Memorization to Guided Integration between scientific content and practice, with most of them in the higher levels of cognitive demand. Three of the sixteen tasks involved coding, and all three tasks are in the medium-high level of cognitive demand

for coding. In other words, three tasks (4S6, 4S7, 4S15) were integrated in a way that addressed standards in both science and coding while expecting high levels of thinking in coding.

We now briefly describe task 4S7, which is cognitively demanding in both science and coding and provides opportunities for students to make and justify science and coding claims (see Figure 1). In task 4S7 students are given a grid with a few organisms depicted on various grid locations. Pairs of students are asked to program a small educational robot to travel across the grid, stopping at each organism and performing a pre-determined action that would indicate the organism's classification as a producer, consumer, or a decomposer. One of the organisms on the grid is a vulture, a type of bird whose diet is animal carcasses. Our partner teachers expected some students would classify it as a consumer, while others might classify it as a decomposer. Therefore, they wrote into the unit that "If students classified it [vulture] as a decomposer, they should be able to explain that it is because they eat dead animals. If they classified it as a consumer, they should be able to explain that it is because they eat other animals." This signals the teacher that (a) the central science concept is the characteristics of the roles of organisms, and (b) students are expected to make and justify their claim of a specific role using a characteristic of that role.

For coding, after a student pair successfully coded their robot, they would then run another student pair's code. Students would discuss with their partner similarities and differences between their own code and the other code. Students would then suggest one way to make the other code more accurate or efficient. Here, students make claims about accuracy and efficiency of some code. Although they do not necessarily justify this claim to another student-pair, they justify the claim between themselves before writing it down as a suggestion. Overall, we see task 4S7 as an exemplar of high-quality integration because of the cognitively demanding focus on standards and opportunities for students to make and justify their claims in both science and coding.

Implications

Instead of two disciplines being put in competition with each other to be the instructional focus, our analysis illustrates that students can be meaningfully engaged in higher cognitive demand tasks to learn content in multiple disciplines simultaneously. It is not necessary to completely separate learning in science and learning in coding. Rather than Venville et al.'s (2002) zero-sum game, intentional attention to cognitive demand can result in teacher-designed activities that require thinking about important concepts in, for example, both science and coding. Additionally, attending to principles of argumentation can result in students having opportunities to make and justify claims

both within and across disciplines. All of this can be done with attention to state and national curriculum standards. Future attention to the combination of cognitive demand and argumentation across disciplines may result in higher-quality interdisciplinary lessons and units.

References

Ayalon, M., & Hershkowitz, R. (2018). Mathematics teachers' attention to potential classroom situations of argumentation. *The Journal of Mathematical Behavior*, *49*, 163–173. https://doi.org/10.1016/j.jmathb.2017.11.010

Bieda, K. N., Ji, X., Drwencke, J., & Picard, A. (2014). Reasoning-and-proving opportunities in elementary mathematics textbooks. *International Journal of Educational Research*, *64*, 71–80. https://doi.org/10.1016/j.ijer.2013.06.005

Bloodworth, A., Conner, A., Miller, C., Franco, L., Foutz, T., & Hill, R. B. (2023). Robotics and coding: A framework for examining cognitive demand. *Journal of Technology Education*, *35*(1), 7–31. https://doi.org/10.21061/jte.631

Cartier, J. L., Smith, M. S., Stein, M. K., & Ross, D. K. (2013). *5 practices for orchestrating productive task-based discussions in science*. National Council of Teachers of Mathematics.

Choresh, C., Mevarech, Z. R., & Frank, M. (2009). Developing argumentation ability as a way to promote technological literacy. *International Journal of Educational Research*, *48*(4), 225–234. https://doi.org/10.1016/j.ijer.2009.10.002

CSTA K12 Standards. (2017). *CSTA K–12 Computer Science Standards*. https://csteachers.org/k12standards/interactive/

Estapa, A. T., & Tank, K. M. (2017). Supporting integrated STEM in the elementary classroom: A professional development approach centered on an engineering design challenge. *International Journal of STEM Education*, *4*(1), Article 6. https://doi.org/10.1186/s40594-017-0058-3

Guzey, S. S., Moore, T. J., & Harwell, M. (2016). Building up STEM: An analysis of teacher-developed engineering design-based STEM integration curricular materials. *Journal of Pre-College Engineering Education Research (J-PEER)*, *6*(1), Article 2. https://doi.org/10.7771/2157-9288.1129

McNeill, K. L., González-Howard, M., Katsh-Singer, R., & Loper, S. (2017). Moving beyond pseudoargumentation: Teachers' enactments of an educative science curriculum focused on argumentation. *Science Education*, *101*(3), 426–457. https://doi.org/10.1002/sce.21274

Moore, T. J., Johnston, A. C., & Glancy, A. W. (2020). STEM integration: A synthesis of conceptual frameworks and definitions. In C. C. Johnson, M. J. Mohr-Schroeder, T. J. Moore, & L. D. English (Eds.), *Handbook of research on STEM education* (pp. 3–16). Routledge. https://doi.org/10.4324/9780429021381-2

Moore, T. J., Stohlmann, M. S., Wang, H.-H., Tank, K. M., Glancy, A. W., & Roehrig, G. H. (2014). Implementation and integration of engineering in K–12 STEM education. In Ş. Purzer, J. Strobel, & M. E. Cardella (Eds.), *Engineering in pre-college settings: Synthesizing research, policy, and practices* (pp. 35–60). Purdue University Press. https://doi.org/10.2307/j.ctt6wq7bh.7

National Research Council. (2013). *Next Generation Science Standards: For states, by states*. The National Academies Press. https://doi.org/10.17226/18290

Stylianides, G. J. (2009). Reasoning-and-proving in school mathematics textbooks. *Mathematical Thinking and Learning*, *11*(4), 258–288. https://doi.org/10.1080/10986060903253954

SuitCASE. (2023). *Computer Science - Georgia Standards of Excellence*. Georgia Department of Education, SuitCASE. https://case.georgiastandards.org/00fcf0e2-b9c3-11e7-a4ad-47f36833e889/

SuitCASE. (2025). *Science - Georgia Standards of Excellence*. Georgia Department of Education, SuitCASE. https://case.georgiastandards.org/27a08dc6-416e-11e7-ba71-02bd89fdd987/27a08dc6-416e-11e7-ba71-02bd89fdd987

Tekkumru-Kisa, M., Stein, M. K., & Doyle, W. (2020). Theory and research on tasks revisited: Task as a context for students' thinking in the era of ambitious reforms in mathematics and science. *Educational Researcher*, *49*(8), 606–617. https://doi.org/10.3102/0013189X20932480

Tekkumru-Kisa, M., Stein, M. K., & Schunn, C. (2015). A framework for analyzing cognitive demand and content-practices integration: Task analysis guide in science. *Journal of Research in Science Teaching*, *52*(5), 659–685. https://doi.org/10.1002/tea.21208

Venville, G. J., Wallace, J., Rennie, L. J., & Malone, J. A. (2002). Curriculum integration: Eroding the high ground of science as a school subject? *Studies in Science Education*, *37*(1), 43–83. https://doi.org/10.1080/03057260208560177

Acknowledgements

This paper is based on work supported by the National Science Foundation under Grant No. 2318287. Opinions, findings, and conclusions in this paper are those of the authors.

TASK PERCEPTION AS A DIMENSION OF STEM TEACHER IDENTITY: INSIGHTS FROM INSTRUCTIONAL ARTIFACTS IN A PHYSICS-TO-CLASSROOM CAREER EXPLORATION PATHWAY

Terrie M. Galanti
University of North Florida
terrie.galanti@unf.edu

W. Brian Lane
University of North Florida
.lane@unf.edu

J. Caleb Speirs
University of North Florida
j.caleb.speirs@unf.edu

Abstract

The shortage of high school physics teachers in the United States has increased interest in alternative certification pathways for STEM graduates. Undergraduate STEM programs can cultivate student interest in these pathways by providing opportunities to explore teaching. This qualitative study elaborates task perception, a dimension of STEM teacher identity, through analysis of instructional artifacts developed in a Physics-to-Classroom career exploration initiative. Analysis revealed three complementary conceptions of the teacher's role: designing student-centered learning environments, facilitating reasoning and sensemaking, and brokering disciplinary learning. These findings extend understandings of task perception and have implications for designing STEM teaching career exploration pathways.

Keywords: STEM teacher identity, task perception, productive disciplinary engagement, teacher education, physics education

Introduction

Persistent shortages of well-prepared secondary mathematics and science teachers have prompted increasing reliance on alternative certification pathways for STEM graduates. Building a STEM teaching workforce is increasingly difficult due to declining enrollment in teacher preparation programs and persistent perceptions that teaching is a less prestigious and less desirable career than other STEM professions (Larsson et al., 2021). At the same time, undergraduate STEM majors often have few opportunities to explore teaching because intensive disciplinary coursework leaves little room for education coursework or authentic teaching experiences. While school districts seek to hire individuals who have strong STEM backgrounds, these graduates often lack pedagogical training and may be less likely to persist in the teaching profession (Solis Rodriguez, 2021).

In response to these challenges, we developed a Physics-to-Classroom career exploration pathway to provide undergraduate physics majors with structured opportunities to engage with secondary STEM content and pedagogy. This pathway was designed to support the development of

STEM teacher identity (Galanti & Holincheck, 2022) by encouraging students to see themselves as both capable STEM disciplinary thinkers and future STEM teachers. Participants earned digital badges focused on physics content for teaching and computer programming for STEM teaching, engaged in a STEM teaching methods course, and worked closely with both STEM and education faculty mentors. This program structure created opportunities for undergraduate STEM majors to integrate their content expertise with effective teaching practices.

Objectives of the Study

The purpose of this study was to examine how instructional artifacts developed by participants in the Physics-to-Classroom career exploration pathway provide insight into task perception, one dimension of STEM teacher identity. Through analysis of participants' STEM lesson planning and computational modeling activities, we sought to elaborate the concept of task perception by describing how participants positioned the teacher in supporting students' STEM reasoning, sensemaking, and computational thinking. This analysis allows us to answer the research question: *How do program artifacts express participants' task perception as a dimension of STEM teacher identity?*

Theoretical Framework

Understanding how STEM majors come to view teaching as a viable professional pathway requires attention not only to what they learn, but also to how they see themselves within the profession. STEM teachers who emphasize student-centered and pedagogical-centered identities, rather than subject-centered identities, may be more likely to persist in their teaching careers (Olitsky et al., 2020). Consequently, STEM teacher identity is valuable for understanding how the STEM majors in our study made sense of the work of teaching.

Task Perception as a Dimension of STEM Teacher Identity

We conceptualize STEM teacher identity as the extent to which undergraduate STEM majors see themselves as legitimate participants in both disciplinary communities and the teaching profession. Teacher identity has been shown to influence teacher-student relationships, teacher well-being, persistence, and enthusiasm for the teaching profession (Tsang & Jiang, 2018; Zembylas, 2003). Identity is a multidimensional construct comprising learner dimensions (performance, competence, interest, and recognition) and teacher dimensions (self-image, self-efficacy, teaching interest, motivation, recognition, and task perception) (Galanti & Holincheck, 2022).

Task perception refers to teachers' beliefs about how they create meaningful learning opportunities and is especially relevant to this study because it reflects how STEM majors exploring teaching conceptualize their roles in a future classroom. Undergraduate STEM majors are likely to enter the pathway with well-developed STEM learner identities but fewer opportunities to develop identities as teachers. While disciplinary competence and interest are reinforced through STEM coursework, teacher identity is cultivated through opportunities to enact teaching practices, receive feedback, and gain recognition as educators (Izadinia, 2015).

Describing Task Perception Through the Lens of Productive Disciplinary Engagement

We draw on the Productive Disciplinary Engagement ([PDE], Engle & Conant, 2002) framework to further specify how teachers create opportunities for students to interact with content and how students are positioned to take up these opportunities. In student-centered classrooms, STEM teachers take on roles as facilitators who encourage students to take ownership for meaningful learning and ask guiding questions. They model problem solving as a collaborative, iterative process (Keiler, 2018). PDE describes how STEM learning environments should be constructed around four interrelated principles: problematizing, authority, accountability, and resources. Teachers **problematize** learning by designing intellectually challenging tasks that invite multiple solution pathways, encourage productive struggle, and navigate uncertainty through reasoning and justification (Holincheck et al., 2024; Koretsky et al., 2021). Creating an environment in which students have the **authority** to propose, test, and refine ideas while making their thinking visible to others is balanced with **accountability** as teachers establish norms for all students to challenge ideas collaboratively and build understanding (Agarwal & Sengupta-Irving, 2018). Teachers also provide time, tools, and representations as **resources** to promote disciplinary sensemaking. The STEM teacher identity framework identifies task perception as a dimension of teacher identity but does not specify how meaningful learning opportunities are created in practice. PDE provides an analytic structure for describing how teachers position students for meaningful disciplinary learning. In this study, we were particularly interested in how STEM majors conceptualized their roles and responsibilities in supporting high school students' reasoning and sensemaking.

Methodology

This study used a qualitative approach (Merriam & Tisdell, 2015) to understand how participants' task perception was developed within the digital badges and STEM methods

coursework in a Physics-to-Classroom Teaching Exploration Pathway. We report our findings for the first of two cohorts for this federally funded physics teacher preparation initiative.

Participants and Context

In Spring 2025, participants were recruited into the inaugural cohort of a Physics-to-Classroom career exploration pathway designed to provide undergraduate STEM majors with structured opportunities to explore secondary STEM teaching. Eligible participants included STEM-interested students from regional two-year colleges, undergraduate STEM majors, and undergraduate STEM education majors. The first cohort included four participants: physics majors Brendan and Felicia and transfer students Maura (biology) and Robert (mathematics).

The pathway integrated physics content and computational thinking (Lane et al., 2023) with STEM pedagogies. Participants completed asynchronous self-paced Canvas modules in *Physics Content for Teaching* badge (Summer 2025) and *VPython Visualization for STEM Learning and Teaching* badge (Fall 2025). Weekly Zoom cohort meetings were hosted to monitor and motivate individual progress. This pathway was designed to strengthen disciplinary knowledge for teaching while encouraging participants to consider how physics concepts, computational modeling, and student reasoning could be represented in classrooms. The VPython badge culminated in the creation of an original three-dimensional physics visualization and presentation to project faculty over Zoom. Participants engaged with physics content and computational modeling from the perspective of teaching rather than solely as disciplinary learners.

Our participants also completed a three-credit STEM teaching methods course (Fall 2025), *Problem-based Instruction in Math and Science,* taught by the first author. The course emphasized problem-based instruction, computational thinking, and student-centered STEM pedagogy through teaching rehearsals and reflection. As a culminating assignment, participants developed a three-lesson instructional unit in the STEM discipline of their choice. Rather than designing lessons from scratch, participants adapted lessons from OpenSciEd or Open Up Resources Mathematics, allowing them to focus on internalizing and enacting high-quality instructional practices embedded within these curricula. Lesson planning was further scaffolded using templates structuring the teachers' role in orchestrating productive discussions (Smith et al., 2008).

Data Collection and Analysis

The primary data sources were participants' STEM instructional units developed in the methods course and their VPython three-dimensional physics visualizations and presentation

transcripts. Participants responded to formal questions about STEM teaching in the lesson design assignment and informal questions during the VPython presentation. These artifacts required participants to articulate explicit instructional decisions about supporting secondary students' STEM learning and therefore provided evidence of task perception. Participant interviews and transcripts of weekly cohort meetings were used as secondary data sources to triangulate the findings.

We collaboratively coded the two assignments, using both categorizing and connecting strategies (Maxwell & Miller, 2008) and resolving differing interpretations through discussion. The first author used her insights as the course instructor to provide interpretive context while the second author used his positioning outside the course to interrogate interpretations. In our first cycle of analysis, we used the four PDE principles as provisional codes (Saldaña, 2021). In our second cycle of analysis, we used pattern coding to condense and organize our first-cycle codes, identifying themes describing how participants conceptualized the role of a STEM teacher.

Results and Discussion

Although the four PDE principles served as our initial analytic codes, participants rarely expressed these principles independently. Instead, their instructional artifacts revealed three broader conceptions of the teacher's role that provide insight into task perception.

STEM Teacher as Designer of Student-Centered Learning Environments

The artifacts represented the teacher as an intentional designer of learning environments that create opportunities for students to engage meaningfully with STEM ideas. Participants took up this role by selecting or adapting inquiry-driven tasks and structuring experiences and resources to support students in thinking like scientists or mathematicians. Maura and Felicia designed unit plans that intentionally positioned students as active participants in disciplinary inquiry. Felicia's physical sciences unit plan **problematized** the analysis of real-world exoplanet data to infer environmental conditions and to design a hypothetical life form capable of surviving those conditions. She conferred intellectual **authority** and planned to hold students **accountable** for explaining and justifying their reasoning by asking questions such as, "How did two groups use the same data but reach different conclusions?" and "What makes a claim stronger in science: creativity, confidence, or evidence from data?"

Similarly, Maura designed a biology lesson in which science notebooks, chart-paper models, gallery walks, and oral presentations served not simply as instructional activities but as **resources** for students to explain, critique, and refine one another's thinking. Reflecting on her lesson design, she

explained, "These supports ensure students have multiple entry points into the scientific practice of modeling and can reason about natural selection even if they struggle with reading or writing." Maura also demonstrated design orientation in her VPython presentation. When asked how she might use computational modeling with students, Maura described how she would **problematize** the subject of rotational motion by providing starter code (**resource**) and scaffolding problem complexity to support students in refining their own models (**authority**). These artifacts suggest that participants viewed lesson design as an intentional process of creating conditions for PDE. Their task perception was evidenced in structuring meaningful resources and designing opportunities for collaborative STEM reasoning more than organizing the delivery of content.

STEM Teacher as Facilitator of Reasoning and Sensemaking

Lesson plans depicted the teacher as a facilitator of reasoning and sensemaking authentic to the discipline. For our participants, the instructional decisions embedded within these artifacts reflected a vision for how teachers would actively facilitate students' reasoning. Brendan elaborated on how he **problematized** studying cloud formations and what actions he would take to hold his Grade 6 students **accountable** for learning beyond the prompts in the template. He described selecting and sequencing student calculations of rainfall surface area and volume during a whole-class discussion. He expected students to compare solution strategies, critique one another's representations, and refine their understanding of how both rainfall depth and surface area determine total volume. He held his students **accountable** for thinking like scientists by asking, "What questions do you have about why some places have heavier storms than others?"

Maura expected her students to collaboratively construct a consensus model of evolutionary change (**accountability**, **authority**) as they "negotiate, debate, and combine ideas to create one large model that represents shared understanding." Reflecting on this activity, she emphasized that it required "explanation, argumentation, and collaboration—mirroring real scientific practice." These artifacts illustrate how participants saw the role of the teacher in actively orchestrating disciplinary discourse rather than simply engaging students in interesting STEM activities. Participants viewed themselves as responsible for facilitating discussion and collaborative sensemaking to position their students as producers of scientific and mathematical ideas.

STEM Teacher as Broker of Disciplinary Knowledge

In some instances, participants saw the teacher as a broker of disciplinary learning who helps students access, interpret, and connect complex STEM concepts. We saw evidence of this role

throughout Robert's unit plan, which provided real-world contexts and multiple representations as **resources** to build student understanding of piecewise functions. Robert's plan emphasized the importance of explaining connections between representations for his students (Sandoval & Possani 2016). He positioned himself as a content expert, using language such as "Students must write and describe equations." This approach might limit opportunities for true **problematizing** with little space for uncertainty, justification, and argumentation. Similarly, during his VPython presentation visualizing a Taylor-Maclaurin series, Robert needed suggestions from faculty to think about the investigatory potential of his dynamic graphical model to accomplish more than illustrating a set number of terms in a series. Even after the first author prompted him to consider the **problematizing** potential for student exploration in an AP Calculus classroom, Robert continued to see his model as a tool to show students a mathematical concept.

When discussing their VPython projects, both Brendan and Felicia expressed hesitation about using their models with students because of the limits of their own knowledge. Felicia would only provide students with "fully formed code that they can just do minor tweaks on so they can see the differences." Brendan explained, "Although I can code things, I code in a very slow and tedious manner, because typically I run into errors, and then I go, okay, what happened? And then I try and fix the errors." While modeling iterative development and debugging is an important theme of computational thinking instruction (Weller et al., 2022), in Brendan's and Felicia's cases, carrying out these processes with students might undermine their role as the local context expert. This emphasis may constrain opportunities to shift problem-solving **authority and accountability** to students.

Implications

Our findings show how these participants conceptualized meaningful learning opportunities through three overlapping roles that can vary depending on the topic or lesson. Even among this small sample, this heterogeneity suggests that an individual's task perception can encompass multiple roles. This outcome points to task perception being a manifold quality delineated by PDE. These roles position teaching as the purview of both a content expert (a broker mediating STEM concepts for students or a designer creating opportunities to meaningfully engage with STEM ideas) and a facilitator trained in pedagogical practices (a designer creating student-centered learning environments or a facilitator of student reasoning). Teacher educators can make these roles explicit as STEM majors explore how their content expertise can support meaningful disciplinary learning. This exploration positions the teaching profession as an intellectually rigorous and socially impactful

application of disciplinary expertise, potentially strengthening the degree to which STEM majors see themselves as both content experts and impactful teachers.

References

Agarwal, P., & Sengupta-Irving, T. (2018). Power in the digital age: A critical revision to productive disciplinary engagement (PDE). In J. Kay and R. Luckin (Eds.). *Rethinking learning in the digital age: Making the learning sciences count.* In 13th International Conference of the Learning Sciences (Vol. 2). International Society of the Learning Sciences. https://repository.isls.org//handle/1/504

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and Instruction, 20(*4), 399–483. https://doi.org/10.1207/S1532690XCI2004_1

Galanti, T. M., & Holincheck, N. (2022). Beyond content and curriculum: Conceptualizing the cultivation of integrated STEM teacher identity. *International Journal of STEM Education, 9*, Article 43. http://doi.org/10.1186/s40594-022-00358-8

Holincheck, N., Kraft, T., Galanti, T. M., Baker, C. K., & Nelson, J. K. (2024). "Everybody was included in the conversation": Teachers' perceptions of student engagement as a path to transdisciplinary STEM learning in diverse elementary schools. *Education Sciences, 14*(3), Article 242. https://doi.org/10.3390/educsci14030242

Izadinia, M. (2015). A closer look at the role of mentor teachers in shaping preservice teachers' professional identity. *Teaching and Teacher Education, 52*, 1–10. https://doi.org/10.1016/j.tate.2015.08.003

Keiler, L. S. (2018). Teachers' roles and identities in student-centered classrooms. *International Journal of STEM Education, 5,* Article 34. https://doi.org/10.1186/s40594-018-0131-6

Koretsky, M. D., Vauras, M., Jones, C., Iiskala, T., & Volet, S. (2021). Productive disciplinary engagement in high- and low-outcome student groups: Observations from three collaborative science learning contexts. *Research in Science Education, 51*(Suppl 1), 159–182. https://doi.org/10.1007/s11165-019-9838-8

Lane, W. B., Galanti, T. M., & Rozas, X. (2023). Teacher re-novicing on the path to integrating computational thinking in high school physics classrooms. *Journal for Research in STEM Education, 6*, 302–325. https://doi.org/10.1007/s41979-023-00100-1

Larsson, J., Airey, J., & Lundqvist, E. (2021). Swimming against the tide: Five assumptions about physics teacher education sustained by the culture of physics departments. *Journal of Science Teacher Education, 32*(8), 934–951. https://doi.org/10.1080/1046560X.2021.1905934

Maxwell, J. A., & Miller, B. A. (2008). Categorizing and connecting strategies in qualitative data analysis. In S. N. Hesse-Biber & P. Leavy (Eds.), *Handbook of emergent methods* (pp. 461–477). The Guilford Press.

Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation.* John Wiley & Sons.

Olitsky, S., Perfetti, A., & Coughlin, A. (2020). Filling positions or forging new pathways? Scholarship incentives, commitment, and retention of STEM teachers in high-need schools. *Science Education, 104*, 113–143. https://doi.org/10.1002/sce.21552

Saldaña, J. (2021). The coding manual for qualitative researchers. Sage.

Sandoval, I., & Possani, E. (2016). An analysis of different representations for vectors and planes in $\mathbb{R}^3$: Learning challenges. *Educational Studies in Mathematics, 92(*1), 109–127. http://www.jstor.org/stable/24830680

Solis Rodriguez, J. (2025). A systematic review of the recent empirical literature on math and science teacher recruitment and retention. *Education Sciences, 15(*8), Article 1073. https://doi.org/10.3390/educsci15081073

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*(4), 313–340. https://doi.org/10.1080/10986060802229675

Tsang, K. K., & Jiang, L. (2018). Positive emotional experiences in teaching, teacher identity, and student behaviors: A symbolic interactionist perspective. *Schools, 15*(2), 228–246. https://doi.org/10.1086/699890

Weller, D. P., Bott, T. E., Caballero, M. D., & Irving, P. W. (2022). Development and illustration of a framework for computational thinking practices in introductory physics. *Physical Review Physics Education Research, 18*(2), Article 020106.

Zembylas, M. (2003). Emotions and teacher identity: A poststructural perspective. *Teachers and Teaching, 9*(3), 213–238. https://doi.org/10.1080/13540600309378

Acknowledgements

This work is supported by the National Science Foundation and the Physics Teacher Education Coalition under grant no. 2325980. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

SECTION II: MATHEMATICS

A FRAMEWORK TO ASSESS PROSPECTIVE TEACHERS' ASSESSMENT OF STUDENT THINKING

Nama Namakshi
University of Arkansas
namakshi@uark.edu

Hiroko Warshauer
Texas State University
hw02@txstate.edu

Abstract

We present the findings of a study that investigated 44 PSTs' assessment skills using an adapted framework from Science Education. Subtle differences in the PSTs' assessment skills along 8 dimensions were captured. Results show that PSTs' overall assessment skills are at the emerging level while slightly favoring the interpretation category. Our findings suggest that by capturing the subtle differences in assessment skills teacher educators can better support the development of PSTs' assessment reasoning.

Introduction and Objective of the Study

Formative assessment is critical for improving teaching effectiveness and student learning outcomes (Black & Wiliam, 2009). However, prospective teachers need support in developing such an important pedagogical skill which involves interpreting students' mathematical thinking (Boerst et al., 2019). Our study aims to characterize and differentiate how Pre-Service Teachers (PSTs) interpret students' mathematical thinking when analyzing students' written work on a mathematical task in the domain of fractions. The mathematics content course which serves as the setting for this study introduces PSTs to the required mathematics content with an eye towards teaching and assessing elementary students' mathematical thinking. Our study is guided by the research question: What are the levels of the PSTs' overall assessment skills along the different dimensions within the categories of evaluation and interpretation when analyzing students' mathematical thinking on a written task in the domain of fractions?

Literature and Conceptual Framework

Assessment of student work entails understanding and interpreting student thinking. Researchers (e.g., Namakshi et al., 2022; Lee & Lim, 2020; Wilson et al., 2013) have examined formative assessment and the value of examining student work. Wilson, Mojica & Confrey (2013) suggest that examining student work provides a way for teachers to build a model of student

knowledge. Thus, the use of student work gives PSTs insight into students' mathematical thinking and a way to learn how to assess students' thinking. In the area of fractions, studies show the challenges in teaching and learning fractional concepts that encompass both procedures and concepts (Copur-Genturk, 2021; Perry, 2023; Thurtell et al., 2019). For example, in the study conducted by Philipp et al. (2007), PSTs struggled with the use of representations such as number lines or area diagrams. In addressing these challenges, Thanheiser et al. (2014) found that opportunities to examine models and different perspectives about fractions provided ways to support PSTs' conceptual understanding of fractions and relate them to the procedures. Sleep (2012) also found that the examination and discussion of student work supported PSTs' pedagogical content knowledge as well as assessment. In our previous work (Namakshi et al., 2022), the qualitative findings identified the value of allowing PSTs to enact the tasks themselves, then share and discuss their strategies in class. This grounded the PSTs in their understanding of the content and the different approaches that students may use. These studies inform our current work on the development of assessment skills among PSTs.

Our prior work (Namakshi et al., 2022) used the Teacher Noticing Framework by Jacobs et al. (2010) to examine the Attend and Interpret dimensions of teacher noticing. However, we recognized a need to differentiate PSTs' assessment reasoning that can better capture the subtle differences in the PSTs' assessment reasoning skills, particularly in interpreting students' thinking (Boerst et al., 2019). Below we describe the adapted/modified version of the Talanquer (2015) Science Education Framework that we are using to analyze data in our current study. The Talanquer Framework provides more layers of differentiation in the scoring of 'Attend' and 'Interpret' features of the Jacobs et al. (2010) teacher noticing framework thereby allowing us to better capture more detailed differences in PSTs' assessment capabilities.

Talanquer et al. (2015) Framework Background and Explanation of the Adapted Version

In our framework there are four evaluative dimensions (evaluation stance, specificity in the evaluation, analysis of coherence, use of evidence) and four interpretive dimensions (quality of interpretation, productive thinking, scope of the evaluation, mathematical accuracy) which we have aligned with the 'Attend' and 'Interpret' features of the Jacobs et al. (2010) teacher noticing framework respectively. Jacobs et al. (2010) describe the 'Attend' component of their framework as the extent to which teachers attend to the mathematical details in children's strategies. By extending upon this definition in the context of assessing students' written work, we define 'Evaluative criteria'

as a mix of descriptive and inferential analysis of students' written work which includes assessing student work along the following four dimensions:

- Evaluative stance - identifying and unpacking the strategies used by the student.
- Specificity in evaluation - generating specific claims about the student's mastery of the concepts/learning goals that are addressed in the problem.
- Analysis of coherence - identifying and unpacking consistencies and/or inconsistencies with respect to conceptions or mistakes in the student response.
- Use of evidence – supporting the assessment of student work along the three dimensions described above by evidence gathered from student's answers.

Jacobs et al. (2010) describe the 'Interpret' component of their framework as "how teachers interpret children's understandings as reflected in their strategies" (p. 172). By further extending upon this definition, we define the 'Interpretive criteria' as largely inferential analysis of students' written work upon which interpretations of students' mathematical understanding of concepts addressed by the problem or otherwise are built. By interpretation of student understanding we mean, if the PSTs can explain what understanding/misconception of the mathematical concept led the student to the response they gave. Additionally, interpretation of student work requires PSTs to draw conclusions about students' mathematical understanding in a broader sense of concepts that are (or are not) being specifically targeted in the problem. This includes assessing student work along the following four interpretive dimensions:

- Quality of interpretation – interpreting student's mathematical understanding of concepts addressed by the problem and the overall task in a way that can be reasonably inferred from student work.
- Productive thinking – focusing on aspects of student work that are most productive for assessing and interpreting student understanding of concepts being addressed in the problem/overall task.
- Scope of the evaluation – recognizing relevant understandings/ misunderstandings in areas explicitly targeted by the problem and in a broader sense of concepts that are (or are not) being specifically targeted in the problem and overall task.
- Mathematical accuracy – building mathematically accurate interpretations when assessing student work along the three dimensions described above.

Tables 1 and 2 provide a sample adaptation for each of the evaluative and interpretive dimensions by levels of sophistication in PSTs' assessment of student thinking.

Table 1

*Evaluative Dimensions (*Soundness of Evaluations Characterized Along These Dimensions*) and Levels of Sophistication in PSTs' Assessment of Student Strategies*

Evaluative dimensions	Novice [0,1]	Emerging (1,2]	Advanced (2,3]
Evaluative stance	Evaluation is mostly descriptive based on students' statements, without attending to student thinking: Focus is on correct/ incorrect answer, on what students' fail to do or say, without searching for meaning in students' answers or strategies used.	Evaluation is a mix of descriptive and inferential statements: Looking for right answers but finding some meaning in students' responses or strategies used.	The core of the evaluation focuses on making inferences of student thinking: Analyzing students' answers and strategies used - looking to understand the reasoning behind them, sometimes independently of whether they are right or wrong.

Table 2

*Interpretive dimensions (*soundness of interpretations characterized along these dimensions*) and levels of sophistication in PSTs' assessment of student understanding.*

Interpretive dimensions	Novice [0,1]	Emerging (1,2]	Advanced (2,3]
Quality of the interpretation	Interpretations cannot be supported with the evidence available, or claims are minimally supported by re-stating students' answers without unpacking the meaning in students' responses.	Interpretations are built, but the evidence provided is limited, superficial or formulaic.	Reasonable interpretations of student understanding are built given the available evidence.

Methodology

Participants and Setting

We conducted this study in a large Hispanic-Serving public university in the southern United States during the Fall 2021 semester in a mathematics content course for prospective elementary and middle school teachers. The course is the first of a two-course mathematics sequence for PSTs in preparation for elementary and middle school teaching. The goal of the course is to develop PSTs' proficiency in understanding the underlying concepts of school mathematics with an emphasis

placed on examining a variety of standard and non-standard algorithms and using multiple strategies for problem solving. Building a holistic and conceptual understanding of fractions and rational numbers through various models such as geometric/visual models and the number line is one of the main objectives of the course. The course is intended to provide a strong foundation in the PSTs' common content knowledge while also focusing on the mathematical knowledge needed for teaching. The participants in the study consisted of 44 PSTs (with vast majority listed as elementary education majors) enrolled in two sections of the mathematics content course taught by the same instructor with no prior experience with assessing students' mathematical thinking by analyzing student work.

Data Collection

The data consisted of PSTs' (N=44) written assessment of a middle school student's work on a formative assessment task (Schoenfeld, 1999) focused on identifying fractional values of area shapes. The student work that PSTs were asked to analyze was focused on identifying the fractional parts of a square subdivided into four equal smaller squares and each of the smaller squares divided into fractional pieces where not all the fractional pieces were congruent shapes. The PSTs had an opportunity to do the activity in class and then share and discuss their work as a whole class prior to completing their analysis and written reports on the work of a middle-school student.

Data Analysis

To answer the research question, we analyzed the PSTs' written assessments of student work using deductive content analysis (Krippendorff, 2019) based on our adapted version of Talanquer's Framework (Talanquer et al., 2015) as described in the previous section. Using [0, 3] as our scoring interval, we designated [0,1) for Novice, [1,2) for Emerging, and [2,3] for Advanced level in evaluating and interpreting student thinking as noted in Tables 1 and 2. The authors first met to establish the modified codes described above and coded six PSTs' written assessment reports together to gain consensus on the codes using the framework. We then separately coded the remaining 38 participants' written assessments. After coding, we met to reconcile any differences in each of the 38 assessments and came to an agreement. Throughout the coding process we maintained 90% or better reliability.

Findings

Using Talanquer's Framework (2015), our findings were able to capture the subtle differences in the PSTs' assessment reasoning skills along four dimensions within the category of evaluation/attend and four dimensions within the category of interpretation in analyzing students' mathematical thinking on a written task in the domain of fractions. Furthermore, we were able to differentiate our PSTs' assessment skills into three levels of sophistication for evaluating and interpreting student thinking along these dimensions. On average PSTs' (N = 44) score within the interval [0, 3] on the construct of overall assessment skills came out to be 1.53, which puts them on average in the middle of the emerging level spectrum. Table 3 provides a summary of the average scores of the PSTs along different dimensions within the two categories of the overall assessment skills - evaluation/attend and interpretation.

When differentiating the PSTs' overall assessment skills into three levels of sophistication for evaluating and interpreting student thinking along these dimensions, we found that 34% (15 PSTs) were at the Novice level, 39% (17 PSTs) were at the Emerging level, and 27% (12 PSTs) were at the Advanced level. Furthermore, PSTs' overall interpretation skill average score came out to be slightly more (4.7% more) than their overall evaluative/attend skill average score.

Table 4 provides sample excerpts of the PSTs' analysis at different levels of sophistication along the 'Quality of the interpretation' dimension using the coding scheme given in Table 1.

Table 3

PST Average Scores on Evaluative and Interpretive Dimensions

Evaluative/ attend dimensions	Evaluative stance	Specificity in evaluation	Analysis of coherence	Use of evidence	Overall evaluative/ attend skill
Average Score	1.54	1.46	1.51	1.46	1.49
Interpretive dimensions	Quality of interpretation	Productive thinking	Scope of evaluation	Math accuracy	Overall interpretation skill
Average Score	1.60	1.54	1.47	1.61	1.56

Table 4

Examples of levels of assessment skills among the PSTs

Levels of assessment	Excerpts from PSTs' written evaluations
Novice	"None of the fractions in the square are the actual fraction that it needs to be, but in the mind of the child, I understand how student B got most of the answers." (PST 415)
Emerging	"It is clear that they (Student B) had difficulty understanding larger fractions and reasoning for dividing up a whole piece into many smaller and uneven pieces." (PST 428)
Advanced	"I noticed one main recurring issue, which was them thinking of it as four individual squares rather than one as a whole. For example, in part one, they incorrectly stated that A and B were each equal to one half. They are correct that A and B are one half of one fourth, however if Student B was looking at the big picture, letters A and B would be equal to one eighth. Those two eighths would be equal to one fourth of the entire square. As you can see, Student B solved the square in four different individual parts. Each fourth of the entire square was equal to one, when in reality the entire square should have been equal to one." (PST 421)

Discussion and Implications

The Talanquer Framework (Talanquer et al., 2015) adapted from science education allowed us to capture the subtle differences in PSTs' assessment skills. By using the Talanquer Framework (Talanquer et al., 2015), we were able to go beyond a broad perspective of the assessment skills that the Jacobs et al's Teacher Noticing framework affords (2010). While the Attend and Interpret categories of students' mathematical thinking give us a way to examine PSTs' skills in those areas (Namakshi et al., 2022), the Talanquer Framework (Talanquer et al., 2015), adapted from science education provides an even more detailed description within these two categories. For example, based on our findings the use of evidence and the specificity of evaluation are areas of Attend category that can be further developed to strengthen PSTs' assessment skills. Similarly in the Interpretive dimensions, the PSTs in our study demonstrated accuracy in their mathematics and quality of their interpretation of the student work, while the scope of the evaluation that captures the student thinking in a broader sense of concepts in the context of the overall task was less developed. Mathematics teacher educators can gain insight into the dimensions that PSTs

demonstrate in levels of assessment capabilities and design curriculum to support their skills in a more targeted and intentional way.

Additionally, we provide one example of how a framework can be interdisciplinary in nature as we adapted the Talanquer Framework used in science education for use in mathematics education to gain insight into how PSTs can strengthen their assessment of student thinking.

In our ongoing and future work, we plan to investigate any connections between PSTs' assessment reasoning skills and their Mathematical Knowledge for Teaching in the domain of fractions. We also plan to later expand this project into other math domains.

References

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*, 5–31.

Boerst, T. A., Blunk, M., Shaughnessy, M., Farmer, S. O., DeFino, R., Pfaff, E., & Pynes, D. (2019). Preparing teachers to formatively assess: Connecting the initial capabilities of preservice teachers with visions of teaching practice. In C. Martin, D. Polly, & R. Lambert (Eds.), *Handbook of research on formative assessment in pre-K through elementary classrooms* (pp. 89–116). IGI Global.

Copur-Genturk, Y. (2021). Teachers' conceptual understanding of fraction operations: Results from a national sample of elementary school teachers. *Educational Studies in Mathematics, 107*(3), 525–545.

Jacobs, V. R., Lamb, L. L. C., & Philipp, R. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education, 41*, 168–202.

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). Sage Publications, Inc.

Lee, M. Y., & Lim, W. (2020). Investigating patterns of pre-service teachers' written feedback on procedure-based mathematics assessment items. *International Electronic Journal of Mathematics Education, 15*(1), Article em0561.

Namakshi, N., Warshauer, H. K., Strickland, S., & McMahon, L. (2022). Investigating preservice teachers' assessment skills: Relating aspects of teacher noticing and content knowledge for assessing student thinking in written work. *School Science and Mathematics, 122*(3), 142–154.

Perry, C. (2023). Elementary preservice mathematics teachers fraction knowledge: An integrative review of research. *Educational Considerations, 49*(1), 1–14.

Philipp, R. A., Ambrose, R., Lamb, L. L. C., Sowder, J. T., Schappelle, B. P., Sowder, L., Thanheiser, E., & Chauvot, J. (2007). Effects of early field experiences on the mathematical content knowledge and beliefs of prospective elementary school teachers: An experimental study. *Journal for Research in Mathematics Education, 38*(5), 438–476.

Schoenfeld, A. (1999). Balanced assessment for the mathematics curriculum: Middle grades assessment (Vol. 1). Dale Seymour.

Sleep, L. (2012). The work of steering instruction toward the mathematical point: A decomposition of teaching practice. *American Educational Research Journal, 49*(5), 935–970.

Talanquer, V., Bolger, M., & Tomanek, D. (2015). Exploring prospective teachers' assessment practices: Noticing and interpreting student understanding in the assessment of written work. *Journal of Research in Science Teaching, 52*(5), 585–609.

Thanheiser, E. (2015). Developing prospective teachers' conceptions of the place value structure of the base-ten number system. *Journal of Mathematics Teacher Education, 18*(2), 105–128.

Thurtell, E., Forrester, T., & Chinnappan, M. (2019). Building conceptual knowledge of fraction operations among pre-service teachers: Effect of a representation-based teaching approach within a teacher education program. *Mathematics Teacher Education and Development, 21*(1), 100–124.

Wilson, P. H., Mojica, G., & Confrey, J. (2013). Learning trajectories in teacher education: Supporting teachers' understandings of students' mathematical thinking. *The Journal of Mathematical Behavior, 32*, 103–123.

CATEGORIZING NON-ROUTINE ALGEBRA DIFFICULTIES TO SUPPORT STEM LEARNERS

Eunmi Joung
Utah Valley University
ejoung@uvu.edu

Kuo-Liang Chang
Utah Valley University
kchang@uvu.edu

Abstract

Non-routine problem solving supports the mathematical reasoning STEM students develop, yet little research has systematically identified the specific difficulties students face in non-routine algebra. This qualitative case study examined how 32 undergraduates in Intermediate Algebra, a gateway course for STEM majors, expressed difficulties when solving problems with unfamiliar contexts but familiar solution methods. Using a modified APOS framework, their writing was categorized into Process, Object, and Schema stages. Students managed single objects such as radicals, fractions, and exponents but struggled most when objects combined into a composite object. The study offers a framework and instructional implications for supporting STEM learners.

Keywords: non-routine problems, problem-solving difficulties, APOS theory, undergraduate algebra, STEM learners

Introduction

Higher order thinking has been promoted in college mathematics education over the past decade to improve students' mathematical reasoning and performance (AMATYC, 2018; Freeman et al., 2014; MAA, 2015). Related activities include giving students unfamiliar problems, discovery-based or thought-provoking questions to work on through investigation, discussion, or group problem solving that promotes higher-order thinking. Non-routine problems play a central role in this effort because they provide unfamiliar problem contexts or do not have predetermined solution methods, which encourages students to devise original and/or flexible solution paths (Beghetto, 2017; Chang, 2023; Pólya, 1966; Schoenfeld, 1992). A subject who can successfully solve a non-routine problem exhibits the ability to overcome fixed or rigid thinking (e.g., algorithmic reasoning for solving routine problems), to break from established mindsets, and to produce divergent solutions such as multiple, alternative, and valid solutions (Haylock, 1997; Lithner, 2008). However, little is known about the challenges students face when solving non-routine problems. Therefore, this study investigates the difficulties students encounter while solving such problems.

Objectives of the Study

This study analyzes and categorizes undergraduate students' difficulties when solving non-routine algebra problems in which the context is unfamiliar, but the solution method is one they have recently been taught and practiced. Identifying and categorizing these difficulties can help educators recognize the specific challenges students face and provide more targeted support. Two research questions guided the study:

1. What difficulties do students encounter, with respect to APOS theory, when solving algebra problems with unfamiliar contexts?
2. What patterns emerge in the expressed difficulties across non-routine algebra problems?

Theoretical Framework

Non-Routine Problems: Same Solution Method with an Unfamiliar Context

A routine problem is one in which the solver is familiar with both the context and the solution method (e.g., prior experiences in textbooks or assignments), and a non-routine problem is one in which the solver is unfamiliar with the context, the solution method or both (Chang, 2023; Schoenfeld, 1992). This study focuses on problems whose contexts are unfamiliar to students (e.g., without much familiar contextual information or clues about how to solve them), but whose solution method students have already learned. This focus is informed by the instructors' classroom experience that students may draw a blank when they encounter a problem they have not seen before, even when the unfamiliar feature is a different number or a change in sign.

Brown and Walter (1990) propose the *What-If-Not* method for posing a problem by modifying existing features of a problem. For instance, one might construct a problem in which students are familiar with the solution method but unfamiliar with the context. The method involves four steps: (a) Attribute Listing, (b) What-If-Not-ing, (c) Question Asking, and (d) Analyzing a Problem. As an example, consider a routine textbook problem: "Calculate the area of a rectangle given that the width is 2 meters and the length is 3 meters" (p. 69). To alter this problem, one first lists its attributes and then applies What-If-Not to them. One attribute is that the width and length are specified. Applying What-If-Not, one might ask: "What if only the length is specified?" or "What if the sum of the length and width is given?" Another way to modify a routine problem is to embed it in a richer context such that the to-be-practiced skills for the routine problem are contained in a task with a bigger goal (Boaler, 2015). For example, Boaler (2015) describes a "Pepperoni Pizza" task in which a student rolls a 4 and draws four pizzas, then rolls a 3 and places three pepperonis on

each pizza. The student then uses multiplication, 4×3=12, to determine that there are 12 pepperonis in total. Thus, multiplication practice is embedded within the larger game-like task.

APOS Theory

To understand the difficulties that students encounter when solving a non-routine problem, Action, Process, Object, and Schema (APOS) theory provides a systematic framework for analyzing learners' problem-solving challenges through mental constructions (Arnon et al., 2014). These four stages represent developmental levels in mathematics learning (Arnon et al., 2014; Asiala et al., 1996; Tsafe, 2024). At the Action level, learners depend on external guidance and follow step-by-step procedures, performing operations only through specific instructions or formulas. A student at this level might factor $x^2 - 5x + 6$ only by following instructions and checking each step. The Process level emerges when students internalize these external instructions through repetition and reflection, allowing them to manipulate concepts mentally without external prompts (Clark et al., 1997). Such a student can carry out the factoring mentally and anticipate the result without writing every step. At the Object level, students treat processes as complete entities that can be transformed. A student at this level can treat a quadratic expression as a single object, for instance recognizing $x^4 - 5x^2 + 6$ as quadratic in form. At the Schema level, students connect actions, processes, and objects into a coherent structure applicable to new situations. For example, faced with the unfamiliar equation above, a student may recognize its quadratic-in-form structure and coordinate factoring and the zero-product property to reach a solution.

Students' written descriptions were analyzed as evidence of reasoning associated with the Process, Object, and Schema categories. Mathematical writing has been identified as an essential way of communicating mathematical ideas and performing mathematical reasoning (NCTM, 2000). It provides opportunities for students to explore ideas, explain their reasoning, construct arguments, and elaborate on mathematical ideas (Casa et al., 2016).

At the Process level, students typically describe procedural obstacles or an inability to recall steps. At the Object level, students describe difficulty when two or more objects combine, such as not knowing how to handle a single object built from different kinds of objects. At the Schema level, students describe difficulties in coordinating processes, objects, and methods, and adapting their knowledge to unfamiliar situations (Asiala et al., 1996).

Methodology

A qualitative descriptive case study explored the difficulties students encounter when solving non-routine algebra problems. Students completed two online discussion tasks, each built around a single non-routine problem, solving it and describing in writing the immediate difficulties they met.

Participants

The study involved 32 students enrolled in two sections of Intermediate Algebra during the spring semester of 2024. Participants were mostly first- and second-year undergraduates who had completed the prerequisite courses or placed into the course, drawn from the two authors' classes. Intermediate Algebra marks a critical transition where students build on basic algebraic knowledge while meeting more complex concepts, and it serves as a gateway course for many STEM majors, making the students well suited for a study of difficulties with non-routine problems.

Tasks and Procedure

Students completed two discussion tasks on the university's learning management system, with Discussion 1 assigned during Week 9 and Discussion 2 assigned during Week 11 and provided electronic consent. Students had one week to complete each task. Using the What-If-Not method, two tasks were created by modifying the problem contexts of two routine algebra problems that students had learned and practiced in the classroom and the textbook. Specifically, the coefficients of equations were changed from whole numbers to radicals or fractions. The exponents were changed to higher-order exponents.

The first task (Discussion 1) presented the equation $\sqrt{2}x^4 - 8\sqrt{2}\,x^2 + 16\sqrt{2} = 0$ after students had learned to solve polynomial equations by factoring. Students first read a short teacher-student dialogue about solving a quadratic equation by a related method. Then they were asked to work on the problem and express difficulties they encountered. The second task (Discussion 2) asked students to find the vertex of the function $f(x) = 2x(x - 2)/3 + 2$ and express difficulties encountered after learning factoring, the line of symmetry, completing the square, and a formula for finding the vertex of a quadratic function. The problem is non-routine because its context was unfamiliar, featuring a fractional form containing an embedded polynomial expression.

Data Analysis

A modified APOS framework guided the analysis. The Action level was excluded from the coding framework because the research design did not examine students following step-by-step

instructions from a teacher or textbook. The analysis therefore focused on the Process, Object, and Schema levels. At the Process level, difficulties involve a single mathematical object and fall into two types. The first type is Single Object with difficulties in Definition, for example, difficulties in understanding or interpreting what a single object is. The second type is Single Object with difficulties in Operations, for example, difficulties in carrying out operations on a single object. At the Object level, difficulties involve a composite object formed from different types of objects, such as an equation with radical coefficients or a fraction with a polynomial numerator. This again falls into two types. The first type is Multiple Types of objects without operations mentioned, for example, difficulties in interpreting or recognizing composite objects. The second type is Multiple types of objects with operations mentioned, for example, difficulties attempting to operate on composite objects. At the Schema level, difficulties involve coordinating processes, objects, and methods toward a goal. The coding rubric included two Schema subcategories: Operations with a Goal, in which the student's actions indicated a goal, and Transforming to Final Goal. For reporting purposes, these two subcategories were combined under a single Schema category in Tables 1 and 2. Together with the Process- and Object-level criteria described above, these Schema criteria formed the coding rubric. Both authors independently coded all student responses using this rubric, resolved discrepancies through discussion until reaching full agreement, and documented agreed-upon principles for later coding. For example, the two coders initially disagreed on Student L20's response, "The biggest difficulty for me in this problem is that there are variables in the square root." One coder classified the response as Process—Single Object, whereas the other classified it as Object—Multiple Type Objects. Using the coding rubric, the coders agreed that the response involved more than one type of mathematical object (e.g., a variable and a radical). Because the student did not describe performing an operation, the response was ultimately coded as Object—Multiple Type Objects (No Operations Mentioned).

Results and Discussion

Difficulties in Discussion 1

In Discussion 1, the highest percentages of difficulty occurred at the Process stage of operations on single objects (42.86%) and at the Object stage of composite objects with no operation mentioned (42.86%). These accounted for 85.71% of reported difficulties (see Table 1). The single-object operational difficulties involved three objects: the square root of 2, the variable terms x^2 and x^4, and the whole-number coefficients -8 and 16. For the square root of 2, several

students described not knowing how to handle the radicals. Student L03 wrote, "The first difficulty I run into is the square roots, I'm not sure how to combine them." Student L04 wrote, "At first glance the part that confuses me would be how to get rid [of] two square roots." Student L21 wrote, "I am not very good with square roots, and I am not sure how to get rid of them." For the exponents, Student L01 wrote, "The part of the equation that would puzzle me most is getting rid of the exponents, as they aren't just squares, the first x has an exponent of 4," and Student L10 wrote, "The hardest part of this equation for me would be getting rid of the exponent 4." For the coefficients, Student L22 wrote, "My first question would be about the -8 and 16, what do I do with them?"

The Object-stage difficulties centered on a single composite object, a polynomial in x with irrational, radical coefficients, which students had to treat as one combined structure. Student E02 wrote, "The problem would be difficult because there is a combination of square roots as well as integers raised to the power of another number." Student E08 wrote, "The main part of the equation that puzzles me is the root over the $2x$ to the 4th power. These things tend to confuse me because I feel like we are mixing opposites." Student L20 wrote, "The biggest difficulty for me in this problem is that there are variables in the square root."

A few students were able to do higher order thinking by noticing the composite object and tried to coordinate the radicals and the variable terms. Student L15 wrote, "I can see that all of the variables in the equation have a like term of the square root of 2. I just am unsure of the process of getting rid of those so that the equation could become easier to look at." This response reflects Schema-stage reasoning because the student recognized a possible way to transform the equation, even though the student did not carry the plan through to a complete solution.

Table 1

Number and Percentage of Difficulties in Discussion 1 by APOS Stage

	Process		Object		Schema	Total
	Single object, definition	Single object, operation	Multiple types, no operation	Multiple types, with operation		
Number (Percent)	2 (7.14%)	12 (42.86%)	12 (42.86%)	1 (3.57%)	1 (3.57%)	28 (100%)

Difficulties in Discussion 2

In Discussion 2, the highest percentage again occurred at the Object stage, with composite objects where no operation was attempted (54.17%). The next highest was split between operations on single objects at the Process stage (16.67%) and the Schema stage (16.67%). Difficulties that involve higher order thinking, which combine the Object stage with operation (8.33%) and the Schema stage (16.67%), accounted for 25.00% (see Table 2). The Object-stage difficulties concerned the quadratic function in its non-standard, fractional form, either the fraction with a polynomial in the numerator or the equation not being in standard form.

For the fraction with a polynomial numerator, Student L08 wrote, "the thing that seems the most difficult to me is the numerator." Student L13 wrote, "The immediate issue would be the fraction, as it has so many parts on top." Student L01 wrote, "The immediate difficulties I notice are the items in the parenthesis and the 3 under the fraction." For the non-standard form, Student L02 wrote, "the equation isn't in standard form." Student L05 wrote, "not having the function in quadratic form, and also that part of the function is being divided by 3." Student L19 wrote, "The initial difficulty is the equation is not in the proper form." A few students were able to do higher order thinking. They tried to operate on the fraction or to coordinate the steps needed to reach standard form, though they still encountered difficulty. Student L18 wrote, "I wouldn't know if I need [to] divide the whole thing or to multiply it first." Student L14 wrote, "the function is neither in vertex [n]or general form, so the primary difficulty we face is trying to get h and k, or a and b." A smaller number of students had Process-stage difficulties with operations on the polynomial in the numerator. Student L17 wrote, "the difficulties for me is knowing if I can distribute the parenthesis."

Table 2

Number and Percentage of Difficulties in Discussion 2 by APOS Stage

	Process		Object		Schema	Total
	Single object, definition	Single object, operation	Multiple types, no operation	Multiple types, with operation		
Number (Percent)	1 (4.17%)	4 (16.67%)	13 (54.17%)	2 (8.33%)	4 (16.67%)	24 (100%)

Across both discussions, the difficulty students expressed most often was not recognizing or manipulating a single mathematical object but coordinating objects once they appeared together

within a composite object in an unfamiliar context. In these tasks, students were unable to treat and analyse the composite objects based on their mathematical properties; instead, they were puzzled by the appearance of the combined objects. One possible interpretation is that these students may have been constrained by fixed assumptions about what a routine problem normally looks like and how it should be approached, although these assumptions were not directly examined in this study. To deal with the constraint, Sriraman (2005) and Haylock (1997) highlight the importance of moving beyond familiar methods and finding new ways of thinking about a problem. Likewise, considering unfamiliar viewpoints (Ennis, 2011) and recognizing or questioning hidden assumptions (English, 2023; Facione, 2015) both depend on this same capacity to move beyond isolated objects. For example, in Discussion 1, students may assume that radical coefficients or exponents greater than 2 prevent them from solving the equation. A teacher could ask what method they would use if the coefficients were whole numbers or the highest exponent were 2. After students identify factoring, the teacher could prompt them to apply the same method to the unfamiliar equation: "Can you factor out the common factor and then factor the remaining expression as a quadratic in x^2?"

Students who tried to operate on a composite object, or to coordinate objects, processes, and methods, did more than identify individual elements: they built networks of connected objects, which Tall (2008) treats as a hallmark of advanced mathematical thinking. This aligns with the upper levels of the revised Bloom's taxonomy, where analyzing and creating sit above remembering and applying (Anderson & Krathwohl, 2001).

Implications

The findings showed that students experienced particular difficulty when familiar mathematical objects were combined within unfamiliar composite objects and when they had to coordinate objects, processes, and solution methods. The categories may also help instructors identify whether students need support with an individual mathematical object or operation, interpreting a composite object, or coordinating objects, processes, and solution methods across a problem. The following instructional sequence is proposed based on these findings; it was not tested as an instructional intervention in the present study. Teachers may emphasize familiar methods, such as factoring in Discussion 1, while helping students look beyond unfamiliar surface features to recognize the underlying structure. For example, a teacher may start with a routine equation such as $x^2 - 5x + 6 = 0$. By changing the coefficients and exponents, the equation $\sqrt{3}x^4 - 5\sqrt{3}\,x^2 + 6\sqrt{3} = 0$ is still quadratic in form but looks less familiar. The method students already know still

applies. The teacher may help students recognize the common factor $\sqrt{3}$ and the quadratic form, then ask how they would factor out the common factor and solve the remaining equation using a familiar method.

This study provides mathematics educators with a framework for identifying whether students' difficulties involve individual objects, composite objects, or coordinating multiple ideas and methods. Because students struggled most with unfamiliar composite objects, these categories can help instructors identify the source of difficulty and provide targeted support. Similar difficulties may occur in K–12 algebra, although this was not examined in the present study, so the framework may also help K–12 teachers identify where students need support.

References

American Mathematical Association of Two-Year Colleges. (2018). *IMPACT: Improving mathematical prowess and college teaching*. https://www.amatyc.org/mpage/IMPACT

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Arnon, I., Cottrill, J., Dubinsky, E., Oktaç, A., Roa Fuentes, S., Trigueros, M., & Weller, K. (2014*). APOS theory: A framework for research and curriculum development in mathematics education*. Springer.

Asiala, M., Brown, A., DeVries, D. J., Dubinsky, E., Mathews, D., & Thomas, K. (1996). A framework for research and development in undergraduate mathematics education. *Research in Collegiate Mathematics Education, 2*, 1–32.

Beghetto, R. A. (2017). Lesson unplanning: Toward transforming routine tasks into non-routine problems. *ZDM Mathematics Education, 49*(7), 987–993. https://doi.org/10.1007/s11858-017-0885-1

Boaler, J. (2015). *Fluency without fear: Research evidence on the best ways to learn math facts*. YouCubed at Stanford University. https://www.youcubed.org/evidence/fluency-without-fear/

Brown, S. I., & Walter, M. I. (1990). *The art of problem posing* (2nd ed.). Lawrence Erlbaum Associates.

Casa, T. M., Firmender, J. M., Cahill, J., Cardetti, F., Choppin, J. M., Cohen, J., Cole, S., Colonnese, M. W., Copley, J., DiCicco, M., Dieckmann, J., Dorl, J., Gavin, M. K., Hebert, M. A., Karp, K. S., LaBella, E., Moschkovich, J. N., Moylan, K., Olinghouse, N. G., . . . Zawodniak, R. (2016). *Types of and purposes for elementary mathematical writing: Task force recommendations*. University of Connecticut. https://mathwriting.education.uconn.edu/wp-content/uploads/sites/1454/2021/09/Elementary-Mathematical-Writing-White-Paper.pdf

Chang, K. (2023). Developing flexibility and originality in mathematical problem solving. *MathAMATYC Educator*, *14*(2), 45–51.

Clark, J. M., Cordero, F., Cottrill, J., Czarnocha, B., DeVries, D. J., St. John, D., Tolias, G., & Vidakovic, D. (1997). Constructing a schema: The case of the chain rule. *The Journal of Mathematical Behavior, 16*(4), 345–364. https://doi.org/10.1016/S0732-3123(97)90012-2

English, L. D. (2023). Ways of thinking in STEM-based problem solving. *ZDM Mathematics Education, 55*, 1219–1230. https://doi.org/10.1007/s11858-023-01474-7

Ennis, R. H. (2011). *The nature of critical thinking: An outline of critical thinking dispositions and abilities*. University of Illinois.

Facione, P. A. (2015). *Critical thinking: What it is and why it counts*. Insight Assessment.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Haylock, D. (1997). Recognising mathematical creativity in schoolchildren. *ZDM, 29*(3), 68–74. https://doi.org/10.1007/s11858-997-0002-y

Lithner, J. (2008). A research framework for creative and imitative reasoning. *Educational Studies in Mathematics, 67*(3), 255–276. https://doi.org/10.1007/s10649-007-9104-2

Mathematical Association of America. (2015). *2015 CUPM curriculum guide to majors in the mathematical sciences*. https://maa.org/resource/cupm-guide/

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*.

Pólya, G. (1966). On teaching problem solving. In E. G. Begle (Ed.), *The role of axiomatics and problem solving in mathematics* (pp. 123–129). Ginn and Company.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334–370). Macmillan.

Sriraman, B. (2005). Are giftedness and creativity synonyms in mathematics? *The Journal of Secondary Gifted Education, 17*(1), 20–36.

Tall, D. (2008). The transition to formal thinking in mathematics. *Mathematics Education Research Journal, 20*(2), 5–24. https://doi.org/10.1007/BF03217474

Tsafe, A. K. (2024). Effective mathematics learning through APOS theory by dint of cognitive abilities. *Journal of Mathematics and Science Teacher, 4*(2), em058. https://doi.org/10.29333/mathsciteacher/14308

ETHNOMODELING: A VEHICLE FOR CATALYZING CHANGE IN MATHEMATICS TEACHER PREPARATION

Siddhi Soni
Eastern Connecticut State University
siddhi.mathed@gmail.com

Alyssa Harbin
University of Central Florida
alyssa.harbin@ucf.edu

Trena L. Wilkerson
Baylor University
trena_wilkerson@baylor.edu

Abstract

Many have called for critical conversations in teacher preparation programs to value contributions from diverse backgrounds, integrate opportunities to use math for liberation, and connect math and lived experiences of marginalized and minoritized students. Conversations and awareness of the structures that inhibit opportunities for learning in schools and systems are critical for PTs to engage in, which includes envisioning alternative structures that can support equitable learning structures, such as through the engagement of ethnomodeling tasks. In this manuscript we share student examples and data to illustrate how an ethnomodeling approach can be a vehicle for catalyzing change in teacher preparation.

Keywords: ethnomodeling, teacher preparation, elementary, secondary, ethnomathematics

Introduction

In culturally relevant situations within our mathematics curriculum, we often start with a "notice/wonder" task. While these conversations are a start, such opening approaches are not sufficient to authentically attend to and value teachers, students, and our community's diverse cultures and identities. Instead of settling for "surface" mathematical tasks that fail to adequately and authentically delve deeply into human experiences, in this paper we will discuss how the integration of the ethnomodeling framework (Desai et al., 2022; Orey & Rosa, 2010) afforded us the opportunity to delve deeper into both attending to students' cultures and identities while also maintaining a focus on rigorous mathematical content during our work with prospective teachers (PTs). According to Rosa and Orey (2010) ethnomodeling is "a process of elaboration of problems and questions growing from real situations that form an image or sense of an idealized version of the *mathema"* (p. 18). Rosa and Orey (2013) go on to explain ethnomodeling as "the study of

mathematical ideas and procedures elaborated by members of distinct cultural groups" (p. 65). Through such intentional integrations of the ethnomodeling framework in teacher preparation programs, teacher educators can advocate for and model the teaching and learning of mathematics in ways that focus on ethnomathematical and cultural perspectives to reflect, sustain, and extend mathematical situations that require rigorous analysis and attention to structures. In this paper, we share results from two studies - one in an elementary math education PT course and another in a secondary math education PT course to exemplify how through an ethnomodeling approach we can start to have these critical conversations. Using a lens of the four recommendations from *Catalyzing Change* (National Council of Teachers of Mathematics [NCTM], 2018, 2020a, 2020b), we discuss experiences of these PTs in engaging in ethnomodeling tasks.

Objectives of the Study

The purpose of this study was to understand PTs conceptualization of broadening the purposes of learning mathematics, creating equitable structures in mathematics teaching and learning, implementing equitable mathematics instruction, and developing a deep mathematical understanding through engaging in an ethnomodeling approach to learning and teaching mathematics.

Related Literature

Critical conversations to initiate changes across our PK–12 curriculum and in teacher preparation programs, particularly in connection to valuing contributions from mathematicians of color and from diverse backgrounds, integrating opportunities to use mathematics for liberation, and connections between mathematics and the lived experiences of historically marginalized and minoritized students have been advocated by both NCTM (2018, 2020a, 2020b) and the Association of Mathematics Teacher Educators (AMTE, 2017). In the *Catalyzing Change* series, NCTM (2018, 2020a, 2020b) advocates for four key recommendations: (1) broaden the purposes of learning mathematics, (2) create equitable structures in mathematics, (3) implement equitable mathematics instruction, and (4) develop deep mathematical understanding. Similarly, in *Standards for Preparing Teachers of Mathematics*, AMTE (2017) supports that a deep, integrated focus on equity in every program that prepares teachers of mathematics is necessary to ensure the success of teachers in helping to foster classrooms, schools, and communities that are equitable and inclusive for all students. Indicator C.4.4 in AMTE (2017) advocates for "beginning teachers of mathematics [to] understand the roles of power, privilege, and oppression in the history of mathematics education

and are equipped to question existing educational systems that produce inequitable learning experiences and outcomes for students." We advocate that such equitable learning structures and instructional practices could be accomplished through engagement in ethnomodeling tasks.

An ethnomodeling approach bridges academic mathematical ideas and practices and cultural aspects that have continued to become eroded or eliminated due to the siloing of mathematical concepts. A key objective of using ethnomodeling involves moving away from traditional perspectives of doing, learning, and teaching mathematics towards learning and gaining perspectives on local practices to understand how mathematical ideas are modeled and organized from the point of view of the people engaging in these practices. The goal is to engage and empower people to further understand, analyze, and mathematize their own realities, such as in the tasks described in this paper.

Within teacher preparation programs, conversations and awareness of the structures that inhibit opportunities for learning in schools and systems are critical for PTs to engage in, which includes envisioning alternative structures that can support equitable learning structures such as through the engagement of ethnomodeling tasks. The integration of ethnomodeling in teacher preparation programs offers a powerful way to value, respect PTs' and MTEs' experiences and backgrounds and gives them tools to become critical participants in our communities. Analyzing elementary and secondary PTs experiences in ethnomodeling tasks provided us a lens into their conceptualization of the *Catalyzing Change* (NCTM 2018, 2020a, 2020b) recommendations.

Methodology

A qualitative narrative methodology was used as the focus of the analysis was on the interpretations and experiences of participants. These research studies took place at two different universities in mathematics education courses. The elementary study took place in a school of education at a small private metropolitan university in the northeastern U.S. during the Fall 2023-Spring 2024 semesters. The secondary study took place in a college of education at a large Hispanic-serving public metropolitan research university in the southeastern U.S. during the Spring 2022 semester as a part of a dissertation study. A combination of observations, interviews, pre-/post-surveys, written artifacts, and audio/video recordings allowed for the use of different data sources to validate and triangulate findings.

Following each task, PTs completed reflection questions about their ethnomodeling experience. The first and second authors analyzed these responses through repeated reading,

aligning and categorizing them into one of the four *Catalyzing Change* recommendations based on NCTM descriptors (2018, 2020a, 2020b). The authors met to resolve any categorization difficulties through collaborative discussion.

Results and Discussion

While the study analyzed both elementary and secondary PTs conceptions for each of the four recommendations from Catalyzing Change, this paper highlights one grade band per recommendation. Future papers will discuss more detailed analysis.

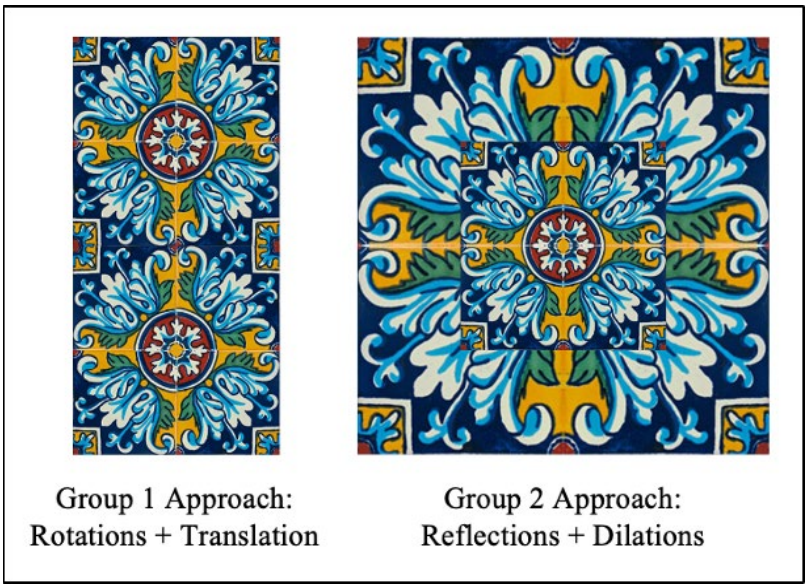
Broaden the Purposes of Learning Mathematics

Recommendation 1: Broaden the purposes of learning mathematics, emphasizes that "Each and every student should learn the Essential Concepts in order to expand professional opportunities, understand and critique the world, and experience the joy, wonder, and beauty of mathematics" (NCTM, 2018, p. 9). The mathematics used by mathematicians is seen as a creative, visual, connected endeavor (Boaler, 2016; Su, 2021), and such opportunities that value and connect the learning of academic mathematics to classroom experiences are of paramount importance for each and every student to access relevant, equitable, and high-quality mathematics education (Aguirre et al., 2013). Through ethnomodeling tasks, PTs engaged in culturally responsive experiences that connected to their backgrounds, cultures, and identities while simultaneously integrating rich mathematical content. In each of the tasks the secondary PTs engaged in, they were able to broaden their perspectives of what it means to do and learn mathematics.

For instance, through exploring tiles and tiling patterns around the world PTs were able to notice geometric and mathematical aspects such as shapes, lines, and symmetry, and their noticings and wonderings also extended to cultural, historical, and design aspects. All too often, traditional high school geometry instruction strictly adheres to the study of two- and three-dimensional figures, postulates, and measurements (NCTM, 2018), which results in students feeling a disconnect between themselves and the relevance of mathematics (Desai & Safi, 2023). Through explorations with Mexican Talavera tiles (see Figure 1), secondary PTs engaged in interesting conversations about mathematical content and context.

Figure 1

Talavera Tile Explorations



While at first glance, there may not appear to be a noticeable difference between the two approaches, upon further examination, a design difference in the leaf pattern can be seen. The distinctions in the designs led to an important discussion about using different transformations, their impacts on orientation, if/when it matters, and how these ideas connect to real-world contexts such as the planning, design, and creation process regularly contemplated by artisans. The intentional connections and integrations of ethnomodeling and geometry afforded PTs' opportunities to link mathematical domains enriched by the wonder, joy, and beauty of geometry in our lives, communities, and cultures.

Create Equitable Structures in Mathematics

Recommendation 2 advocates for mathematics instruction to dismantle inequitable structures, including tracking teachers as well as the practice of ability grouping and tracking students into qualitatively different courses (NCTM, 2020b). Boaler (2011) and Larnell (2016) describe that current school structures are often inequitable for students of color, particularly African American, Latinx, and children living in poverty. AMTE (2017) and the National Council of Supervisors of Mathematics & TODOS: Mathematics for ALL (NCSM & TODOS, 2016) both advocate for beginning teachers of mathematics to understand, have conversations about, and be aware of structures that inhibit opportunities for learning in schools and systems. Beyond these

conversations, PTs should also consider alternative structures that promote equitable learning structures such as ethnomodeling tasks.

Secondary PTs highlighted heterogeneous grouping as a key equitable structure. Mathematical learning can occur within and across groups, but structures such as ability grouping and tracking students into varied math courses, common at the secondary level, limit opportunities for students to learn from one another's diverse cultural, linguistic, and lived experiences (NCTM, 2020b). When engaging in the various ethnomodeling tasks, secondary PTs worked in small groups of 4–5 that were each diverse in terms of race, gender, and experiences. The power of such grouping was noticed and mentioned by multiple students. For instance, in one of the groups, a group member was a former architecture major who had a deep knowledge and understanding of the history and culture of many different architectural aspects; a different member of this same group was the only student in the school who had formal teaching experience as a current classroom teacher. Each of them brought unique perspectives to the conversation, and the heterogeneous grouping highlighted the importance of valuing multi-dimensional aspects of students' identities. In their final reflection, a PT shared, "working in heterogeneous groups gave us multiple perspectives to view the tasks from, and as each group member had different strengths, we were able to learn from each other and build something great together." When we track students as well as place them into within-class homogeneous grouping, we rob them of meaningful experiences to learn from the brilliance and life experiences of their peers. Such examples illustrate the importance of ethnomodeling tasks in working towards creating equitable structures and providing students' opportunities to learn about their own and their peers' cultures and identities in ways that are rich in mathematical explorations.

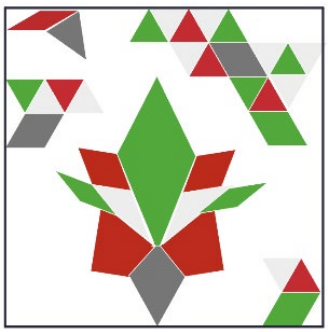
Implement Equitable Mathematics Instruction

The third recommendation related to implementing equitable mathematics instruction states that "Classroom instruction should be consistent with research-informed and equitable teaching practices" (NCTM, 2018, p. 25). Equitable teaching practices can be traced to the eight practices found in *Principles to Actions* (NCTM, 2014). Two of these practices that were clearly seen in one ethnomodeling task, the *American Barn Quilt* (Soni & Harbin, 2026), were to 'use and connect mathematical representations' and 'facilitate meaningful mathematical discourse.' This task consisted of learning about the history of the barn quilts, creating personal barn quilts using geometric shapes that represented their culture and family, and then sharing their designs with their peers. This

allowed elementary PTs to feel a sense of belonging while gaining an appreciation of one another's cultures and backgrounds. One PT noted that, "this task showed me that bringing personal and cultural experiences into learning can make activities more interesting and relatable. I think it's important to include their [students'] identities in the classroom so they feel seen and appreciated."

Figure 2

Two Examples of PT's Barn Quilts



This quilt represents my family and my community as the colors red, white/gray, and green are representative of the national flag of Mexico, where I am from. I attempted to make a sort of flower, geometric design that represents the passing/generosity of traditions from generation to generation, and love.



I am very inspired by the sun. It brings us life, the days we live and gives us energy. I am grateful for every day I live. The six hexagons represent the six most important things in my life: my daughter, my family, my career, my health and fitness, my safety and my future. The trapezoids represent all the traveling routes I have taken and will continue to take - north, south, east, west and in between.

In this ethnomodeling task, PTs designed personal barn quilts, using geometric representations for a deeper purpose than merely describing shapes. These representations fostered meaningful discourse about geometrical concepts alongside the history and utility of barn quilts (see Figure 2). Engaging in these activities and conversations strengthens PTs' mathematical understanding, preparing them to implement equitable instruction in their future classrooms.

Develop Deep Mathematical Understanding

This recommendation focuses on the key concepts important to developing a deep mathematical understanding. The recommendation states that "Each and every student should learn the Essential Concepts in order to expand professional opportunities, understand and critique the world, and experience the joy, wonder, and beauty of mathematics" (NCTM, 2018, p. 37) which ties directly to Recommendation 1. One way to develop deep mathematical understanding is to promote

reasoning and problem solving (NCTM, 2014). For example, in the *Chocolate Around the World Task,* elementary PTs explored the global history and uses of cacao, then analyzed quantitative data to determine the optimal sugar content for a newly invented chocolate bar for their island.

When sharing their justifications for how they arrived at the amount of sugar, PTs shared a variety of mathematical skills that assisted them in reaching a conclusion. This included proportional reasoning, multiplication, bar graphs, and interpreting data in various ways (mean, median, and mode). Sharing and justifying their reasoning with others allowed them to see different ways this task could have been approached or solved. However, their justifications went beyond just mathematical calculations, as one PT shared,

> I come from a background where we usually focus on numbers and getting the "right answer" quickly. So when I started the task, I automatically tried to figure out how to make the cocoa production more efficient. But then I realized that I also had to think about what the islanders valued, like tradition and the environment. That made me slow down and consider different solutions, not just the most profitable one.

This example highlights PTs' mathematizing both their own realities and those of others, and recognizing that an 'optimal' solution can extend beyond mere calculations.

Implications

Consistent with redefining traditional ways of doing and learning mathematics as recommended by NCTM's Catalyzing Change series (2018, 2020a, 2020b), as well as by the work of D'Ambrosio (2001), mathematics needs to be thought of as more than, and indeed different from counting, calculating, sorting, or comparing. Through a lens of the catalyzing change recommendations, we highlight PTs' conceptions of these recommendations in connection to their experiences with ethnomodeling tasks. PTs come into our courses with unique skill sets, backgrounds, and experiences; however, until and unless we create opportunities for them to integrate these into the mathematics they are learning, we are not able to leverage the full potential of the students. By providing these transformative learning experiences, teacher preparation programs can better prepare PTs to become catalysts for more equitable and culturally sustaining mathematics classrooms.

References

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

Association of Mathematics Teacher Educators. (2017). *Standards for Preparing Teachers of Mathematics*. https://amte.net/standards

Boaler, J. (2011). Changing students' lives through the de-tracking of urban mathematics classrooms. *Journal of Urban Mathematics Education*, *4*(1), 7–14.

Boaler, J. (2016). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching*. Jossey-Bass.

D'Ambrosio, U. (2001). What is ethnomathematics, and how can it help children in schools? *Teaching Children Mathematics, 7*(6), 308–310.

Desai, S., & Safi, F. (2023). Model it! Building and sustaining cultural traditions with ethnomodeling. *Mathematics Teacher: Learning and Teaching PK–12, 116*(10), 782–787.

Desai, S., Safi, F., Bush, S. B., Wilkerson, T., Andreasen, J., & Orey, D. C. (2022). Ethnomodeling: Extending mathematical modeling research in teacher education. *Investigations in Mathematics Learnin*g, 14(4), 305–319.

Larnell, G. V. (2016). More than just skill: Examining mathematics identities, racialized narratives, and remediation among Black undergraduates. *Journal for Research in Mathematics Education, 47*(3), 233–269.

National Council of Supervisors of Mathematics & TODOS: Mathematics for ALL. (2016). *Mathematics education through the lens of social justice: Acknowledgement, actions, and accountability* [Joint Position Paper].

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

National Council of Teachers of Mathematics. (2018). *Catalyzing change in high school mathematics: Initiating critical conversations*.

National Council of Teachers of Mathematics. (2020a). *Catalyzing change in early childhood and elementary mathematics: Initiating critical conversations.*

National Council of Teachers of Mathematics. (2020b). *Catalyzing change in middle school mathematics: Initiating critical conversations*.

Orey, D. C., & Rosa, M. (2010). Ethnomodeling: A pedagogical action for uncovering ethnomathematical practices. *Journal of Mathematical Modelling and Application, 1*(3), 58–67.

Rosa, M., & Orey, D. C. (2010). Ethnomodeling as a pedagogical tool for the ethnomathematics program. *Revista Latinoamericana de Etnomatemática*, *3(*2), 14–23.

Rosa, M., & Orey, D. C. (2013). Ethnomodeling as a research theoretical framework on ethnomathematics and mathematical modeling. *Journal of Urban Mathematics Education, 6*(2), 62–80.

Soni, S., & Harbin, A. (2026). Barn Quilts Ethnomodeling Task. *Mathematics Teacher: Learning and Teaching PK–12*, *119*(1), 16–20. https://doi.org/10.5951/MTLT.2025.0012

Su, F. (2021). *Mathematics for human flourishing*. Yale University Press.

Acknowledgements

Portions of the data collected for this task were part of the Mathematics Education Trust & Association of Mathematics Teacher Educators Early Career Research Grant for which the first author was a recipient.

LONGITUDINAL DEVELOPMENT OF ELEMENTARY STUDENTS' WRITTEN MATHEMATICAL ARGUMENTS

Madelyn W. Colonnese
University of North Carolina at Charlotte
mcolonn1@charlotte.edu

Fabiana Cardetti
University of Connecticut
fabiana.cardetti@uconn.edu

Christie Martin
University of South Carolina, Columbia
martinc1@mailbox.sc.edu

Abstract

Students' ability to compose mathematical arguments is supported by opportunities to share their reasoning with others. However, there is limited understanding about how students' written mathematical reasoning matures throughout the elementary years. This study reports on the analysis of elementary students (n = 5) who participated in 3 to 4 consecutive summers of a mathematics program focused on argumentative writing. Students' written reasoning was analyzed using a content analysis approach. Findings revealed that students followed non-linear developmental trajectories and four characteristics that describe the longitudinal development of written mathematical argumentation.

Keywords: mathematical argumentative writing, written reasoning, elementary mathematics

Introduction

Children begin developing social practices for arguing at a very young age. These early forms of argumentation lay the foundation for how students engage in sharing their mathematical reasoning within the classroom (Civil & Hunter, 2015). Classroom opportunities for students to engage in sharing their reasoning with others and hearing the reasoning of their peers supports the development of higher-quality mathematical arguments (Rojas-Drummond & Zapata, 2004). Recognizing the natural trajectory and learning benefits of these discursive practices, national policy frameworks have long positioned student-led reasoning as an essential part of mathematics learning.

Consistent with the view of argumentation as a social practice grounded in mathematical reasoning, we conceptualize mathematical argumentation using Toulmin's framework (1958/2003) in which arguments emerge through social interaction and consist of a claim, evidence, and warrant. We drew on Toulmin's (1958) argumentation framework alongside sociocultural theory to understand how collaborative, classroom-based reasoning shapes students' individual written mathematical writing overtime. Because mathematical argumentation is an expectation across

students' primary and secondary education, it is anticipated that students will progress in their ability to compose a written mathematical argument. This progression is not only mathematical in nature but also interwoven with students' ability to appropriately use the language of mathematics and communicate with increasing sophistication. In spite of its importance, relatively few studies have examined how written mathematical argumentation develops across grade levels, particularly from integrated social/cognitive perspectives (Lehmann & Friend, 2025). Considering the need to attend to written mathematical arguments and understanding how mathematical argumentative writing matures over time, the purpose of this study is to explore how U.S. primary students' written mathematical arguments develop through repeated participation in a classroom community where this practice is routinely expected and cultivated under carefully designed instruction.

Objectives of the Study

Kosko and Zimmerman (2019) identified a set of classifications to understand how students' mathematical writing develops across grades K–3 (see Table 1). The classifications distinguish increasingly explicit ways of communicating mathematical reasoning, from statements and recounts to descriptions and explanations. This study seeks to build on and extend this area of research in two ways: (1) to analyze students' written mathematical arguments after engaging in instruction specifically designed to support mathematical argumentative writing; and (2) to analyze the longitudinal development of individual student's written mathematical arguments. Given these objectives, the research question guiding the study is: What characterizes the longitudinal development of elementary students' written mathematical arguments across years of participation in a summer mathematics program?

Table 1

Kosko and Zimmerman's (2019) Classifications of Mathematical Argumentative Writing

Classification	Brief description
Statement	States a claim or answer with little or no reasoning.
Recount	Describes what the student did; the mathematics remains largely implicit.
Procedure	Makes the mathematical actions or operations explicit, but the justification remains implicit.
Detailing	Integrates information given in the task into the mathematical procedure or argument.
Description	Provides an explicit justification for why the procedure or strategy works in the particular situation.
Explanation	Provides a justification that explicitly invokes a mathematical rule, definition, or property.

Theoretical Framework

The theoretical framework guiding this study is sociocultural theory (Lerman, 2001). Sociocultural theory provides a lens for the research since primary students' mathematical argumentation is situated within and shaped by social and cultural practices across multiple years of a summer program. Consequently, in this study mathematical argumentation is viewed as not only an individual act of reasoning, but as a collective communicative practice within the classroom community. Toulmin's theory of argumentation offers a model for understanding the composition of a mathematical argument –the evidence leads to a conclusion supported by a warrant and backing (1958). Within this study, sociocultural theory explains how norms, social interactions, and classroom culture mediate students' reasoning, while Toulmin's framework serves as our analytical tool to capture how these social, collective practices are internalized by individual students into structured written mathematical arguments over time.

Methodology

Participants

The students (n = 5) involved in the study attended a summer mathematics program for 3 to 4 consecutive summers. Two of the students began the summer academic program after completing first grade. Three of the students began the academic program after completing second grade. The students all attended the same elementary school located in a large southeastern city during the academic year. All of the students at the school were invited to participate each summer in the academic program. Because of this, the participants in this study reflect a range of learners. The project began when students were in first or second grade and as a result participation varied: students who started in second grade attended three summer programs, whereas those who started in first grade attended four summer programs. The school was a Title 1 school and 95% of students qualified for free and reduced lunch. Students eligible to attend the school during the academic year had to meet at least one of the following requirements: at-risk academically or behaviorally, at-risk for or lacking stable housing, or coming from an underperforming school based on state metrics.

Mathematical Argumentative Writing Instruction

Each summer program consisted of 19 days of instruction. The mathematics instruction was designed to reinforce previous grade level content through oral and written mathematical communication. For example, the students who had completed first grade focused on instructional

activities to support their understanding of grouping by tens. The first week of instruction focused on introducing students to mathematical writing with emphasis on learning the components of argumentative writing. During the second week of instruction, students participated in a four-lesson sequence designed to provide students an opportunity to practice writing a mathematical argument and receive feedback from their peers on their argument (see Table 2). The third week of instruction focused on composing a mathematical explanation. At the end of the 19-day unit, students completed a post-assessment task designed to elicit a mathematical argument. The tasks were adapted from state assessments for Grades 1 and 2 and large-scale assessments including NAEP (National Assessment of Educational Progress, 2019) and SBAC (Smarter Balanced Consortium, 2022) for Grade 3 and 4. These post-assessments served as the primary data source for the current study and remained consistent in terms of mathematical content and difficulty across years.

Table 2

Sequence of Mathematical Argumentative Writing Instruction (Week 2)

Lesson	Instructional Activity
Lesson 1	Students engaged in a mathematical game or investigation to explore the mathematical concept.
Lesson 2	Students discussed how representations and strategies used during the mathematical game or investigation supported their understanding of the mathematical concept.
Lesson 3	Students translated ideas developed through the mathematical game or investigation and discussion into written mathematical reasoning.
Lesson 4	Students engaged in peer feedback on their written mathematical reasoning.

Data Analysis

To conduct an analysis of students' progression of written mathematical arguments, we employed a content analysis methodology (Krippendorff, 2018) of students' post-assessments. We used Krippendorff's (2018) definition of content analysis – a "technique for making replicable and valid inferences from texts to the contexts of their use" (p. 24). To operationalize this analysis, we adopted Kosko and Zimmerman's classification scheme for mathematical argumentative writing. Grounded in Toulmin's (1958) theory of argumentation, this framework was selected because it provides a developmental continuum tailored specifically to early elementary (K–3) learners.

First, we deductively applied Kosko and Zimmerman's (2019) six classifications: statement, recount, procedural, detailing, description, and explanation. While coding we also allowed for sub-categories or variations of the existing classifications to emerge from the dataset. Inter-rater reliability across the three raters was evaluated using Fliess' Kappa (κ). The analysis yielded a Kappa

coefficient of $\kappa = 0.74$, indicating substantial agreement according to standard interpretation benchmarks (Landis & Koch, 1977). We then collaboratively conducted an inductive analysis of the students' written arguments to identify characteristics that elaborate on the existing classifications and illustrate the students' longitudinal development. To conduct the inductive analysis, the first two authors separately analyzed each student's writing, generating initial open codes for emergent features. The authors then met to compare notes and refine category definitions until reaching consensus. Characteristics were determined to be distinct when they captured non-overlapping structural or conceptual features of argumentation. The inductive analysis revealed four characteristics of students' written mathematical arguments.

Results and Discussion

Findings from the deductive analysis (see Table 3) revealed that students did not move uniformly through Kosko and Zimmerman's (2019) classifications. Rather, each student followed an individualized developmental trajectory. For example, Student 3 progressed gradually from recount (2022–2023) to procedural (2024), whereas Student 5 demonstrated a more substantial shift from emergent writing in 2022 to description in 2023.

Table 3

Classification of Students' Written Mathematical Arguments

Student	2022	2023	2024	2025
1	Statement	Procedural	Procedural	Detailing
2	Procedural	Detailing	Procedural	Procedural
3	Recount	Recount	Procedural	n/a
4	Procedural	Description	Detailing	n/a
5	Emergent	Description	Description	n/a

These individualized trajectories suggested that movement through the classifications alone did not fully capture students' longitudinal development, motivating the subsequent inductive analysis. This analysis identified four characteristics of the longitudinal development of elementary students' written mathematical argumentation: drawing on mathematical structure, responding to mathematical demands, writing for an audience, and composing individualized arguments.

Drawing on Mathematical Structure

Student 1 illustrates how mathematical thinking became increasingly visible over time. In 2022, when describing why $54 > 59$, Student 1 wrote "Trus [true], greater". Although Student 1

recognized that the symbol > means greater and believed the statement to be true, the mathematical relationship (54 > 59) remained implicit. In 2023, the student made this thinking visible by drawing 54 circles, crossing out 10, and writing, "I made 54 circles and took away 10." The interplay of representation and writing illustrates a developmental shift in how the student drew upon mathematical structure to communicate her reasoning. Rather than simply stating a conclusion, the student increasingly relied on mathematical structure, in this case a representation, to make visible the relationships underlying her claim. Although students drew upon different mathematical structures, their written arguments gradually came to be organized around mathematical relationships rather than isolated statements or procedures. This finding extends Kosko and Zimmerman's (2019) descriptions of written mathematical arguments by illustrating how students used mathematical resources to communicate their reasoning.

Responding to Mathematical Demands

As mathematical content became increasingly sophisticated across grade levels, students faced greater demands in communicating their mathematical reasoning. Student 2 illustrates this characteristic. Between 2022 and 2023, the student's writing shifted from procedural to detailing; however, the 2024 and 2025 arguments focused on fractions. Although Student 2 remained at a procedural level, increasingly sophisticated fraction concepts made the reasoning more difficult to communicate in writing. Nevertheless, the student continued to make their thinking visible. For example, in 2025, the student wrote "⅖ + ⅗ + ⅘ I agree with 9/15". Although mathematically incorrect, the response communicates the student's strategy and connects it to the prompt. Students' sustained efforts to make their mathematical thinking explicit likely reflect the communicative culture of the summer program, which emphasized sharing writing with both peers and authentic external audiences. This suggested that increasingly sophisticated mathematics does not necessarily correspond to increasingly sophisticated written arguments. This finding contributes to the emerging literature on longitudinal development in mathematical argumentative writing (Lehmann & Friend, 2025) by suggesting that changes in students' written arguments should be interpreted alongside the increasing mathematical sophistication of the tasks they are asked to communicate about.

Writing for an Audience

Students demonstrated a distinct shift toward audience awareness. Initially, students' writing appeared to be directed primarily toward themselves. In Student 4's 2022 response (Figure 1), the procedure is outlined without sufficient context or justification for an external reader to understand

why 87 is greater than 86. By Grade 4 (2024, Figure 2), the student explicitly justified the solution by explaining the denominator, making the reasoning accessible to others. This shift illustrates how students transitioned over time from recording internal thoughts to crafting arguments designed for an external reader, reflecting norms that were introduced and cultivated through the summer program. Over time, audience awareness became a defining characteristic of students' written mathematical argumentation. This finding aligns with sociocultural perspectives that position argumentation as a communicative practice developed through participation with others (Civil & Hunter, 2015; Rojas-Drummond & Zapata, 2004), illustrating how attention to audience becomes increasingly evident in students' written mathematical arguments.

Figure 1

Argumentative Writing from 2022

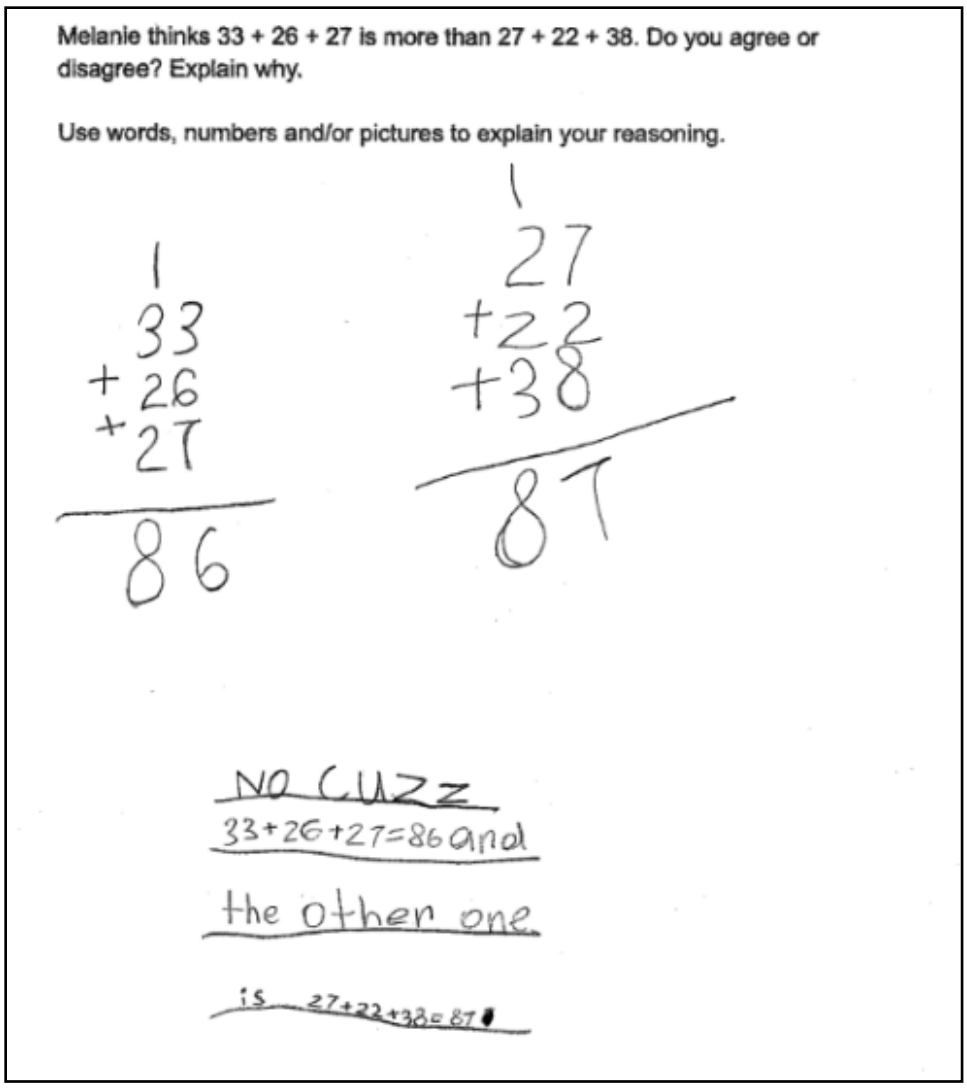
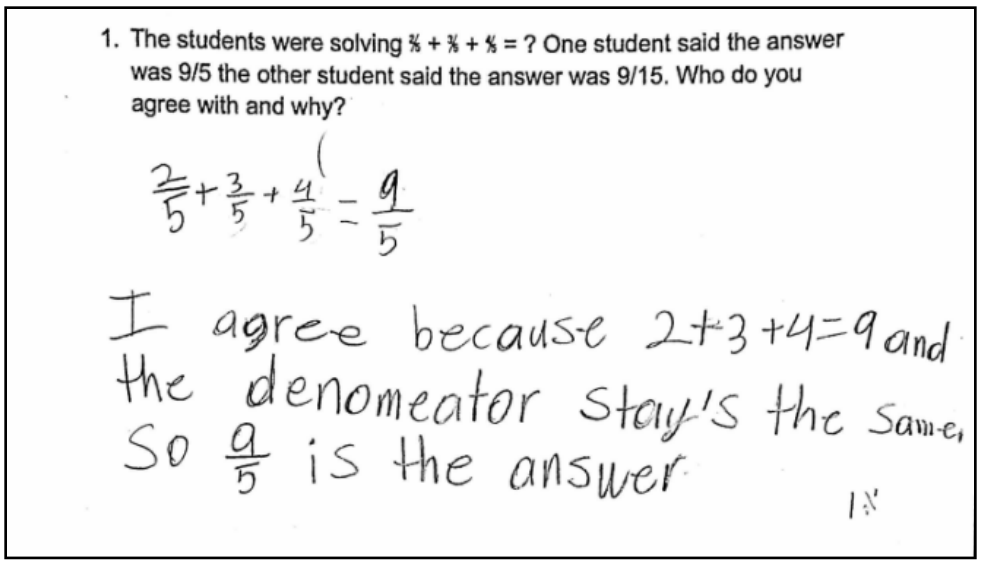


Figure 2

Argumentative Writing from 2024



Composing Individualized Arguments

In analyzing the student samples, we observed that the students' writing was not formulaic; rather, each student expressed their ideas uniquely, reflecting individualized ways of composing mathematical arguments. Responses to a common prompt (Figure 3; Table 4) illustrate this characteristic. Although all responses included claims and supporting evidence, their structures and underlying fraction understandings differed. For instance, some students stated their claim with a simple "No," while others used a formal sentence frame such as "I agree/disagree." Furthermore, students drew upon different conceptual models of fractions—some focused on equal-sized parts, whereas others centered their reasoning on the number of shaded parts. Together, these examples suggest that students appropriated written mathematical argumentation in individualized ways while drawing on different mathematical understandings to support their claims.

Figure 3

Grade 3 Argumentative Writing Prompt

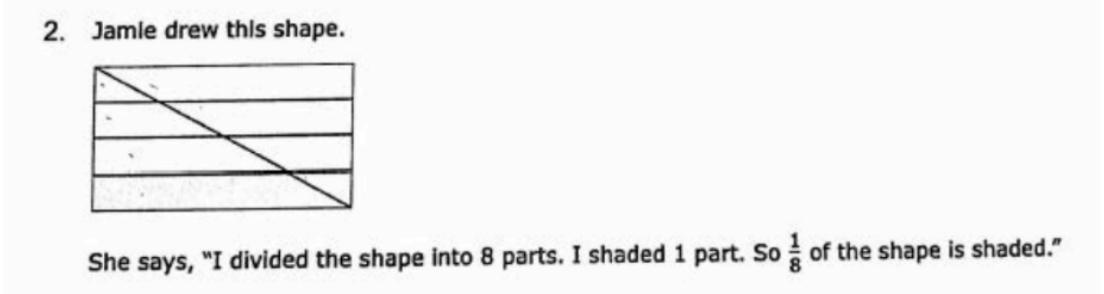


Table 4

Students' Mathematical Arguments to the Prompt in Figure 3

Student	Response
1	"It is incorrect because it show ¼."
2	"I agree 7 not shaded 1 is shaded."
3	"I disagree because it has to many lines in it shade in one."
4	"No because it is not equal."
5	"No because she don't have 8 equivalent parts."

Conclusion

The purpose of this study was to characterize the longitudinal development of elementary students' written mathematical arguments across repeated participation in summer mathematics programs that emphasized written mathematical argumentation. Kosko and Zimmerman's (2019) framework provided a developmental lens for characterizing students' written mathematical arguments, while the inductive analysis extended this lens by identifying four characteristics of their longitudinal development. By examining students' writing across three to four consecutive summers, this study provides a rare longitudinal perspective on written mathematical argumentation within an instructional context intentionally designed to cultivate these practices. Although the sample size is small, the findings suggest that the longitudinal development of written mathematical argumentation is characterized by evolving ways of drawing upon mathematical structure, responding to increasing mathematical demands, attending to an audience, and composing individualized written arguments.

Implications

The findings suggest that supporting students' written mathematical arguments involves more than moving students toward increasingly sophisticated arguments. Repeated opportunities to engage in written mathematical argumentation within classroom communities that routinely value and expect mathematical argumentation may support students in drawing upon mathematical structure, attending to an audience, and making their mathematical reasoning accessible to others. It is also important to acknowledge that the students received instruction throughout the school year that may have influenced changes in their arguments. For research, the findings demonstrate that longitudinal analyses can reveal characteristics of development that extend beyond movement through existing developmental classifications. Future work should examine these characteristics across broader populations and instructional contexts to better understand how written mathematical argumentation develops over time.

References

Civil, M., & Hunter, R. (2015). Participation of non-dominant students in argumentation in the mathematics classroom. *Intercultural Education*, *26*(4), 296–312.

Kosko, K. W., & Zimmerman, B. S. (2019). Emergence of argument in children's mathematical writing. *Journal of Early Childhood Literacy*, *19*(1), 82–106.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). Sage Publications.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Lehmann, T. H., & Friend, L. (2025). Mathematical argumentative writing in K–12 education: A systematic literature review. *International Journal of Science and Mathematics Education*, 23, 3679–3705.

Lerman, S. (2001). Cultural, discursive psychology: A sociocultural approach to studying the teaching and learning of mathematics. *Educational Studies in Mathematics*, *46*(1), 87–113.

National Center for Education Statistics. (n.d.). *NAEP questions tool* [Grade 3 & 4 Sample Assessment Items]. U.S. Department of Education, Institute of Education Sciences. Retrieved [2022, 2023], from https://www.nationsreportcard.gov/nqt/

Rojas-Drummond, S., & Zapata, M. P. (2004). Exploratory talk, argumentation and reasoning in Mexican primary school children. *Language and Education*, *18*(6), 539–557.

Smarter Balanced Assessment Consortium. (n.d.). *[Grade 3 & 4 Sample Assessment Items]*. https://sampleitems.smarterbalanced.org/

Toulmin, S. (1958). *The uses of argument*. Cambridge University Press.

ELICITING MATHEMATICAL THINKING: TEACHERS' QUESTIONING DURING READ-ALOUDS

Mahtob Aqazade
Illinois State University
maqaza1@ilstu.edu

Laura Bofferding
Purdue University
lbofferd@purdue.edu

Beatriz Ruiz Ordoñez
Universidad Iberoamericana
beatriz.ruiz@correo.uia.mx

Lisa Hawley
Grand Valley State University
hawleyL@gvsu.edu

Jelena Byers
Purdue University Northwest
byers40@purdue.edu

Abstract

Interactive read-alouds provide opportunities for children to engage in mathematical conversations through teachers' purposeful questioning. However, little is known about how teachers use questioning during interactive read-alouds. This systematic literature review examined 47 studies to characterize teacher questioning during read-alouds. We analyzed the questions by mathematical content area and level of abstraction. Most questions focused on number and operations and occurred at levels 2 and 3 of abstraction, while relatively few addressed geometry, measurement, and mathematical practices. We found limited reporting of teachers' follow-up interactions. Future research should document questioning practices and development of mathematical conversations during interactive read-alouds.

Objectives of the Study

Using children's literature in mathematics provides "a vehicle for communicating mathematical ideas" (National Council of Teachers of Mathematics, 1989, p. 5), through which children can engage in discussions that "foster joy and wonder for mathematics and literacy" (Hintz & Smith, 2022, p. 4). Research studies on using children's literature in mathematics instruction identify beneficial outcomes for children, including increased mathematics achievement (Capraro & Capraro, 2006; van den Heuvel-Panhuizen et al., 2016), interest in and a more positive attitude towards mathematics (e.g., McAndrew et al., 2017), and engagement in mathematical discourse and use of mathematical vocabulary (e.g., Hassinger-Das et al., 2015; Purpura et al., 2017). Interactive read-alouds are a powerful and common instructional practice that allow children to explore mathematical ideas in the story and connect to their own experiences (Anderson et al., 2004). During interactive read-alouds, teachers can stop at key points, invite children to share their noticings and wonderings, and incorporate their responses in discussion across literacy and mathematical ideas

Anderson-Pence, K., & Shelton, R. N. (Eds.). (2026). *Proceedings of the 125th annual convention of the School Science and Mathematics Association* (Vol. 13). Chicago, IL: SSMA.

(e.g., Hintz & Smith, 2022; Walsh & Hodge, 2018). Thus, teacher questioning becomes essential to not only draw children's attention to such ideas but also extend and investigate them (Linder & Bennett, 2020). While questions, such as "What do you notice? What do you wonder?" can elicit children's thinking (Hintz & Smith, 2022), using mathematical questions anchored in the story, like "How many cats are there now?" can lead to a meaningful and integrated discussion (Anderson et al., 2004). Despite the recognized importance of teacher questioning through a mathematical lens during interactive read-alouds, research provides limited insight into how these questioning practices are developed and implemented across elementary grades. In this study, we conducted a systematic literature review to examine how teachers or researchers incorporated questioning during interactive read-alouds and the instructional practices that followed. The following research question guided this study: How do studies describe teachers' use of questioning during interactive read-alouds in mathematics instruction?

Theoretical Framework and Related Literature

By engaging in interactive read-alouds, children can construct knowledge through reciprocal interaction between their social environment and their internal interpretation of the story. Some of these interactions can emerge during children's explorations (e.g., questions that they ask an adult or peers), some might be embedded in the story for children to act upon (e.g., stories that have questions in them), and some can reflect a teacher's (or adult's) conversations with children (e.g., questions that a teacher or adult asks children). Thoughtfully designed mathematical questions (posed by teachers, researchers, or children) can initiate meaningful conversations. Despite prior work illuminating teacher questioning in mathematics instruction, research is scarce in the context of interactive read-alouds. To characterize the role of teacher questioning in these conversations, Torbeyns et al.'s (2024) framework described categories of mathematical questions by level of abstraction during read-alouds. Mathematical questions provoke conversations about number and operations (e.g., counting, comparing and ordering), measurement (e.g., act of measuring), or geometry (e.g., shapes, patterns, spatial relations). Level 1 questions draw children's attention to salient, page-level information by asking them to name or notice objects (e.g., What number is this?). Level 2 questions ask children about a scene or to name characteristics of objects, which require focusing on specific aspects and integrating separate components of salient information in the story (e.g., How many ducklings do you think there are altogether?). Level 3 questions involve reordering or restructuring information from parts of the story that are not saliently visible, including making

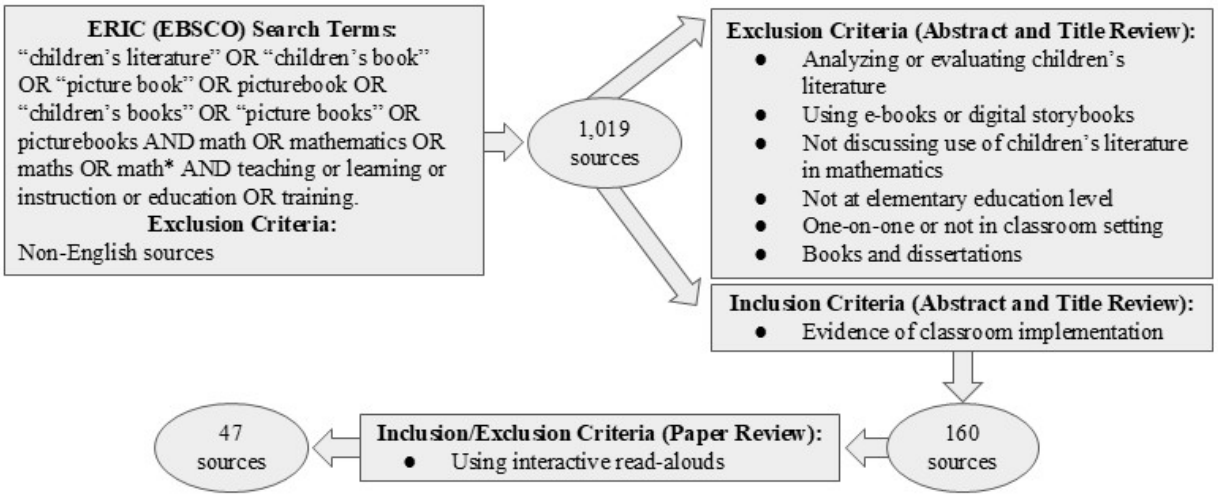
judgements or summarizing (e.g., How many cats are there now? How do you know?). Level 4 questions ask for children's predictions about the story or solving problems (e.g., How many ways could we spend a nickel buying candy marked one cent?). In this paper, we systematically reviewed prior scholarly work to explore how questions are used in interactive read-alouds to support mathematical conversations and learning, including who generated and asked the questions.

Methods

In early 2025, we used specific search terms, as shown in Figure 1, in the educational database ERIC (EBSCO), with no date-of-publication restrictions and initially excluding non-English papers. Although we continued our search in other databases to extend our work, we initially used ERIC (EBSCO) because of its comprehensive indexing of education-specific research. After at least two researchers reviewed the abstracts and titles, we excluded papers that focused on analyzing or evaluating children's literature; only used e-books or digital storybooks; did not focus on children's literature in mathematics; had participants beyond sixth grade; used one-on-one or non-classroom settings; or were books or dissertations. We focused on papers that involved evidence of classroom implementation of children's literature. Next, we reviewed the papers to determine whether they used interactive read-alouds, which resulted in 47 sources (see Figure 1).

Figure 1

Inclusion and Exclusion Criteria and Results



The first two authors identified the questions involved in 47 sources. We then together classified the mathematical questions by content area (e.g., number and operations, measurement, geometry, etc.) and level of abstraction (Torbeyns et al., 2024). In our analysis, we found additional

questions that did not align with any of the levels of abstraction, were not content-focused, or were non-mathematical questions but led to additional mathematical conversations and questions. We discuss these sources with examples in the results. Finally, we noted who asked these questions (i.e., teachers or researchers) and who generated the questions (i.e., teacher, researcher, or story).

Results and Discussion

Out of 47 sources, 32 sources were studies conducted in the United States, and the rest came from Belgium, Sweden, the Netherlands, Greece, Canada, Australia, Turkey, China, and Singapore. They targeted pre-kindergarten to sixth grade. Five sources did not provide any examples of questions. Two sources contained only non-mathematical questions (e.g., "What do you think the story will be about?"; Shatzer, 2008).

Twenty-two sources solely focused on mathematical questions, with 14 primarily related to number and operations. Examples include, "What do you see and how many?" (Gonulates & Gilbert, 2001, p. 8), "Do you see 10?" (Ang et al., 2024, p. 14), "Did the caterpillar eat more pears than oranges?" (Green & Towson, 2022, p. 93), and "We need to know how many boxes there are. How can we learn that?" (Lewis et al., 1993, p. 471). Six sources included questions about geometry and measurement, such as "If we put those triangles back together, what shape would we have?" (Jenner, 2002, p. 168) and "What are two places that you find this shape outside your classroom?" (Mack, 2007, p. 242). The last two sources related to recognizing patterns asked, "What pattern would you put the party hats in?" (Wijns et al., 2023, p. 738). Seventeen of the 47 sources contained both mathematical and non-mathematical questions. Often, the interactive read-alouds began by asking non-mathematical questions, such as "What do you think this book will be about?" and followed up with mathematical questions, "What number is this? Let's say it together" (Alfes et al., 2021, p. 131). From these 17 sources, 13 focused on number and operations and four on geometry and measurement. In addition to mathematical questions directly connected to content areas, a few sources contained questions related to mathematical practices, such as "Can you use the picture in some way to keep track?" (Jenner & Anderson, 2000, p. 546) and "What does it mean to look for structure?" (Young et al., 2021, p. 135).

Within the sources, we found 49 occurrences of non-mathematical questions. Sometimes, these non-mathematical questions led to mathematical conversations and further mathematical questions at different levels of abstraction. Some questions that may not appear mathematical when viewed in isolation can take on mathematical meaning within a specific classroom context. For

example, in Whitin and Whitin (2006), the teacher initiated the interactive read-aloud by asking, "What do you notice?" for each illustration in *More, Fewer, Less*. The teacher then extended children's mathematical thinking with questions such as, "How many cheeses have orange edges? How many have red edges? Are there more cheeses with orange edges or more with red edges?" (p. 197). These questions exemplify level 2 of abstraction, as they direct children's attention to salient features of the illustrations while prompting them to identify, count, compare, and describe mathematical relationships. Additionally, the question, "What kind of questions could we explore about this story?" (Wickett, 1997, p. 476), is sufficiently open-ended that it could be interpreted as mathematical or non-mathematical. However, based on their class routines and expectations, students in their study understood the question as an invitation to explore mathematical ideas.

Overall, the majority of mathematical questions were at levels 2 and 3, with each level represented 54 times across the sources. This was followed by levels 1 and 4, with each occurring 28 times across the sources. These sources, however, were not explicit on how these levels of questions elicited students' mathematical thinking, as they rarely provided excerpts of the mathematical conversations followed by these questions. Table 1 shows examples of these mathematical questions.

Table 1

Examples of Levels of Abstraction in Mathematical Questions

Levels of Abstraction	Examples
Level 1	"What number is this?" (Ang et al., 2024, p. 29); "How much is that rose in the window?" (Harris, 1999, p. 522); "What did these shapes remind him of?" (Skoumpourdi & Mpakopoulou, 2011, p. 201)
Level 2	"How many small cars?" (Whitin et al., 1994, p. 170); "What's the pattern that you see?" (Jenner & Anderson, 2000, p. 545)
Level 3	"If we put those triangles back together, what shape would we have?" (Jenner, 2002, p. 168); "What if there were three pigs in another village? How many pigs would there be altogether?" (Margerm, 1999, p. 87)
Level 4	"If Benny buys a rose for a penny, how many pennies will he have left?" (Harris, 1999, p. 522); "What are two places that you find this shape outside your classroom?" (Mack, 2007, p. 242)

Overall, teachers asked the majority of questions, while researchers asked the questions in six sources, and children asked *some* of the questions in two of the sources. In four sources, the

questions were written in the book as part of the story, and in one source, teachers asked questions developed by researchers. Students' responses to teacher-initiated questions were largely focused on aspects of the stories that teachers highlighted. In contrast, a *wonder* orientation (e.g., Hintz & Smith, 2022) could encourage students to explore the mathematics they see in the stories.

Implications

Overall, our systematic literature review results illustrated limited information about the questions asked during mathematics-focused interactive read-alouds, making it difficult to understand how teachers center mathematical ideas through discussion. Additionally, we could only unpack mathematical conversations followed by teacher questioning in a few of these sources through excerpts they provided. Future studies should document not only how the questions are posed but also students' responses and teachers' follow-up interactions to provide a more comprehensive picture of how mathematical conversations and other instructional practices develop around a story. The instructional practices teachers use to facilitate children's responses influence the quality of these conversations during interactive read-alouds. One such practice that could support children's mathematical thinking and learning is talk moves (i.e., revoicing, repeating, reasoning, adding on, and waiting; Chapin et al., 2009). A shift toward talk moves could also encourage more teachers to encourage students to pose questions. Further, examining read-aloud interactions through the lens of talk moves (Chapin et al., 2009), for example, could further illuminate how teachers extend, clarify, and deepen students' mathematical thinking during this practice. Additionally, because the majority of reported questions focused on number and operations, further research capturing questions related to geometry, measurement, and data analysis is needed to better characterize the scope of read-aloud questioning, which could then lead to investigations of the efficacy of different questions to support students' thinking, especially in upper elementary grades. Several questions were difficult to classify, or in the case of mathematical practices, could not be categorized using the framework proposed by Torbeyns et al. (2024). These findings contribute to refining and expanding this framework to better capture a broader range of questions that emerge during interactive read-alouds that future research can address. For example, although level 4 questions extend beyond the immediate context of the story and may be more abstract than level 2 questions, they may not be more advanced in terms of difficulty than level 2 questions.

References

Anderson-Pence, K., & Shelton, R. N. (Eds.). (2026). *Proceedings of the 125th annual convention of the School Science and Mathematics Association* (Vol. 13). Chicago, IL: SSMA.

*Alfes, C., Missall, K., & Meeker, K. A. (2021). Shared book reading effects on number naming in preschoolers at risk for numeracy delays. *Infants & Young Children*, *34*(2), 122–140.

Anderson, A., Anderson, J., & Shapiro, J. (2004). Mathematical discourse in shared storybook reading. *Journal for Research in Mathematics Education, 35*(1), 5–33. https://www.jstor.org/stable/30034801

*Ang, D., Sun, B., & Cheung, P. (2024). The effect of book genre in eliciting teacher math talk. *Mind, Brain, and Education*, *18*(4), 408–416. https://doi.org/10.1111/mbe.12433

Capraro, R. M., & Capraro, M. M. (2006). Are you really going to read us a story? Learning geometry through children's mathematics literature. *Reading Psychology*, *27*(1), 21–36. https://doi.org/10.1080/02702710500468716

Chapin, S. H., O'Connor, M. C., & Anderson, N. C. (2009). *Classroom discussions: Using math talk to help students learn, Grades K–6*. Math Solutions.

*Gonulates, F., & Gilbert, J. (2021). Early mathematics within rich story and conversation. *Dimensions of Early Childhood*, *49*(2), 6–14.

*Green, K. B., & Towson, J. P. (2022). Easy as 1, 2, 3, ABC: Integrating number sense and shared storybook readings. *Young Exceptional Children*, *25*(2), 88–100.

*Harris, J. (1999). Links to literature: Interweaving language and mathematics literacy through a story. *Teaching Children Mathematics*, *5*(9), 520–524. https://doi.org/10.5951/TCM.5.9.0520

Hassinger-Das, B., Jordan, N. C., & Dyson, N. (2015). Reading stories to learn math: Mathematics vocabulary instruction for children with early numeracy difficulties. *The Elementary School Journal*, *116*(2), 242–264. https://doi.org/10.1086/683986

Hintz, A., & Smith, A. T. (2022). *Mathematizing children's literature: Sparking connections, joy, and wonder through read-alouds and discussion*. Stenhouse.

*Jenner, D. M. (2002). Experiencing and understanding mathematics in the midst of a story. *Teaching Children Mathematics*, *9*(3), 167–171. https://doi.org/10.5951/TCM.9.3.0167

*Jenner, D. M., & Anderson, A. G. (2000). Experiencing mathematics through literature: The story of Neil. *Teaching Children Mathematics*, *6*(9), 544–547. https://doi.org/10.5951/TCM.6.9.0544

*Lewis, B. A., Long, R., & Mackay, M. (1993). Fostering communication in mathematics using children's literature. *The Arithmetic Teacher*, *40*(8), 470–473. https://doi.org/10.5951/AT.40.8.0470

Linder, S. M., & Bennett, A. (2020). Leveraging read alouds for mathematical connections. *Mathematics Teacher: Learning and Teaching PK–12, 113*(4), 317–321. https://doi.org/10.5951/MTLT.2019.0215

McAndrew, E. M., Morris, W. L., & Fennell, F. (2017). Geometry-related children's literature improves the geometry achievement and attitudes of second-grade students. *School Science and Mathematics*, *117*(1–2), 34–51. https://doi.org/10.1111/ssm.12202

*Mack, N. K. (2007). Gaining insights into children's geometric knowledge. *Teaching Children Mathematics*, *14*(4), 238–245. https://doi.org/10.5951/TCM.14.4.0238

*Margerm, P. (1999). Links to literature: An old tale with a new turn-and flip and slide. *Teaching Children Mathematics*, *6*(2), 86–90. https://doi.org/10.5951/TCM.6.2.0086

Purpura, D. J., Napoli, A. R., Wehrspann, E. A., & Gold, Z. S. (2017). Causal connections between mathematical language and mathematical knowledge: A dialogic reading intervention. *Journal of Research on Educational Effectiveness, 10*(1), 116–137. https://doi.org/10.1080/19345747.2016.1204639

*Shatzer, J. (2008). Picture book power: Connecting children's literature and mathematics. *The Reading Teacher*, *61*(8), 649–653. https://doi.org/10.1598/RT.61.8.6

*Skoumpourdi, C., & Mpakopoulou, I. (2011). The prints: A picture book for pre-formal geometry. *Early Childhood Education Journal*, *39*(3), 197–206. https://doi.org/10.1007/s10643-011-0454-0

Torbeyns, J., Eynde, E. O., Depaepe, F., & Verschaffel, L. (2024). Preschool teachers' mathematical questions during shared picture book reading. *ZDM, 56*(5), 907–921. https://doi.org/10.1007/s11858-023-01544-w

van den Heuvel-Panhuizen, M., Elia, I., & Robitzsch, A. (2016). Effects of reading picture books on kindergartners' mathematics performance. *Educational Psychology*, *36*(2), 323–346. https://doi.org/10.1080/01443410.2014.963029

*Whitin, D. J., Mills, H., O'Keefe, T., & Thiessen, D. (1994). Exploring subject areas with a counting book. *Teaching Children Mathematics*, *1*(3), 170–175.

*Whitin, P., & Whitin, D. J. (2006). Making connections through math-related book pairs. *Teaching Children Mathematics, (13)*4, 196–202. https://www.jstor.org/stable/41198910

*Wickett, M. S. (1997). Serving up number sense and problem solving: "Dinner at the Panda Palace." *Teaching Children Mathematics, 3*(9), 476–480. https://www.jstor.org/stable/41198516

*Wijns, N., Purpura, D., & Torbeyns, J. (2023). Stimulating preschoolers' repeating patterning ability by means of dialogic picture book reading. *Journal of Educational Psychology*, *115*(5), 732–746. https://psycnet.apa.org/doi/10.1037/edu0000756

*Young, T. A., Monroe, E. E., & Roth-McDuffie, A. (2021). Picturebook biography read-alouds and standards for mathematical practice. *The Reading Teacher, 75*(2), 135–146. https://doi.org/10.1002/trtr.2019

* Reflects several papers identified in the systematic literature review process.

THE USE OF AI IN DESIGNING MATHEMATICS INTERVENTIONS FOR STUDENTS WITH DISABILITIES

Mahtob Aqazade
Illinois State University
maqaza1@ilstu.edu

Sebnem Tekin
University of Saint Joseph
stekin@usj.edu

Matt Mauntel
Tarleton State University
mmauntel@tarleton.edu

Abstract

A key affordance of GenAI is its capacity to rapidly adapt instructional materials for individual learners, which can benefit teachers who design individualized interventions. This study explored how in-service teachers used GenAI to design mathematics interventions for students with learning disabilities using the universal design for learning (UDL) math framework. We identified three approaches to AI use: producers, adapters, and explorers. We found that while GenAI sometimes implicitly incorporated UDL elements, intentional and iterative prompting resulted in more purposeful, individualized interventions. Teachers noted GenAI helps generate and refine ideas for task design and assessment, but lacks knowledge of students' needs.

Objectives of the Study

Inclusive mathematics instruction requires moving away from deficit-based perspectives that position students, particularly those with learning disabilities, as less capable of complex mathematical thinking, often resulting in an emphasis on procedural fluency over conceptual understanding (Kurtz et al., 2014; Lambert, 2024). However, research at the intersection of mathematics and special education highlights the importance of recognizing students' diverse approaches to problem-solving and providing meaningful, intellectually challenging learning opportunities (Lambert, 2024). At this intersection, Lambert (2021) developed the universal design for learning (UDL) Math framework, including engagement, representation, and strategic action to support the design of accessible instructional materials for *all* students. With increasing attention to generative artificial intelligence (GenAI), studies have shown increases in mathematics achievement and decreases in mathematics anxiety of students with learning disabilities (e.g., Lessing & Ogbonnaya, 2025; Polydoros et al., 2025). Although these studies have demonstrated the capability of GenAI tools in providing targeted interventions addressing students' needs (Yang et al., 2025), research remains scarce regarding how teachers can use GenAI tools to design such interventions.

Anderson-Pence, K., & Shelton, R. N. (Eds.). (2026). *Proceedings of the 125th annual convention of the School Science and Mathematics Association* (Vol. 13). Chicago, IL: SSMA.

In this study, we explored four teachers' design of mathematics interventions for students with learning disabilities using GenAI tools. We particularly focused on how they utilized the UDL Math framework to prompt the GenAI tools to develop high-quality tasks tailored to their students' needs. Additionally, we investigated these teachers' reflections on their experience using GenAI tools. The following research questions have guided our study:

1. In what ways do teachers utilize the UDL Math framework to prompt GenAI tools when designing mathematics interventions for students with disabilities?
2. What do teachers' reflections reveal about the application of GenAI tools when designing mathematics interventions for students with disabilities?

Theoretical Framework and Related Literature

To challenge deficit-based perspectives on students with learning disabilities and develop meaningful and challenging mathematical learning opportunities, Lambert (2021) developed the UDL Math framework with three design elements: engagement, representation, and strategic action. The engagement element highlights making mathematics relevant to students. The representation element refers to the multimodality of opportunities to understand core mathematical ideas. The strategic action element includes supporting students' identities as learners of mathematics and planning equitable assessment to inform student learning. Given the crucial role of learning goals informing the lesson planning process, Lambert (2024) proposed that the individualized education program (IEP) goals should (a) be measurable, (b) be ambitious, (c) be toward access to the general education curriculum, (d) have a high-leverage focus on core mathematical ideas and practices, and (e) include stakeholders, including the student, family, and the general education teacher.

Research has shown that GenAI tools can assist teachers in generating and adapting instructional tasks and assessments that support personalized learning and address students' needs (Trust et al., 2023). This capability may be particularly relevant for special education teachers, who frequently need to develop IEP goals and interventions by improving access to general education curriculum for students with learning disabilities (Yang et al., 2025). However, teachers often report feeling unprepared to use GenAI tools due to their rapid growth, the variety of these tools, and a lack of guidelines on effective teaching practices for GenAI incorporation (Naresh et al., 2025). One way to support teachers is to provide opportunities for integrating GenAI with frameworks such as the UDL Math framework that promote inclusive mathematics instruction. Building on prior work, our study explores the use of GenAI in developing IEPs aligned with the UDL Math framework.

Methodology

Study Context

This study took place at a private university in the Northeastern United States. Ten in-service teachers pursuing a special education certification were enrolled in an eight-week course focused on evidence-based practices for designing mathematics interventions for students diagnosed with autism or neurodiversity. As part of the course, they interviewed two K–12 students diagnosed with learning disabilities to explore their experiences with mathematics and their mathematical thinking. Then, they chose one of these students and developed a two-day IEP intervention plan–IEP goal, relevant tasks, and assessments. While developing the IEP intervention plan, they used a GenAI tool of their choice, shared their interactions with it, and reflected on their experiences (e.g., what challenges/affordances did you encounter?). Out of 10 in-service teachers, four, with a range of teaching experience, grade levels, and prior experiences with GenAI, participated in this study.

Data Sources and Analysis

In addition to the reports of two interviews and IEP intervention plans, we collected the participants' chatlogs with the GenAI tool, including their prompts, the GenAI's responses, and their reflections on using the tool. We first present the results by introducing the in-service teachers and their focus students. To answer the first research question, we coded the prompts to identify their purpose for using GenAI and the presence of the UDL Math framework. A thematic analysis (Saldaña, 2021) of this coding process resulted in three cases of potential AI users: producers (Ellie and Melody), adapters (Joy), and explorers (Sophie). In each case, we discussed how they incorporated the UDL Math framework (Lambert, 2021) in their prompts. To answer the second research question, we described their reflections on the affordances and challenges with GenAI.

Results

Producer: The Cases of Ellie and Melody

Melody and Ellie taught early elementary grades, including mathematics, with less than three years of teaching experience. Ellie's experience with mathematics as a student was positive, helping her enjoy teaching and learning mathematics. Ellie worked with a second grader, Charlie, whose primary diagnoses were other health impairment and autism. Because Charlie relied on direct

modeling and trial and error to solve equal-sized group problems during the interview, Ellie's IEP intervention focused on such problems to support Charlie in developing more efficient strategies.

Melody had mixed experiences with mathematics. While mathematics did not come easily to her in elementary and middle grades, her teachers built her confidence by taking the time to support her and making the material more accessible in high school. Melody worked with a second-grade student, Sam, who received Speech and Language services. Sam demonstrated strong number flexibility and mental strategies, but could not articulate the regrouping in the base-ten number system. Thus, Melody focused on real-life subtraction problems that incorporated base-ten blocks, Unifix cubes, visuals, and opportunities for Sam to draw her mathematical thinking to develop a more robust understanding of grouping by 10 for the IEP intervention plan.

Ellie and Melody were producers because they used prompts that generated an IEP goal or an IEP intervention plan, which were not revised. Ellie used ChatGPT to "give a clear description of two instructional tasks that I could use with a material and instruction list" aligned with her IEP goal. Although Ellie did not utilize any elements of the UDL Math framework in her prompt, ChatGPT's first task involved a concrete-to-representational sequence, and the second task included a representational-to-abstract sequence. These instructional task sequences are aligned with the representation element of the UDL Math framework, which Ellie noticed and highlighted in her IEP intervention plan report when making a connection to the framework. Along with producing the two tasks, ChatGPT offered to "create a data collection sheet aligned directly to this goal for progress monitoring," to which Ellie responded, "Yes please."

Melody used Google Gemini to produce a "high-leverage and ambitious" IEP goal for addition and subtraction problems "within 100." She then prompted Google Gemini to "create examples of the specially designed instruction tip activity" and later, focusing on "a set of addition and subtraction problems within 20." Although Melody prompted Google Gemini to "make the swap shop activity around the UDL principle," she did not ask for specific elements of the UDL Math framework, nor did she highlight them in her IEP intervention plan report. Her remaining interactions involved producing visuals and an exit slip for the activity.

Adapter: The Case of Joy

Joy had five years of high school teaching experience. Although she had prior experience teaching mathematics, she was not teaching mathematics at the time of this study. Joy expressed her interest in learning about different strategies and skills to improve her mathematics teaching. She

worked with Peter, a high school junior. He was a neurodivergent student with an autism diagnosis. When Peter worked on identifying patterns and determining the general expression, he struggled to see "the relationship between the shape number and the number of toothpicks." Thus, Joy focused on finding general expressions for visual patterns with word problems in her IEP intervention plan.

Joy leveraged AI to adapt her plan with a clear initial IEP goal, representing an adapter user of AI. She used Google Gemini to adapt her IEP goal, specifically aligning better with "measurable, ambitious, towards access, high leverage," which all reflect the qualities of an effective IEP goal (see Lambert, 2024). Although Joy prompted Google Gemini to generate a two-day IEP intervention plan, she focused on the "multimodal" element of the UDL Math framework (see Lambert, 2021) and specified "hands-on when possible…while incorporating his [Peter's] interests." Joy continued her interactions with Google Gemini to adapt the activity for "visually determin[ing] constant vs. the slope," while asking "should he have more days of finding and identifying this in a word problem before he starts graphing?" She also wanted to "provide multiple ways for him to solve the patterns" and support Peter's "progression naturally." These prompts are informed by the elements of the UDL Math framework, including representations, engagement (e.g., meaningful tasks for students' interest), and strategic action (e.g., assessment informing growth). Thus, Joy explicitly incorporated the UDL Math design elements in her prompting.

Explorer: The Case of Sophie

Sophie taught middle school, including mathematics, with more than 10 years of teaching experience. As a student, Sophie shared that she put a lot of effort into understanding mathematics, but she could not grasp concepts until she started teaching mathematics to students with learning disabilities as a paraeducator. Sophie worked with a sixth-grade student, Sarah, who received services for attention deficit disorder/attention deficit-hyperactivity disorder. In the IEP plan, Sophie focused on division of whole numbers to build Sarah's "confidence using box area arrays for multiplication to expand her knowledge of division of large quantities." Sarah could use the standard algorithm for long division but "got confused with each step and where to put numbers."

Sophie is an explorer user of AI because she explored different methods for whole number divisions and how those methods are connected. Sophie used SchoolAI, which is specifically trained to create materials for K–12 educational settings. Sophie's initial prompt was a "lesson task to help students reach an IEP objective given explicit math instruction; the student will divide a multidigit whole number using the standard algorithm." Although she did not incorporate the UDL Math

framework elements in her initial prompt, because of her AI choice, SchoolAI provided different supports and approaches, including multiplication charts, skip counting, and partial products. Sophie decided to go beyond the provided material and prompted whether "it would be confusing to introduce the box method for division alongside the standard algorithm?" and followed up by asking for suggestions about how to teach students the "connection between the box array method for division of multi-digit numbers and the standard algorithm." Sophie focused on the core idea of division with multi-digit numbers and different representations reflecting the elements of the UDL Math framework (see Lambert, 2021), despite not directly using the elements in her prompts. Her prompts also implicitly reflect the engagement element, where she aimed to build meaningful connections between different methods of doing long division.

Reflections on GenAI Applications

In her reflection, Ellie wrote that ChatGPT was helpful as "it broke each of the two tasks down into different steps" where she "was able to vision what the teacher and what the student would be doing throughout the task." Although Ellie stated that ChatGPT "did not help identify any potential barriers to students' learning," she did not prompt ChatGPT for any revisions to strengthen the two tasks in her IEP intervention plan. She solely relied on ChatGPT and said, "There was nothing that I disliked." Melody noted that Google Gemini's second task contained "UDL components and high leverage aspects," but "had to be very specific with the grade level." Thus, she ended up "rework[ing] some components on my own offline" from the "baseline they generated." Melody appreciated Google Gemini's offer on "data collection tools, guiding questions, and other support strategies for every component asked." Similar to Ellie, Melody was unsure whether AI can produce "an individualized plan for a student they don't know," but liked "utilizing it [AI] to help generate ideas and create formative assessment tools." Thus, for Ellie and Melody, knowing their students was a crucial input for designing effective IEP plans, which AI tools lacked.

Joy viewed AI as both supportive and challenging in developing her IEP plan. She leveraged Google Gemini to aid "in developing problems that were relevant to my students' specific interests" and provide "feedback on my goal and helping me to make my goal more ambitious." Unlike Ellie and Melody, she explained how AI "was pretty good at incorporating the strengths and challenges of the student when helping me to cater tasks to the student." She added how her interactions with Google Gemini improved the IEP goal to be "quite measurable and a little more ambitious than simply understanding linear equations." However, Joy found her use of AI challenging when

uploading her original task and needed to explain the task and the student's current level to Google Gemini instead. Although Joy believes that "AI has a place," she does not think AI can replace the "teacher's ability to know how to sequence instructions in a way that will be well received by their students" because teachers better know their students.

Sophie started her reflection by noting that she found "AI a bit scary." However, she noted the need for understanding AI as her school had "initiated a great focus" on using AI in the classroom. She found AI useful to generate ideas, but she had issues with framing her prompt and wrote that "changing my wording or including key words changed the direction significantly." She observed that AI could be successfully leveraged to "explore different methods and broaden my understanding of outside-the-box thinking to best support my learners." However, she was concerned about AI's potential to impact "independent thinking, learning and critical skills." Sophie also mentioned how she learned to solve mathematical tasks in particular ways (e.g., standard division algorithm) and leveraged AI to explore alternative solution strategies (e.g., box method).

Conclusion, Discussion, and Implications

Across the producers, adapters, and explorers, teachers varied in how explicitly they incorporated the UDL Math framework. Ellie and Sophie did not reference the framework in their prompts, although their AI-generated outputs sometimes implicitly reflected its elements. Ellie's use of ChatGPT primarily as a producer, combined with her positive but limited prior experience with AI, may have encouraged greater reliance on initial outputs and less critical revision. Similarly, although Melody asked Google Gemini to incorporate the UDL Math framework, her broad prompts positioned her primarily as a producer. One way to support producer users of AI is to facilitate discussions around AI being prone to making mistakes and the importance of continuous revisions to improve GenAI's outputs (Naresh et al., 2025; Sawyer & Aga, 2025). In contrast, Joy used targeted prompts related to specific elements within the UDL Math framework, such as multimodality, and she adapted IEP goals to be measurable, ambitious, towards access, and high leverage. Her case as an adapter user of AI illustrates how iterative and purposeful prompting can support more meaningful and intellectually challenging mathematical learning opportunities. Sophie's exploratory use of School AI also suggests that sustained interactions with GenAI can help teachers refine both their pedagogical ideas and their prompting techniques. Research similarly emphasizes the value of interactive and purposeful prompting for improving interactions with GenAI. Viewed alongside Lambert's (2024) conceptualization of IEP goals, these cases suggest that

teachers can benefit from drawing on all three users. As producers, they can generate materials for measuring progress toward IEP goals; as adapters, they can modify those materials to address students' individual needs and increase access to the curriculum; and as explorers, they can investigate mathematical connections and identify ambitious, high-leverage ideas. Thus, supporting teachers' use of GenAI tools requires more than introducing these tools and instructional frameworks; it also requires developing their pedagogical decision-making when using GenAI and prompting techniques to critically evaluate and adapt AI-generated outputs.

In their reflections, teachers acknowledged that these tools are useful for generating and refining ideas. However, they also recognized that GenAI lacks a nuanced understanding of their students' individual needs, interests, and learning contexts. Consequently, it is crucial to provide detailed and thoughtful prompts to develop IEP intervention plans tailored to individual students, despite the fact that doing so can be both time-consuming and labor-intensive. Additionally, teachers expressed concerns about the extent to which they could share information about their students with AI due to data privacy considerations. Thus, when GenAI tools are used, it is important for teachers to be provided with clear instructions surrounding data privacy protection while maximizing their abilities to generate, adapt, and explore learning opportunities for students.

References

Kurz, A., Elliott, S. N., Lemons, C. J., Zigmond, N., Kloo, A., & Kettler, R. J. (2014). Assessing opportunity-to-learn for students with disabilities in general and special education classes. *Assessment for Effective Intervention*, *40*(1), 24–39.

Lambert, R. (2021). The magic is in the margins: UDL math. *Mathematics Teacher: Learning and Teaching PK–12, 114*(9), 660–669.

Lambert, R. (2024). *Rethinking disability and mathematics: A UDL math classroom guide for grades K–8*. SAGE Publications.

Lessing, M. J., & Ogbonnaya, U. I. (2025). Artificial intelligence in mathematics education of students with special educational needs. *Problems of Education in the 21st Century, 83*(6), 786–806.

Naresh, N., Yilmaz, Z., Johnson, I. D., & Cockerham, D. (2025). Exploring potential uses and concerns of GenAI use in problem solving in the mathematics classroom. *Ohio Journal of School Mathematics, 101,* 31–51.

Polydoros, G., Galitskaya, V., Pergantis, P., Drigas, A., Antoniou, A. S., & Beazidou, E. (2025). Innovative AI-driven approaches to mitigate math anxiety and enhance resilience among students with persistently low performance in mathematics. *Psychology International, 7*(2), Article 46. https://doi.org/10.3390/psycholint7020046

Saldaña, J. (2021). *The coding manual for qualitative researchers.* SAGE Publications Limited.

Sawyer, A. G., & Aga, Z. G. (2026). A mile high and an inch deep: Exploring ChatGPT as a mathematics curriculum development tool. *School Science and Mathematics, 126*(1), 60–74.

Trust, T., Whalen, J., & Mouza, C. (2023). Editorial: ChatGPT: Challenges, opportunities, and implications for teacher education. *Contemporary Issues in Technology and Teacher Education, 23*(1), 1–23.

Yang, Y., Chen, L., He, W., Sun, D., & Salas-Pilco, S. Z. (2025). Artificial intelligence for enhancing special education for K–12: A decade of trends, themes, and global insights (2013–2023). *International Journal of Artificial Intelligence in Education, 35*(3), 1129–1177. https://doi.org/10.1007/s40593-024-00422-0

SENSING INTERVENTION OPPORTUNITIES IN SMALL-GROUP MATHEMATICAL ARGUMENTATION

Jennifer Kleiman
University of Georgia
jk51087@uga.edu

Ngutor Tembe
University of Georgia
ngutor.tembe@uga.edu

AnnaMarie Conner
University of Georgia
aconner@uga.edu

Abstract

Teachers facilitating small-group mathematics discussions often sense a group needs intervention before they know why. This makes moments for artificial intelligence intervention difficult to identify from audio alone. We analyze a 12-minute segment of fourth-grade small-group mathematical argumentation through three lenses: phenomenologically-informed dwelling, Conner et al.'s (2014) argumentation support framework with extended Toulmin diagramming, and classroom discourse analysis. Across all three, the argumentation is missing a push for reasoning. We relate felt-sense moments to framework elements and report negative findings: talk-time imbalance flagged the student recording the group's thinking, and many of the most consequential moments had no audio proxy.

Keywords: mathematical argumentation; small groups; teacher intervention; classroom discourse analysis; artificial intelligence; phenomenology

Introduction

Mathematical argumentation is recognized as a core disciplinary practice (Conner et al., 2014; Krummheuer, 2007; Toulmin, 2003). Students can practice peer-to-peer argumentation in small groups. Even in small-group discussions, if a teacher is nearby students tend to direct talk toward the teacher rather than each other, so intervening is somewhat like opening Schrödinger's box; if the teacher is close enough to listen and decide, the peer-to-peer argumentation she wants to foster may already be disrupted. Research in mathematics education says much about how to intervene once a teacher decides to act; for example, Conner et al. (2014) catalogue teacher moves in their Teacher Support for Collective Argumentation (TSCA) framework. It says much less about how to tell when intervention is needed.

The when-question is newly relevant. AI-mediated tools now make it possible to listen to small-group talk in real time and flag candidate moments for teacher attention — but only if intervention opportunity can be characterized in terms of features such tools can detect. Ayalon & Nama (2026) identified critical events that prompt teacher response during small-group

Anderson-Pence, K., & Shelton, R. N. (Eds.). (2026). *Proceedings of the 125th annual convention of the School Science and Mathematics Association* (Vol. 13). Chicago, IL: SSMA.

argumentation, but their framework centers on cognitive noticing (what teachers attend to) rather than on AI-detectable features of the talk. Soysal (2023) relates teacher question types and follow-up moves to student argument quality but in whole-class, teacher-led discourse. D'Angelo & Rajarathinam (2024), analyzing audio of teaching assistants circulating among small groups, found that TAs over-intervened: 81 of 141 interventions were followed by only 20 seconds of student discussion before the next intervention. Defining an intervention opportunity is not only about finding moments-that-go-unaddressed; sometimes the right response is to wait and protect the students' space for productive struggle. The felt sense—that *I should step in*—and the interactional mechanics that might relate to that feeling remain underexamined. This paper takes up that gap.

Objectives of the Study

We apply three analytic lenses to the same segment of fourth-grade small-group mathematical argumentation and ask: 1) What characterizes intervention-opportunity moments in this segment, and how does each lens articulate them? 2) Where do the three lenses converge, and what does that convergence license us to say about connecting felt teacher recognition to observable features of discourse? 3) Which identified features are plausibly detectable from audio alone, and which are not?

Theoretical Framework and Related Literature

We treat *intervention opportunity* in small group mathematical argumentation as an observer's judgment, informed by a normative theory of good argumentation, that some response from a teacher or other adult, including deliberate non-intervention that preserves productive struggle (Kapur, 2008), could plausibly improve discourse quality at that moment. Thus conceived, intervention opportunity is observer-relative: different teachers will find different opportunities depending on the situation and their values. Counting deliberate non-intervention as a response risks making every moment an opportunity, since doing nothing is always available. We therefore distinguish moments an observer recognizes and then chooses to leave alone from moments that pass unnoticed; only the former are opportunities.

Our analytic passes are based on three bodies of work. Phenomenological methodology (Freeman, 2021) calls for attention to lived experience before interpretive categories are applied; we use it to examine the felt sense of, "Oh, I really should step in here." Conner et al.'s (2014) TSCA framework with extended Toulmin diagramming (ETD) provides a structural account of collective argumentation: claims (what is asserted), data (the evidence offered for it), warrants (why the data

support the claim), rebuttals (challenges, or conditions under which the claim would not hold), and the teacher moves that elicit and respond to them. Classroom discourse analysis (Rymes, 2016) examines social context, agency, and interactional resources (e.g., turn-taking, pauses, embodied resources) that influence how mathematical reasoning gets voiced; for one teacher-led exchange we additionally draw on Soysal (2023)'s coding of triadic dialogue patterns, follow-up moves, and question cognitive demand. Reading multiple analyses of the same data against each other follows the logic of crystallization (Ellingson, 2009): each lens illuminates part of the phenomenon, combining ways of knowing into a thick description.

Methodology

Data

We analyzed 12 minutes (51:00–63:02) of video of one group in a fourth-grade classroom in a rural Georgia elementary school, collected as part of a university–school STEM integration partnership. In our focal video, the class is tasked with making a quadrilateral on the floor with masking tape, measuring its sides, and writing pseudocode for a robot to traverse the shape, with color-coding to indicate equal-length sides. The group consists of four students (pseudonyms Jim, Hannah, Michael, and Ava); the teacher, Ms. Kelly (pseudonym), circulates among groups and intervenes briefly with this group at 09:46 (times offset from segment start). Data are used under the project IRB protocol. We selected this segment because it has the highest transcription fidelity in the corpus to date and, in our judgment, is representative of the activities we are studying.

Three analytic passes were conducted: 1) *Phenomenologically-informed dwelling.* Two of the authors watched the segment repeatedly, recorded felt-sense reactions wherever their attention was pulled, then organized that into themes (Freeman, 2021). 2) *TSCA with* extended *Toulmin diagramming.* The transcript was first segmented into argumentative episodes, distinguished from procedural and interpersonal exchanges. Six episodes were identified and diagrammed with claims, data, warrants, backing, and rebuttals marked where present, and teacher support actions coded where visible. 3) *Classroom discourse analysis.* Three micro-moments, selected where felt-sense reactions from Pass 1 clustered within argumentative episodes identified in Pass 2, were examined with granular conversation analysis: turn-taking, latching, pauses, overlap, alignment, and non-verbal elements (Rymes, 2016). Triadic dialogue patterns, follow-up moves, and question cognitive demand were applied illustratively to the segment's one teacher interaction, a twelve-second exchange (Soysal, 2023). The three separate analytic artifacts were then read across each other.

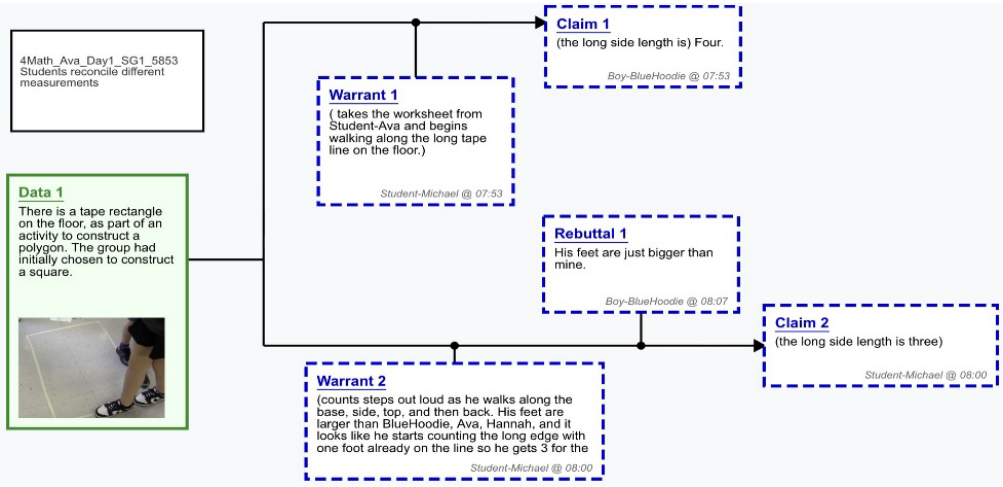
Results and Discussion

Pass 1: Dwelling. Three major themes characterized intervention opportunity as felt before reflection. A primary aspect is that intervention opportunity needs to be understood in *temporal structure* — it emerges in sequence and duration, not in isolated moments. The feeling "I need to intervene now" builds from awareness through watchfulness to urgency only when it is clear the group will not get there on its own, or alternately, when the observer becomes aware that the group is running out of discussion time and could benefit from a push. Another important aspect, which bears on being able to identify intervention opportunity from audio only, is an *embodied-versus-verbal gap*: in this segment much of the mathematical work happens in physical movement (heel-to-toe stepping, kneeling, taping) while the verbal discourse layer is often unrelated to the group's activities. Dwelling also surfaced observer effects: the analyst's own state colors what registers as opportunity (e.g., awareness of remaining class time, attitude about productive struggle, the observer's pedagogical goals).

Pass 2: TSCA/ETD. Across the diagrammed episodes, warrants are almost entirely (a) implicit and never articulated; (b) action-based rather than spoken (heel-to-toe walking *is* the warrant for student measurements of side length in **Figure 1**); or (c) teacher-prompted (Ms. Kelly prompts a definition-level warrant from Michael). Rebuttals are almost entirely absent. A warrant was only articulated once, at 08:07, when two students' step-counts of the same side differ, Jim offers an explanation, shouting (twice) "his feet are just bigger than mine." The spontaneous exchange was part of disagreement over measurement, not deliberative argumentation.

Figure 1

Extended Toulmin diagram for the measurement exchange at 05:58–06:53: embodied action carries the argument; eight words spoken aloud



Pass 3: Discourse analysis. Of the three micro-moments examined, two bear on the signals we develop below. At 07:48–08:07, latching across short turns ("you're making this so complicated" → "why?" → "what is going on" → "I'm so confused") precedes the segment's one unprompted warrant, where Michael takes the pseudocode worksheet and walks the tape himself. At 09:42–09:54 Ms. Kelly intervenes; we analyze the turn pattern with Soysal's (2023) framework. The group had said they would build a square, and the task required coloring different side-lengths differently. The teacher moves are structurally open-ended, where each student response is followed by another question, but her questions stay at the lower-level (recall) of cognitive demand ("What's a rectangle?"), and Michael's responses stay at claim-only level. The form is also less open than it first appears: the open question at 09:46 is not left standing, but it is converted at 09:48 into a yes/no ("Is it four equal sides?") before Michael answers it, a follow-up move that lowers demand rather than raising it. The exchange ends without the group acknowledging that their shape and color were inconsistent and without substantial discussion or argumentation — a missed opportunity.

Excerpt 1

The segment's one teacher-led exchange: open in form, low in cognitive demand (09:42–09:54).

09:42 Ms. Kelly (background, approaching): Well what shape do y'all have?

09:44 Michael: It literally looks like a rectangle.

09:46 [Ms. Kelly joins the group]

09:46 Ms. Kelly: A rectangle? What's a rectangle?

09:48 Ms. Kelly: Is it four equal sides?

09:49 Michael: No.

09:50 Ms. Kelly: No, so you can't make it all the same color?

09:52 Michael: No.

09:53 Ms. Kelly: Or just the same color?

09:54 Michael: That's a rectangle.

Missing reasoning. One finding appears in all three lenses. In dwelling, notes like: "the kids don't ask each other questions or say why they're doing what they're doing" appear repeatedly. In TSCA/ETD, almost all warrants were coded implicit or action-based; only one warrant was explicitly voiced. In the discourse analysis, there is near-zero why-question density, and the segment's one verbal warrant was the result of a disagreement. Missing reasoning was common to many of the intervention opportunities we looked at in this segment, although not all, and the absence of reasoning-talk is not necessarily a call for intervention; it is context-dependent.

Mapping felt sense to audio-detectable signal. Taken together, the three passes characterize intervention opportunity moments in this segment not as isolated events but as reasoning that fails to appear after something has called for it—a conflict, a reversal, the teacher's prompt, a display of confusion, or the end of the task—and that failure registers as opportunity only **when** enough time has passed to make it clear the group will not supply the reasoning itself. We connect chains of signals from what the observer feels (dwelling) to corresponding constructs from the frameworks (ETD or CDA) and consider possible audio proxies. Table 1 presents the chains supported by the analysis.

Table 1

Three intervention signals, from felt sense through framework construct to hypothesized audio proxy

Intervention Signal	Felt sense (dwelling notes)	Framework element	Audio proxy
1. Missing reasons	"the kids don't ask each other questions or say why they're doing what they're doing"	ETD: warrantslot unfilled (implicit, or action-based)	reason-seeking turns that go unanswered; claims without explicit reasoning
1a. unresolved counterclaims	"Why don't they discuss this?"	competing claims with no rebuttals	conflicting claims, then no reconciling
1b. backwards progress without diagnosis	"the redoing indicates they need to strategize"	failure never becomes data for a diagnostic claim	talk about restarting with no discussion of why
1c. no reasoning after intervention	"the group never returns to what shape they had said they would make"	teacher intervention did not prompt argumentation	No reason-talk in the window after departure
2. Unresolved confusion	"I had to watch and see if the students would also become aware"	CDA: confusion display	confusion words open a window; do they re-engage?
3. Off-task after task completion	"This felt like a real intervention opportunity here"	Argumentative episodes cease after task completion	completion talk plus collapse of task-related utterance rate

Signal 1 states missing reasoning as a detection rule: specific moments call for a reason, and absence draws attention. Without context, Signal 1 would fire indiscriminately in this segment. The sub-signals are actionable: each specifies a triggering event, and the absence becomes informative

relative to that trigger. Unresolved counterclaims (1a): three steps vs. four for the same side (06:18–07:53); no one asks why the counts differ aloud, but they are reconciled when one student explains the discrepancy ("His feet are just bigger than mine"). Backwards progress without diagnosis (1b): "Yo, we are restarting that" (03:38) with no discussion of what went wrong, and twenty seconds later "What happened to this side?" / "I don't know" (04:22–04:23). No reasoning after intervention (1c): the teacher leaves at 10:00; the group assigns two colors, but no one says why, and the shape-versus-color inconsistency she obliquely addressed was not discussed. Signal 2, confusion, is seen at 07:48–07:52, but the group resolves it on their own at 07:53 and produces its best argumentation, illustrating how the correct response is sometimes to wait (D'Angelo & Rajarathinam, 2024; Kapur, 2008). Signal 3, task completion, is straightforwardly actionable.

Temporality and determining opportunities from audio. In all frames, for most intervention signals, opportunity *accumulates*, building from awareness, then watchfulness, then urgency only when it is clear the group will not get there on its own. Each signal starts a watch window where time and further dialogue turns provide evidence. If argumentation does not follow, then it is a candidate for intervention (D'Angelo & Rajarathinam, 2024). In our analysis, the felt sense of intervention need did not manifest immediately but only after roughly one to three minutes. We offer this as a preliminary bound on window duration, observed in one segment by two analysts and not yet validated at scale. Durations and operational definitions are empirical questions for future work.

Problematic signals. Per-student talk-time imbalance is used in conversation analytics and is easily computed from audio and used by some computer discourse systems to trigger intervention (de Araujo et al., 2024) but it is misleading here: Ava was silent through the segment's middle stretch while she did the group's central written work, the pseudocode. An audio-only system using this measure may flag productively engaged students. Detecting peer dysregulation as an intervention signal by audio only was also problematic: our dwelling notes, "disruption becomes apparent and then actionable after some time," and in discourse analysis we saw a student taking the enforcer role because no adult is present. Our sense of those intervention opportunities required access to embodied interactions, sociocultural context, and situated knowledge, so no audio proxy signals could substitute, and we do not carry it forward as an intervention signal candidate.

What audio-only analysis excludes. Audio-and-transcript data with coded heuristics leave out much of what our felt sense relied on. Beyond visual phenomena (eye-gaze, a student's

worksheet), our findings drew on classroom norms, the teacher's and students' goals and values, and our own judgment of what constitutes good argumentation.

Implications

For **teacher education**, the three-lens reading suggests that "knowing when to step in" to help small groups argue mathematically could be taught at more than one layer: teacher-educators could offer instruction about discourse features that we found, encouraging teachers to elicit spoken warrants, to use open-form and high-cognitive demanding questions, and to recognize when observant nonintervention is called for.

For **AI tool design**, our pilot analysis suggests that automating the intervention opportunities construct involves defining conditions that will start a timed window and tracking argumentation that occurs within that window. Intervention responses could draw from teacher support actions such as Conner et al.'s (2014) catalog, including "elicit reasoning." Detecting off-task groups should also be within its capabilities. Our segment also bears on intervention move design: a structurally well-formed intervention still produced no argumentation, possibly because its cognitive demand was low. The negative findings inform design constraints: an audio-only system must route moments needing richer perception to the teacher while not crowding out students, and it must be aware of its limitations.

This initial analysis is based on 12 minutes from one video, and these findings are not meant to be generalized. Our ongoing work extends the analysis to additional segments across activity types, tests inter-analyst reliability on moment selection, and evaluates whether the candidate signals can predict moments where we believe an intervention opportunity exists.

References

Ayalon, M., & Nama, S. (2026). Mathematics teachers' noticing of argumentation within small groups. *ZDM – Mathematics Education*, *58*(1–2), 231–245. https://doi.org/10.1007/s11858-025-01741-9

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.

D'Angelo, C. M., & Rajarathinam, R. J. (2024). Speech analysis of teaching assistant interventions in small group collaborative problem solving with undergraduate engineering students. *British Journal of Educational Technology*, *55*(4), 1583–1601. https://doi.org/10.1111/bjet.13449

de Araujo, A., Papadopoulos, P. M., McKenney, S., & de Jong, T. (2024). A learning analytics-based collaborative conversational agent to foster productive dialogue in inquiry learning. *Journal of Computer Assisted Learning*, *40*(6), 2700–2714. https://doi.org/10.1111/jcal.13007

Ellingson, L. L. (2009). *Engaging crystallization in qualitative research: An introduction*. SAGE Publications.

Freeman, M. (2021). Five threats to phenomenology's distinctiveness. *Qualitative Inquiry*, *27*(2), 276–282. https://doi.org/10.1177/1077800420912799

Kapur, M. (2008). Productive failure. *Cognition and Instruction*, *26*(3), 379–424. https://doi.org/10.1080/07370000802212669

Krummheuer, G. (2007). Argumentation and participation in the primary mathematics classroom: Two episodes and related theoretical abductions. *The Journal of Mathematical Behavior*, *26*(1), 60–82. https://doi.org/10.1016/j.jmathb.2007.02.001

Rymes, B. (2016). *Classroom discourse analysis: A tool for critical reflection* (2nd ed.). Routledge.

Soysal, Y. (2023). An exploration of the determinants of middle school students' argument quality by classroom discourse analysis. *Research in Science & Technological Education*, *41*(1), 343–371. https://doi.org/10.1080/02635143.2021.1908981

Toulmin, S. E. (2003). *The uses of argument* (2nd ed.). Cambridge University Press.

Acknowledgments

This work is supported by the National Science Foundation under Grant #DRL2318287. *Opinions, findings, conclusions, or recommendations are those of the authors and do not necessarily reflect the views of the Foundation.*

PRESERVICE TEACHERS' MATHEMATICAL KNOWLEDGE FOR TEACHING: CONTEXTUAL FRACTION MULTIPLICATION

Tegan W. Nusser
Bradley University
tnusser@bradley.edu

Brooke Krejci
University of Wisconsin - River Falls
brooke.krejci@uwrf.edu

Bima Sapkota
University of Kentucky
bsapkota@uky.edu

Bona Kang
Ohio Wesleyan University
bkang@owu.edu

Abstract

This study examines elementary preservice teachers (PSTs)' specialized content knowledge (SCK) through their contextual representations of fraction multiplication across two conditions: solving and pictorially representing them alone, or with additional focus on constructing coherent representations with a focus on referent unit, relationship between factors, and iterating to the original whole. Overall, PSTs wrote more coherent contextual representations requiring multiplicative reasoning following instruction. Furthermore, PSTs who were explicitly prompted to focus on aspects of fraction multiplication in contextual tasks tended to show integrated multiplicative reasoning in their tasks. We discuss how our codebook supported us in unpacking PSTs' SCK through their contextual representations and may be useful to other mathematics teacher educators.

Keywords: Fraction Multiplication, Contextual Representations, Math Knowledge for Teaching

Introduction

Conceptions of fractions, and fraction multiplication in particular, present some of the most persistent difficulties in elementary and middle level mathematics (Askoy & Yazlik, 2017; Sari et al., 2024). Although it has been noted that preservice teachers (PSTs) can compute an accurate product algorithmically (Park et al., 2013), they may require greater support in conceptualizing the relationship between multiplier and multiplicand when solving pictorially (Kim, 2025) and when writing contextual representations (Luo, 2009). As understanding of fractions is a core outcome of both elementary and middle school mathematics (NGA Center & CCSSO, 2010), supporting PSTs in developing content knowledge specific to teaching fractions (i.e., Mathematical Knowledge for Teaching [MKT] Fraction Multiplication) is an important aim of their coursework.

Objectives of the Study

The purpose of this study, part of a cross-institutional project investigating PSTs' conceptions of fraction multiplication, is to enhance PSTs' representational fluency and examine how they construct word problems following instruction. Our research questions are: (1) In what ways do PSTs' contextual representations differ following instruction that explicitly foregrounds referent units? (2) How do PSTs' contextual representations reveal their fraction multiplication reasoning?

Mathematical Knowledge for Teaching Fraction Multiplication

We framed this study using the MKT framework, which Ball and colleagues (2008) defined as "the mathematical knowledge needed to teach mathematics" (p. 395). The framework has six interconnected domains of knowledge that teachers draw upon to perform mathematics-specific instructional tasks. We focus on Specialized Content Knowledge (SCK), which is "mathematical knowledge not typically needed for purposes other than teaching" (Ball et al., 2008, p. 400).

Teachers demonstrate SCK when they connect and coordinate multiple mathematical representations, explain relationships among those representations, unpack mathematical ideas, evaluate unconventional solution strategies, and use concrete models to justify mathematical procedures and reasoning (Ball et al., 2008). Unlike procedural competence, SCK is evidenced when teachers understand the mathematical structures underlying representations and to communicate those structures in ways that support student learning (Sapkota, 2025). Thus, SCK involves representational fluency, including being able to move between various representations (e.g., contextual and visual) through the preservation of underlying concepts (Acevedo Nistal et al., 2009). Thus, we conceptualized PSTs' ability to translate symbolic fraction multiplication expressions into contextual story problems as evidence of SCK. Constructing an appropriate story problem requires PSTs to interpret the multiplicative relationship represented by the multiplicand and multiplier beyond executing an algorithm.

Developing SCK fraction multiplication may be challenging. For example, Luo (2009) shows that creating contextual tasks with meaningfully represented fraction factors requires a high degree of SCK. PSTs' constructed contexts were similar in structure and in units of measure, frequently relying on food-related contexts and interpreting multiplication as "the *n*th multiple of a measure" (Luo, 2009, p. 92). This is problematic in terms of representational fluency as representations are not merely tools for communicating ideas, but instead, are constitutive components of concepts

themselves. One aspect of representational fluency involves representing a contextual situation with an area model, which requires PSTs to coordinate units, partitions, and quantities so that the diagram accurately reflects the mathematical meaning of fraction multiplication (Son & Lee, 2016). In sum, these tasks demand that PSTs move flexibly among symbolic, contextual, and visual representations while preserving the underlying multiplicative relationships. Therefore, examining PSTs' contextual representations provides insight into how they unpack the meaning of fraction multiplication and reveals important aspects of their SCK (Ball et al., 2008).

Methodology

Context and Procedure

For this study, we use data from two Midwestern universities, Institution A and Institution B, due to contrasting instruction regarding writing contextual fraction multiplication tasks. Elementary PSTs from both institutions were enrolled in their first mathematics education course. Institution A included 18 PSTs enrolled in their first of a two-course sequence focused on mathematical knowledge for teaching. Institution B included 17 PSTs also enrolled in their first mathematics teaching course (methods) with their math prerequisite being met with a qualifying standardized test score (e.g., ACT) or completion of a general education mathematics course (e.g., Calculus I).

Both Mathematics Teacher Educators (MTEs) from Institution A and Institution B implemented tasks focused on fraction multiplication in symbolic, contextual and visual representations: (1) whole number × fraction, (2) fraction × whole number, (3) unit × non-unit fraction, (4) unit × unit fraction, (5) non-unit × non-unit fractions, and (6) fractions greater than 1 × non-unit fraction. Institution A's instructor focused instruction on solving and visually representing contextual fraction multiplication tasks over 75 minutes. In addition to solving and visually representing contextual tasks, Institution B's instructor had PSTs write contextual problems, share them with a partner and with the whole class, and then provided instructor feedback over 120 minutes. Specifically, PSTs were encouraged to think about the following questions in writing contextual tasks: "What is the referent unit? What is the "whole" amount that is the size of the group being multiplied? How much of what is it you're trying to find (__ of __)?"

Data and Analysis

For this study, we are focusing on two of the four assessments, pre- and post-assessments, administered throughout the instructional sequence approximately one month apart. Assessments were identical and instructed PSTs to: "Create a fraction multiplication story problem that requires students to find ½ of ¾ ."

We used deductive followed by inductive coding (Saldaña, 2009) to develop a codebook analyzing PSTs' multiplicative reasoning with fraction multiplication as demonstrated through their writing of a contextual task. Deductive codes came from literature informing multiplicative reasoning with fractions (e.g., Son & Lee, 2016), whereas the inductive codes were informed by iterative revision based on the data. Contextual representations were classified based on the following components: underlying context, whether the context was discrete (counts) or continuous (e.g., area), the framing of and relation to the referent unit, the question posed, how the unknown was framed, and the operations required to determine the unknown when solving the problem. We then categorized PSTs' multiplicative reasoning as demonstrated through contextual representations into five categories: Integrated, Partially Integrated, Transitional, Foundation, and Uninterpretable. Three MTEs coded each pre- and post-assessment based on the integration of multiplicative reasoning category with an initial interrater agreement rate of 81.4%, and all disagreements were resolved based on collective consensus. This process of consensus coding in particular helped us to develop this codebook into categories that, while similar, became distinct through revision and iterative coding. Each category had required or possible elements to achieve the corresponding level of integration (see Table 1).

To identify any changes from pre- to post-assessments, we compared overall shifts for each institution by looking at counts and percentages of PSTs' categories of integration. To examine changes across all participants, we assigned numerical values to each category (see Table 2 for numerical equivalents of each category) and calculated the difference between post and pre scores. We determined net shifts and then found the mean number of shifts per PST for each institution. We then selected two PSTs' pre- and post-assessments, one from each institution, that depicted varying levels of integration and varying degrees of change from pre- to post-assessments with representative shifts for each institution.

Table 1

Codebook for Shifts in Multiplicative Reasoning as Demonstrated in Contextual Representations

Integration	Descriptor
Fully Integrated (4)	• Context represents multiplicative reasoning and is meaningful (required) • Mathematically correct stories depict the given symbolic statement (required) • Multiplicand is interpreted as the referent unit (required) • Clear and correct reference whole • The unknown in the story clearly asks for fraction multiplication (required)
Partially Integrated (3)	• Context represents multiplicative reasoning OR ○ Context represents multiplicative reasoning but context may not be meaningful (required) • Mathematically correct stories, but they alter the multiplicand and multiplier in the given symbolic statement (possible) • Multiplier is interpreted as the referent unit (required) • Clear and correct reference whole (required) • The relationship between the multiplicand and whole is implicit (possible) • The unknown in the story clearly asks for fraction multiplication (required) • Question asked does not explicitly imply fractional reasoning in relation to the original whole (implied but would not be clear if asked to a student) (possible)
Transitional (2)	• Context does not consistently represent multiplicative reasoning (e.g., subtraction, addition, and/or division are represented) (possible) • Mathematically correct stories, but they do not represent the given symbolic statements (possible) • Referent units are not correctly used (e.g., counts used) (possible) • Clear and correct referent whole in the context and question asked (possible)
Foundational (1)	• Context does not represent multiplicative reasoning (possible) • Stories are mathematically incorrect and/or require clarification because one or more of the following occur: operations are unclear or mixed, wholes are not specified, referent units are misused, or the unknown in the story problem is not clearly defined or not present (possible) • Unclear or incorrect referent whole used in either the context or question asked (possible)
Uninterpretable	• Did not write a story problem

Results and Discussion

We first report overall shifts comparing the institution with solving instruction (Institution A) versus the institution with instruction for solving and writing fraction multiplication representations (Institution B). Table 2 demonstrates the shifts in multiplicative reasoning in written contextual representations from pre to post.

Table 2

Shifts in Writing Contextual Representations Requiring Multiplicative Reasoning

Level of Integration	Institution A Solving Instruction (n=18)		Institution B Solving and Writing Instruction (n=17)	
	Pre	Post	Pre	Post
Fully Integrated	1 (5.6%)	1 (5.6 %)	2 (11.8%)	9 (52.9%)
Partially Integrated	2 (11.1%)	8 (44.4%)	1 (5.9%)	2 (11.8%)
Transitional	1 (5.6%)	4 (22.2%)	4 (23.5%)	1 (5.9%)
Foundational	12 (66.7 %)	5 (27.8%)	8 (47.1%)	5 (29.4%)
Uninterpretable	2 (11.1%)	0 (0%)	2 (11.8%)	0 (0%)

Institution A saw moderate shifts showing possible growth from exposure to solving contextual representations alone: Uninterpretable decreased from 2 (11.1%) to 0, Foundational decreased from 12 (66.7%) to 5 (27.8%), Transitional increased from 1 (5.6%) to 4 (22.2%), Partially Integrated increased from 2 (11.1%) to 8 (44.4%), and Fully Integrated remained unchanged at 1 (5.6%). Perhaps expectedly, the Institution B showed comparatively greater growth: Uninterpretable decreased from 2 (11.8%) to 0, Foundational decreased from 8 (47.1%) to 5 (29.4%), Transitional decreased from 4 (23.5%) to 1 (5.9%), Partially Integrated increased from 1 (5.9%) to 2 (11.8%), and Fully Integrated increased from 2 (11.8%) to 9 (52.9%). However, looking at shifts in aggregate does not clearly capture the growth. To do so, we counted net shifts and found the mean number of shifts per PST for each institution. Institution A's PSTs averaged 0.944 categorical shifts per PST, while Institution B averaged 1.293 shifts per PST. When controlling for the two PSTs writing an Integrated contextual representation for both the pre- and post-assessments, Institution B averaged 1.466 shifts per PST.

Overall, it is encouraging to see the Uninterpretable or Foundational written representations decreasing in both institutions. While Institution A shifted primarily to the Transitional and Partially Integrated categories, Institution B shifted overwhelmingly to the Fully Integrated category. These shifts indicate a qualitative difference in the fluency demonstrated by PSTs when constructing word problems to answer the research question, "Does specific focus on constructing word problems

result in qualitatively different outcomes?" as measured by the development of our coding schema. While exposure to and solving contextual fraction multiplication representations supported PSTs in writing more multiplicatively integrated contextual representations at Institution A, the focus on referent unit, the whole amount of the group being multiplied, and the relationship between multiplier and multiplicand for Institution B supported PSTs in the details of writing fully integrated representations. The prompts and discourse used by Institution B's MTE may support the high demands of SCK for fraction multiplication (Luo, 2009). The exemplars that follow show these changes for a typical PST from both institutions.

To answer the research question, "How do PSTs' contextual representations reveal their fraction multiplication reasoning?," we present two exemplars that illustrate the most common shift in multiplicative reasoning for each institution. For Institution B, PSTs' contextual representations tended to shift from Foundational to Fully Integrated. For the pre-test, PST 'B' wrote: "A student is making cookies and wants to give away 3/4ths of them. Then she decides to give ½ to her family. How many of the total cookies did she give her family?" Here, the PST specifies the referent unit as a fraction, ¾, of an original unspecified amount x, leading to a possible referent unit of $\frac{3}{4}x$. They then change the multiplier to either be half of the original amount, $\frac{1}{2}x$, or half of the ¾ the student wants to give away, $\frac{1}{2} \times (\frac{3}{4}x)$. The question, then, is phrased as a count (how many) rather than a fraction of a fraction (how much of or what fraction of). Because of the uncertainty a student would have in discerning the unit being multiplied, and phrasing inferring subtraction (e.g., give away), this PST wrote a Foundational contextual representation. On the post-test, they wrote "Bob is designing a park and wants ¾ of the park to be sports fields. Of the sports field area, he wants ½ of it to be soccer fields. How much of the total park will be soccer fields?" Here the referent unit or multiplicand is clear (¾)), the multiplier is clear ($\frac{1}{2} \times \frac{3}{4}$), and the question asks about the fractional product of the original whole, a Fully Integrated contextual representation.

For Institution A, the most frequent shift was Foundational to a Partially Integrated. For the pre-test, PST 'A' wrote: "I want to plant cucumbers in ¾ of my garden. ½ of that will be pickled cucumbers." While the PST introduces the referent unit, ¾, and a possible multiplier as a fractional amount within the referent unit ½, they do not ask a question which would require finding the product. For the post-test they wrote: "I had ¾ of my cake left when I went to bed. When I woke up, I saw Tim ate ½ of the leftover cake. How much cake did Tim eat?" Here, the PST identifies the referent unit clearly, ¾, an action to find the fraction of a fraction (i.e., Tim eating ½ of the leftover cake), but the question requires students to infer that the task is asking how much of the original

cake did Tim eat, thus, the relation to the whole is implied rather than explicit. Because that may not be clear to a student engaging in that task, this contextual representation represents Partially Integrated multiplicative reasoning on the part of the PST writing it.

Overall, we found that our codebook supported us in systematically unpacking PSTs' multiplicative reasoning in regard to the referent unit, relation between multiplier and multiplicand, and phrasing questions in relation to the original whole. The balance between using predetermined categories and emergent descriptors was useful to inform how we assess PSTs' SCK as MTEs.

Implications

Our findings align with some of the prior research that shows how attending specifically to conceptual features like identifying the referent unit, coordinating the relationship between the multiplicand and multiplier, and framing questions in relation to the original whole support the development of representing fraction multiplication (e.g., Izsak, 2008; Luo, 2009; Son & Lee, 2016). Although engaging PSTs in solving contextual fraction multiplication tasks can promote representational fluency to some extent, our findings suggest that instruction should explicitly foreground these underlying conceptual features. Attending to these conceptual features may help strengthen PSTs' ability to construct mathematically meaningful contextual representations of fraction multiplication rather than just producing story problems that correspond procedurally to symbolic expressions. These findings have implications for the design of instructional activities in mathematics methods or content courses. In particular, MTEs can consider explicitly scaffolding PSTs' attention to fraction multiplicative reasoning through instructional prompts and opportunities to compare and discuss with peers, similar to instructional moves at Institution B. Specifically, MTEs can draw PSTs' attention to referent units, multiplicative relationships, and the relation between the product and the original whole when constructing contextual representations.

The findings of this study must be foregrounded in limitations, and among them are the small number of PSTs in each instructional condition (18 and 17 respectively), in addition to a confounded comparison with two institutions, instructors, and course placements. More specifically, the prompts and additional activities meant to support writing contextual representations in Institution B were given to PSTs in a methods course versus a first-semester math content course in Institution A.

This study shows the value of examining SCK by looking at how PSTs demonstrate mathematical reasoning in their representations rather than relying solely on broad distinctions such

as procedural versus conceptual knowledge. Our coding framework captured meaningful differences in how PSTs coordinated units, operations, and quantitative relationships when constructing contextual tasks. These distinctions reveal aspects of SCK that may remain hidden when assessments focus only on symbolic computation or correct answers. Analyzing the mathematical structure embedded within PSTs' representations offers a more nuanced understanding of how they reason about fraction multiplication and where instructional support is needed. Our findings also provide a foundation for future research examining how different representations reveal and support PSTs' SCK. After examining PSTs' contextual representations, we became interested in exploring how instructional differences influenced their visual representations. Thus, our subsequent work investigates the relationship between PSTs' visual and written contextual representations.

References

Acevedo Nistal, A., Van Dooren, W., Clarebout, G., Elen, J., & Verschaffel, L. (2009). Conceptualising, investigating and stimulating representational flexibility in mathematical problem solving and learning: A critical review. *ZDM, 41*(5), 627–636. https://doi.org/10.1007/s11858-009-0189-1

Aksoy, N. C., & Yazlik, D. O. (2017). Student errors in fractions and possible causes of these errors. *Journal of Education and Training Studies*, *5*(11), 219–233. https://eric.ed.gov/?id=EJ1161215

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special?. *Journal of Teacher Education*, *59*(5), 389–407. https://doi.org/10.1177/0022487108324554

Izsák, A. (2008). Mathematical knowledge for teaching fraction multiplication. *Cognition and Instruction, 26*(1), 95–143. https://doi.org/10.1080/07370000701798529

Kim, J. (2025). How representations support young students' conceptual understanding of multiplication? A study on exploring third-grade students' representational reasoning. *Asian Journal for Mathematics Education, 4*(3), 311–331. https://doi.org/10.1177/27527263251335686

Luo, F. (2009). Evaluating the effectiveness and insights of preservice elementary teachers' abilities to construct word problems for fraction multiplication. *Journal of Mathematics Education*, *2*(1), 83–98.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*.

Park, J., Güçler, B., & McCrory, R. (2013). Teaching prospective teachers about fractions: Historical and pedagogical perspectives. *Educational Studies in Mathematics, 82*(3), 455–479. https://doi.org/10.1007/s10649-012-9440-8

Saldaña, J. (2009). *The coding manual for qualitative researchers.* Sage.

Sapkota, B. (2025). Secondary preservice teachers' mathematical knowledge for teaching: Affordances and limitations of peer teaching and reflection. *International Journal of Science and Mathematics Education, 23*, 2349–2373. https://doi.org/10.1007/s10763-025-10555-w

Sari, I. P., Suryadi, D., Herman, T., Dahlan, J. A., & Supriyadi, E. (2024). Learning obstacles on fractions: A scoping review. *Infinity Journal*, *13*(2), 377–392. https://doi.org/10.22460/infinity.v13i2.p377-392

Son, J. W., & Lee, J. E. (2016). Pre-service teachers' understanding of fraction multiplication, representational knowledge, and computational skills. *Mathematics Teacher Education and Development, 18*(2), 5–28. https://eric.ed.gov/?id=EJ1113959

EXPLORING FACTORS THAT SUPPORT STUDENTS' LEARNING IN A SOCIAL JUSTICE MATHEMATICS CLASSROOM

Waleed A. Raja
University of South Carolina
wraja@mailbox.sc.edu

Nicole L. Fonger
Syracuse University
nfonger@syr.edu

Anis Munfarikhatin
Syracuse University
amunfari@syr.edu

Abstract

Social justice mathematics positions mathematics as a tool for understanding and addressing real-world inequities while supporting meaningful student learning (Gutstein, 2003; Kokka, 2022). This practitioner paper synthesizes research to present a framework of five interrelated factors that support student learning in social justice mathematics classrooms: (1) co-creating classrooms, (2) engaging students in problem-solving, investigative research, and collaborative work, (3) implementing curriculum centered on social justice issues, (4) attending to students' emotions, and (5) support from the school. Together, these factors provide educators with practical guidance for designing equitable mathematics learning environments that foster mathematical understanding, student engagement, and critical consciousness.

Keywords: Social justice mathematics; student learning; mathematics education; equitable teaching

Introduction

Mathematics education researchers have long argued for changing how mathematics is taught in classrooms (Bartell, 2013; Battey et al., 2022; Gonzalez, 2009; Gutierrez, 2002; Gutstein, 2003, 2006; Gutstein & Peterson, 2013; Guzmán & Craig, 2019). Gutstein and Peterson (2013) described mathematics as a tool "to help us understand and potentially change the world" (p. 5), noting that it is too often "taught in ways divorced from the real world" (p. 1). Moreover, students perceive mathematics as more relevant when it is taught through activities grounded in social issues (Wright, 2016).

Social justice mathematics (SJM) responds to this call. Gutstein (2003) defined SJM as mathematics education grounded in a vision of social justice that challenges systems of oppression and promotes social change, using real-world examples that connect mathematics to social issues and empower students to work toward change. As mathematics education evolves toward more

equitable, justice-oriented approaches, there is a growing need to understand the conditions that enable meaningful student engagement and learning in these classrooms.

Purpose

This paper explores key factors that support student learning (SSL) in SJM classrooms. Drawing on a comprehensive review of literature in SJM education, we identified five interrelated SSL factors: (1) co-creating classrooms, (2) engaging students in problem-solving, investigative research, and collaborative work, (3) implementing curriculum that centers social justice issues, (4) attending to students' emotions, and (5) support from the school. These factors are not content-specific and may support student learning in other disciplines as well. Together, they offer an instructional framework that promotes academic learning and critical consciousness.

We identified the sources for the literature review by searching Google Scholar and multiple research guides and databases, including JSTOR, ERIC (EBSCO), and Taylor & Francis Online for "social justice mathematics," "social justice pedagogies," "critical mathematics," "teaching mathematics for social justice," and "teaching and learning mathematics for social justice." The search produced an initial list of articles on social justice mathematics. We reviewed abstracts to identify studies that discussed the theoretical framework of social justice mathematics and/or its classroom implementation. We compared and coded selected studies for reported practices, conditions, and outcomes. We grouped recurring practices into the five SSL factors.

Related Literature Grounding SSL Factors

This paper is grounded in Kokka's (2022) social justice mathematics framework, which organizes pedagogical goals into three dimensions: dominant (developing mathematical power and succeeding on conventional measures), critical (using mathematics to understand and challenge relations of power and inequity), and affective (helping students identify and process emotions related to mathematics and injustice) (Kokka, 2022). The five SSL factors elaborated next map onto and support these three dimensions of pedagogical goals for SJM.

The first SSL factor is *co-creating classrooms*, which refers to sharing authority with students by actively involving them in shaping the learning environment and instructional decisions. This approach encourages students to help establish classroom norms, select social issues for investigation, and participate in decisions that shape their learning. Research in SJM suggests that when students' voices and experiences are valued, they are more willing to engage in discussions about complex social issues and participate meaningfully in mathematical inquiry (Frankenstein,

1983; Gutstein, 2003, 2007; Tate, 1995; Turner et al., 2009; Voss & Rickards, 2016; Wright, 2016). Sharing authority with students supports learning in two ways: creating a comfortable classroom environment (Gutstein, 2003) and involving students in selecting the social issues to be studied (Gutstein, 2007; Tate, 1995; Turner et al., 2009; Wright, 2016). This shift in power dynamics helps normalize discussion of difficult topics and positions students' voices and lived experiences as central to learning (Gutstein, 2003). Gutstein (2007) argued that involving students in selecting issues grounded in their own lives supported the development of social agency by helping them connect mathematical knowledge to their communities. Gutstein (2003) argued that students developed more critical awareness of social issues, came to see mathematics as a tool for analyzing and questioning the world, and generally maintained or improved their mathematical performance. Tate (1995) described how students identified problems affecting their community, investigated those issues, and developed strategies for social action, illustrating how shared decision-making can promote both learning and agency. Integrating students' interests helped students recognize the relevance of mathematics in their everyday lives (Turner et al., 2009), supported the development of social agency (Wright, 2016), and improved their engagement and academic achievement (Voss& Rickards, 2016). Although the literature highlights the benefits of co-creating classrooms, it also suggests that balancing student autonomy with productive classroom learning requires careful facilitation. Gregson (2013) reported that providing students with greater decision-making authority sometimes resulted in off-task discussions and challenges in maintaining productive learning. These findings suggest that co-creating classrooms is not simply about transferring authority to students but about developing shared responsibility that supports respectful dialogue, meaningful participation, and sustained mathematical inquiry.

The second SSL factor is *engaging students in problem-solving, investigative research, and collaborative work*. Research suggests that students' learning is facilitated when teachers use problem-solving approaches (Frankenstein, 1990; Wright, 2016), engage students in investigative research (DePietto, 2026; Gutstein, 2003, 2007; Tate, 1995), and encourage collaboration (Avcı, 2020; DePietto, 2026; Tate, 1995; Wright, 2016). This approach positions students as active participants who investigate authentic social issues by asking questions, gathering information, evaluating evidence, and collaboratively developing mathematical solutions. Together, these instructional practices encourage students to construct mathematical knowledge while developing the capacity to analyze real-world problems (DePietto, 2026; Frankenstein, 1990; Gutstein, 2003, 2007; Tate, 1995; Wright, 2016).

The third SSL factor is *implementing curriculum centered on social justice issues*. Curriculum that situates mathematics within social justice contexts enables students to investigate real-world issues, analyze data related to inequities, and use mathematics to understand societal structures (Berry et al., 2020; DePietto, 2026; Frankenstein, 1990; Gutstein, 2003). Social justice curriculum situates mathematical concepts within authentic social contexts, allowing students to apply mathematical reasoning to issues that affect their communities. Such curriculum encourages students to justify their thinking, explore multiple solution strategies, and recognize mathematics as a meaningful tool for understanding social inequities (Frankenstein, 1990; Gutstein, 2003, 2006). Berry et al. (2020) offered guidance on designing lessons driven by students' interests and challenges and on creating social justice math lessons.

The fourth SSL factor is *attending to students' emotions*. Investigating injustice can evoke strong emotional responses (Kokka, 2022). Teachers must create a safe environment where students can express emotions and be supported through healing-informed approaches (Kokka, 2019). Research suggests that when students' emotions are not attended to, they can become disengaged from the mathematics lesson (Battey et al., 2022). Students' emotions play a key role in mathematical learning (DePietto, 2026; Schukajlow et al., 2017) and, when carefully managed, can support students' development as active members of society who use mathematics to address inequities (DePietto, 2026; Gutstein, 2006; Kokka, 2019, 2020, 2022). Creating a classroom environment where students feel comfortable expressing their emotions and where teachers intentionally respond to those emotions can support both students' well-being and their mathematical learning (Battey et al., 2022; Kokka, 2019, 2022). This connects to Kokka's (2022) theorized affective pedagogical goals as a central dimension of SJM. These goals emphasize creating relationships that support students' emotional well-being, helping students identify emotions related to mathematics and social injustice, and supporting them in processing these emotions as they develop social agency.

The fifth SSL factor is *support from the school*. While teachers play a central role in implementing SJM, sustained instructional change is strengthened when schools provide the conditions necessary for justice-oriented teaching. School support includes establishing justice-focused learning goals, providing curricular autonomy, allocating resources, and creating opportunities for professional learning that enable teachers to implement meaningful social justice investigations (Kokka, 2020). Teachers are better able to sustain SJM instruction when their school provides administrative backing, explicit justice-oriented teaching standards, funding for resources and professional development, and curricular autonomy (Kokka, 2020). These forms of institutional

support enable teachers to implement sustained SJM instruction and provide students with opportunities to examine issues of inequity while developing mathematical understanding.

SJM in Practice

Translating these five factors into practice gives mathematics teachers a concrete starting point for designing SJM instruction. To co-create the classroom, teachers can invite students into decisions about classroom norms and setup, and ask students what social issues they would like to explore, using their input as a starting point for curriculum (Frankenstein, 1983). By sharing authority with students and encouraging dialogue around topics often considered taboo, the classroom can become a space where students connect mathematics to their lived experiences and develop a stronger sense of justice and social agency. Teachers can engage students in problem-solving and investigative research by centering instruction on student-generated questions, pairing problem-solving with reflection on solution methods (Frankenstein, 1990), and supporting students in researching community issues (Tate, 1995) while working collaboratively (Avcı, 2020). Incorporating open-ended problem-creating and problem-solving activities helps students recognize their existing mathematical knowledge and develop greater confidence in their mathematical abilities (Frankenstein, 1990). Teachers can focus on encouraging students to raise questions rather than simply answering teacher-generated questions.

Teachers can combine existing curricula with real-world projects situated in a relevant social issue to develop curriculum that supports SJM, ask students to justify their explanations and consider multiple perspectives (Gutstein, 2003). Teachers can build in structured opportunities, such as reflective prompts, for students to identify and regulate their emotions (Kokka, 2019). This can help teachers attend to students' emotions. Advocacy for justice-oriented teaching standards, dedicated planning time, and funding for resources and professional development (Kokka, 2020) can help teachers secure support from the school.

Classroom Examples of SSL in SJM

In this section, the paper shares classroom examples, drawn from the studies reviewed while developing the SSL framework, that illustrate how the five factors have been enacted in practice. In Gutstein's (2003) study, students conducted investigations to examine changes in their own neighborhood by using ratio and proportion to analyze the effect of housing costs and gentrification pressures on the families they knew personally. The students raised issues of housing costs and gentrification, which led to a more open discussion of sensitive topics than might have occurred

under a teacher's learning agenda. Gutstein (2003) shared authority with students and encouraged dialogue around topics often considered taboo. This created a classroom that became a space where students connected mathematics to their lived experiences and developed a stronger sense of justice.

Frankenstein (1990) built a critical mathematical literacy curriculum around problem-posing rather than solving problems from textbooks. Students used statistics they found by themselves meaningful to write their own mathematical questions, then worked in groups to gather supporting information, weigh the credibility of their sources, and solve the problems they had posed. This setting positioned students as producers rather than consumers of mathematical questions. Similarly, Tate (1995) shared how students investigated issues affecting their community by researching local laws, gathering evidence, and collaboratively developing strategies to address a neighborhood concern. Throughout the investigation, students' questions guided the learning process, illustrating the importance of inquiry and collaborative problem-solving in supporting both mathematical learning and social agency. In another example, Avcı (2020) found that students benefited academically from working in groups and that collaborative investigations helped build a stronger classroom community. Students engaged with one another and moved beyond initially working with peers from similar backgrounds and increasingly collaborated across diverse groups. Likewise, Wright (2016) reported that teachers perceived problem-solving approaches and collaborative group work as supporting the development of students' social agency. Similarly, Gutstein (2006) argued that while no curriculum completely aligns with the goals of SJM, existing curricula can be adapted to support justice-oriented instruction. In his earlier work, Gutstein (2003) integrated the Mathematics in Context curriculum with real-world projects and found that students developed multiple solution strategies, reasoned mathematically, communicated their thinking, and viewed mathematics as relevant to their lives. Although the curriculum was implemented alongside authentic investigations, making it difficult to separate the contribution of each component, the findings suggest that the curriculum provides an important foundation for meaningful mathematical inquiry. Gutstein (2003) incorporated social justice directly into required mathematics curriculum content. In one unit, students applied their understanding of area and scale to show that the commonly used Mercator map projection distorts the relative size of continents, enlarging Europe while shrinking Africa. In another, students used unit rates and proportional reasoning to calculate that the cost of a single military aircraft could fund four years of college scholarships for their entire graduating class by using standard mathematical content to reveal underlying societal priorities.

The lesson practice from Kokka (2019) focused on issues like food deserts, financial realities of single-parent households, and living on minimum wage. Students identified and discussed the emotions the topics evoked, including frustration and empathy, and the teacher's role was to facilitate students' discussion by creating a supportive environment; thus, students could engage with these issues without feeling re-traumatized. According to Kokka (2020), sustained SJM instruction depended not only on individual teacher commitment but also on institutional support. School administrators established justice-oriented teaching standards; evaluated teachers against those standards; funded instructional resources and teachers' participation in conferences; and provided additional planning time and curricular flexibility beyond standardized testing requirements. These organizational conditions enabled teachers to implement and sustain justice-oriented pedagogy over time.

The SSL factors are intertwined. Some of the examples above demonstrate how multiple SSL factors are interrelated and work together to support students' learning. For example, in Gutstein's (2003) work, we can see *co-creating classrooms, engaging students in problem-solving, investigative research, and collaborative work,* and *implementing curriculum centered on social justice issues* at play. However, teachers can implement one SSL factor at a time, as demonstrated by Avcı's (2020) example about *engaging students in problem-solving, investigative research, and collaborative work.*

Implications

This review has implications for both practice and research. For practitioners, the five SSL factors offer a practical framework for designing SJM instruction: sharing authority with students, embedding collaborative and investigative work, selecting curriculum around relevant social issues, attending to students' emotions, and securing school-level support. Because these factors are not content-specific, they may also guide teachers in other disciplines.

The strength of evidence varies across factors. Co-creating classrooms, problem-solving and group work, and curriculum are relatively well documented, while attending to students' emotions and securing school support remain newer areas needing further research within SJM contexts (Kokka, 2022). These five factors point to the interdependence of classroom-level practice and school-level conditions: attending to only some of them may limit how far SJM instruction can promote both academic learning and students' critical consciousness.

References

Avci, B. (2020). Collaborative learning within critical mathematics education. *Beta: Jurnal Tadris Matematika, 13*(1), 1–13. https://doi.org/10.20414/betajtm.v13i1.366

Bartell, T. G. (2013). Learning to teach mathematics for social justice: Negotiating social justice and mathematical goals. *Journal for Research in Mathematics Education, 44*(1), 129–163.

Battey, D., Amman, K., Leyva, L. A., Hyland, N., & McMichael, E. W. (2022). Racialized and gendered labor in students' responses to precalculus and calculus instruction. *Journal for Research in Mathematics Education, 53*(2), 94–113. https://doi.org/10.5951/jresematheduc-2020-0170

Berry, R. Q., III, Conway, B. M., IV, Lawler, B. R., & Staley, J. W. (2020). *High school mathematics lessons to explore, understand, and respond to social injustice.* Corwin.

DePietto, M. (2026, Summer). Hoodies, identity, and algebra: Mathematics as student advocacy. *Rethinking Schools, 40*(4). https://rethinkingschools.org/articles/hoodies-identity-and-algebra/

Frankenstein, M. (1983). Critical mathematics education: An application of Paulo Freire's epistemology. *Journal of Education, 165*(4), 315–339. https://doi.org/10.1177/002205748316500403

Frankenstein, M. (1990). Incorporating race, gender, and class issues into a critical mathematical literacy curriculum. *The Journal of Negro Education, 59*(3), 336–347. https://doi.org/10.2307/2295568

Gonzalez, L. (2009). Teaching mathematics for social justice: Reflections on a community of practice for urban high school mathematics teachers. *Journal of Urban Mathematics Education, 2*(1), 22–51. https://doi.org/10.21423/jume-v2i1a32

Gregson, S. A. (2013). Negotiating social justice teaching: One full-time teacher's practice viewed from the trenches. *Journal for Research in Mathematics Education, 44*(1), 164–198. https://doi.org/10.5951/jresematheduc.44.1.0164

Gutierrez, R. (2002). Enabling the practice of mathematics teachers in context: Toward a new equity research agenda. *Mathematical Thinking and Learning, 4*(2–3), 145–187. https://doi.org/10.1207/s15327833mtl04023_4

Gutstein, E. (2003). Teaching and learning mathematics for social justice in an urban, Latino school. *Journal for Research in Mathematics Education, 34*(1), 37–73. https://doi.org/10.2307/30034699

Gutstein, E. (2006). *Reading and writing the world with mathematics: Toward a pedagogy for social justice.* Taylor & Francis.

Gutstein, E. (2007). "And that's just how it starts": Teaching mathematics and developing student agency. *Teachers College Record, 109*(2), 420–448. https://doi.org/10.1177/016146810710900203

Gutstein, E., & Peterson, B. (2013). *Rethinking mathematics: Teaching social justice by the numbers.* Rethinking Schools.

Guzmán, L. D., & Craig, J. (2019). The world in your pocket: Digital media as invitations for transdisciplinary inquiry in mathematics classrooms. *Occasional Paper Series, 2019*(41). https://doi.org/10.58295/2375-3668.1274

Kokka, K. (2019). Healing-informed social justice mathematics: Promoting students' sociopolitical consciousness and well-being in mathematics class. *Urban Education, 54*(9), 1179–1209. https://doi.org/10.1177/0042085918806947

Kokka, K. (2020). Social justice pedagogy for whom? Developing privileged students' critical mathematics consciousness. *The Urban Review, 52*(4), 778–803. https://doi.org/10.1007/s11256-020-00578-8

Kokka, K. (2022). Toward a theory of affective pedagogical goals for social justice mathematics. *Journal for Research in Mathematics Education, 53*(2), 133–153. https://pubs.nctm.org/view/journals/jrme/53/2/article-p133.xml

Schukajlow, S., Rakoczy, K., & Pekrun, R. (2017). Emotions and motivation in mathematics education: Theoretical considerations and empirical contributions. *ZDM Mathematics Education, 49*(3), 307–322. https://doi.org/10.1007/s11858-017-0864-6

Tate, W. F. (1995). Returning to the root: A culturally relevant approach to mathematics pedagogy. *Theory Into Practice, 34*(3), 166–173. https://doi.org/10.1080/00405849509543676

Turner, E. E., Gutiérrez, M. V., Simic-Muller, K., & Díez-Palomar, J. (2009). "Everything is math in the whole world": Integrating critical and community knowledge in authentic mathematical investigations with elementary Latina/o students. *Mathematical Thinking and Learning, 11*(3), 136–157. https://doi.org/10.1080/10986060903013382

Voss, R., & Rickards, T. (2016). Challenges related to teaching mathematics using social justice pedagogies: A secondary school experience. *Journal of Education and Practice, 7*(17), 68–73. https://eric.ed.gov/?id=EJ1105307

Wright, P. (2016). Social justice in the mathematics classroom. *London Review of Education, 14*(2), 104–118. https://doi.org/10.18546/lre.14.2.07

REPRESENTATIONS OF MATHEMATICAL PROOF ACROSS MAJOR U.S. STANDARDS DOCUMENTS

Seyedehkhadijeh Azimi Asmaroud
Virginia State University
sazimi@vsu.edu

Somya N Thakkar
Virginia State University
stha8420@students.vsu.edu

Abstract

This study presents a curriculum analysis examining how mathematical proof is represented across three major standards documents: National Council of Teachers of Mathematics (NCTM), the Virginia Department of Education Standards of Learning (VSOL), and the Common Core State Standards Mathematics (CCSSM). Research suggests that many students encounter proof primarily in a single high school course, which may limit their opportunities to develop sustained reasoning and argumentation skills (e.g., Abbaspour Tazehkand et al., 2025). In addition to direct mentions of proof, the study examined related terms: Conjecture, Justification, Argument, Validation, Explain, Reasoning, Generalization, and Verify. The findings in this study showed that the ratio of the word "proof" used in the VSOL was lower than that of the NCTM (2000) standards but higher than that of the CCSSM (2010). However, the ratios of the total number of proof-related terminology across all three documents indicated a consistent emphasis on proof-related terminology across the three standards.

Keywords: proof, standards, CCSSM, NCTM, Virginia SOLs

Introduction

Proof is a core aspect of mathematics and serves multiple purposes in mathematics (De Villiers, 1999). Despite its foundational role, existing research consistently highlights that proof is often central in theory yet receives limited emphasis in school curricula and practices (Abbaspour Tazehkand et al., 2024; Knuth et al., 2019). Major standards and frameworks, such as CCSSM (2010), address proof primarily within high school geometry, and reasoning-focused initiatives like *Principles to Action* provide little explicit guidance for teachers on how to develop proof skills in students (Abbaspour et al., 2024). Furthermore, several scholars argue that proof should instead be treated as a continuous K–12 strand, beginning with informal justifications and reasoning in the early grades and gradually progressing toward formal proof (e.g., Knuth et al., 2019; Zaslavsky et al., 2012).

This study is part of a broader research project that investigates the representation of proof and proof-related concepts across standards documents, to better understand how they support the development of students' reasoning and proving competencies. In this paper, we specifically report the presence and distribution of explicit references to proof and proof-related terms. The research question is: How do the VSOLs, CCSSM (2010), and NCTM (2000) standards differ in their representation of mathematical proof and reasoning practices across K–12 education?

Literature Review

Proof is a core aspect of mathematics and serves multiple purposes in mathematics (De Villiers, 1999; Rocha, 2019). Despite its critical role in "learning mathematics" (Knuth et al., 2019), existing research consistently indicates that proof often receives limited emphasis in school curricula and classroom practices (Abbaspour Tazehkand et al., 2024; Rocha, 2019) and students of all ages continue to experience difficulties in learning how to develop and understand mathematical proofs (Eldred, 2018).

Researchers in mathematics education argue that reasoning and proof should have a more prominent place in mathematics instruction across all grade levels (Knuth et al., 2019; Zaslavsky et al., 2012). Mathematics curriculum should engage students in authentic experiences with mathematical proof from the beginning of their mathematical education (Stylianides, 2016). NCTM (2000) also makes a strong argument for including proof and reasoning in K–12 mathematics education by mentioning that "Instructional programs from prekindergarten through grade 12 should enable all students to recognize reasoning and proof as fundamental aspects of mathematics" (p. 56). Proof cannot be effectively taught through a single unit on logic or solely by "doing proofs" in geometry (NCTM, 2000). Consequently, examining only explicit references to the word *proof* provides an incomplete picture of how standards support the development of proving. Therefore, this study examines both explicit references to proof and the occurrence of proof-related terminology to better understand how the standards promote the development of students' reasoning and proving competencies. In the next section, we identify the proof-related terms examined in this study.

Related Words to Proof and Reasoning

The literature review guided the selection of terms related to proof and proving for the document analysis. Its major roles include verification, explanation, justification, discovery, systematization, communication, and intellectual challenge (Cadwallader-Olsker, 2011; Rocha, 2019;

Zaslavsky et al., 2012). Argumentation, conjecturing, and mathematical reasoning are also closely connected to proof because arguments developed during problem solving can be organized into valid proofs (Stylianides et al., 2016). Also, NCTM (2000) emphasizes reasoning, conjectures, arguments, and various proof methods, while terms such as *show*, *justify*, *explain*, and *prove* represent different forms of proof in school mathematics (Tall et al., 2012).

Theoretical Perspective

Jerome Bruner played a major role in shaping modern views of education by promoting curriculum reform. Bruner introduced the idea of the spiral curriculum as part of his wider work on how children learn (Gibbs, 2014). Bruner's spiral curriculum suggests that mathematical ideas should be introduced gradually and revisited with increasing depth and complexity (Gibbs, 2014). Revisiting topics with increasing levels of complexity promotes not only deeper understanding but also mastery of those topics at more advanced levels. Therefore, examining proof-related terms across all grade levels is important because practices of these concepts progressively prepare students to engage in formal mathematical proof at higher levels. Research in mathematics education supports the spiral curriculum approach in learning proof by emphasizing that students should engage in meaningful proof experiences from the beginning of their mathematical education to develop deep mathematical understanding (Knuth et al., 2019; Stylianides, 2016). NCTM (2000) also mentioned that proof cannot be taught effectively through a single unit on logic or only through completing proofs in geometry courses.

Proof cannot be taught effectively through a single unit on logic or solely through completing proofs in geometry (NCTM, 2000). Consequently, examining only explicit references to the word *proof* provides an incomplete picture of how standards support the development of proving. Thus, the purpose of this study is to examine how the VSOL prepare students for the concept of proof by analyzing the frequency and approaches of reasoning and proof within the standards and comparing them with the CCSS and NCTM recommendations.

Methods

This study used a mixed-methods document analysis that combined quantitative counts and frequencies of proof-related terminology with a qualitative analysis of how proof and reasoning were positioned across grade levels and mathematical topics. The present study reports the findings of the key proof-related terms including Proof, Conjecture, Justification, Argument, Validation, Explain, Reasoning, Generalization, and Verify in three documents VSOLs, NCTM, and CCSSM. In

examining the frequency of words within the documents, we considered all variations of each word. For example, for the word "generalization," we included related forms such as "generalize," "generalized," and "generalization" in the analysis. In addition, since the total number of pages in each document were different, after counting the frequency of the words, we divided the frequency count by the total number of pages to determine the ratio. Both authors independently counted the words and then compared their results together. In cases where the counts differed, the words were recounted to verify the accuracy of the final totals. The qualitative component examined how proof is conceptually framed within the VSOLs, including the ways reasoning, justification, and argumentation are presented as part of mathematical learning.

Results

In this section, we first present the frequency of proof-related vocabulary across the three standards documents. We then provide a qualitative analysis examining how proof and reasoning practices are conceptually framed within each set of standards. Table 1 presents the frequency of the words and their corresponding ratio.

Table 1

Frequency of the Proof-related Words and their Corresponding Ratio

	Frequency of the words			Ratio of the frequency		
	NCTM	VSOL	CCSSM	NCTM	VSOL	CCSSM
Proof	127	16	5	0.31	0.13	0.06
Conjecture	163	1	3	0.42	0.01	0.03
Justification	77	119	14	0.19	0.99	0.16
Argument	135	5	34	0.34	0.04	0.38
Validation	24	12	8	0.07	0.10	0.09
Explain	133	55	63	0.33	0.46	0.70
Reasoning	386	85	111	0.96	0.71	1.23
Generalization	159	30	10	0.39	0.25	0.11
Verify	15	45	3	0.03	0.37	0.03
# of pages	402	120	90			

Overall, the NCTM standards contained the highest frequency of explicit references to *proof* (127 occurrences; ratio 0.31), substantially exceeding both the VSOL (16 occurrences; ratio 0.13) and the CCSSM (5 occurrences; ratio 0.06). Also, *reasoning* was the most frequently occurring term in both the NCTM standards and the CCSSM, appearing 386 and 111 times, respectively. In contrast, *justification* was the most frequently occurring term in the VSOL, with 119 occurrences. After accounting for document length, the CCSSM demonstrated the greatest emphasis on reasoning

(ratio 1.23), followed by the NCTM standards (ratio 0.96), and the word *justification* (ratio 0.99) was the most frequently occurring term in VSOL. The VSOL included a comparatively strong emphasis on *verify* (45 occurrences; ratio 0.37), whereas this term appeared only 15 times in the NCTM standards (ratio 0.03) and 3 times in the CCSSM (ratio 0.03). The VSOL placed the least emphasis on *argument* (5 occurrence, ratio 0.04) compared to other standards (135 (ratio 0.34) in NCTM and 34 (ratio 0.38) in CCSSM). To provide an overall measure of the presence of proof-related language in the standards, we calculated the total number of occurrences of all proof-related words. Table 2 summarizes these results.

Table 2

Total Frequency of the Proof-related Words and their Corresponding Ratio

	NCTM (2000)	VSOL (2023)	CCSSM (2010)
Total	1,219	368	251
Ratio	3.03	3.07	2.79
# of pages	402	120	90

As Table 2 shows, although the documents differ substantially in total number of pages, the ratios indicate a similar overall emphasis on proof-related terminology in all three standards documents. The VSOL had the highest ratio of proof-related terms per page (3.07), followed closely by the NCTM standards (3.03), and the CCSSM had a slightly lower ratio (2.79). In the VSOL, Highschool Geometry is the first course in which the vocabulary "proof" is introduced; the curriculum "similarity of triangles" and "congruent triangles" is the first instance in which the authors observed the word "proof" in the text. The course emphasizes the development of reasoning skills through the exploration of geometric relationships, including properties of geometric figures, trigonometric relationships, and mathematical proofs. Geometry is designed for students who have successfully completed Algebra I. In this course, students learn to construct mathematical proofs using definitions, postulates, and theorems. They develop the ability to prove and justify that two triangles are congruent or similar by applying deductive reasoning. Throughout the proof process, students are expected to integrate algebraic methods and coordinate geometry to support and verify their arguments. In CCSSM, the word *proof* first appears in Grade 8 within the geometry standard, Pythagorean Theorem, and is later emphasized further in high school geometry courses. The expectation is that students should be able to "explain a proof of the Pythagorean Theorem and its converse" (CCSSM, 2010, p. 56). In high school, students build on earlier geometry

knowledge by moving from informal reasoning to formal proof and using equations and algebraic manipulation to model and prove geometric relationships.

However, in the NCTM (2000), Reasoning and Proof is identified as a separate process standard that should be integrated across all grade levels and students are expected to develop mathematical reasoning beginning with their earliest experiences in mathematics (e.g., pattern creation in PreK–2 to make conjectures). The NCTM standards emphasize that, even in the PreK–2 grades, teachers should encourage reasoning by asking questions such as *"Why do you think that is true?"* to help students learn to support or refute their ideas using evidence. From the earliest school years, students should construct and communicate sound mathematical arguments.

Discussion

The results of our study showed that the ratio of the word *proof* used in the VSOL was lower than that of the NCTM (2000) but higher than that of the CCSSM (2010). A study by Abbaspour Tazehkand et al. (2025) on California, Missouri, and Massachusetts standards shows that California and Massachusetts mention "proof" most frequently compared to Missouri standards. The ratio of the word "proof" in our study for the VSOL (0.13) was even lower than in the Missouri standards (0.18) in Abbaspour Tazehkand et al. (2025) study. In addition, their study showed that proof-related words were concentrated at the high school level, with minimal references in grades K–8, which aligns with the findings of our study on the VSOL. Abbaspour Tazehkand et al.'s (2025) study also found that concepts such as "conjecture" were either barely mentioned or entirely absent in some state standards, consistent with our finding for the VSOL. Also, they showed that the ratio of the word "argument" in the California and Massachusetts standards was approximately 0.33 and 0.45, respectively, which is close to our findings for the NCTM (2000) and the CCSSM (2010), with ratios of 0.34 and 0.38, respectively. This ratio was 0.04 for the VSOL, based on our findings, and 0.03 for the Missouri standards, based on Abbaspour Tazehkand et al.'s (2025) study. Also, consistent with CadwalladerOlsker (2011) and Abbaspour Tazehkand et al. (2025), our findings for the VSOL indicate that students primarily encounter formal proof at the high school level. This may separate proving from constructing mathematical knowledge and may lead students to view proof as a procedural task used only when explicitly required (CadwalladerOlsker, 2011).

Mathematical discourse is an effective way to help students develop their understanding of proof and mathematical reasoning (Eldred, 2018). Through classroom discussions and argumentation, students learn to explain, justify, and support their mathematical ideas, which builds

the foundation for proof. Therefore, as suggested by Eldred (2018), teachers play an important role in creating meaningful opportunities for mathematical discourse. Future studies are needed to explore how explanation and argumentation can be more effectively incorporated into school mathematics to support students' reasoning and proof development. Building on the findings of the present study, our broader research agenda will extend this work through a qualitative analysis of curriculum standards to examine how opportunities for explanation and argumentation prepare students for learning formal proof in the upper grade levels.

Classroom resources, particularly textbooks, shape students' opportunities to learn proof and reasoning. Bieda et al. (2013) found that only 3.7% of tasks in seven U.S. mathematics textbooks for students ages 9–11 supported reasoning and proving, and standards-aligned textbooks provided more such opportunities. Similarly, Nirode and Boyd (2021) found that textbooks rarely required students to construct proofs from general conditional statements and instead provided substantial scaffolding. Therefore, examining only tasks labeled as "proof" may not fully capture students' opportunities to develop proof-related reasoning and proof construction, highlighting the need to investigate how textbooks and online materials align with standards.

References

Abbaspour Tazehkand, S., Harbaugh, A., Ragan, G., Sullivan, P., & Barnett, J. (2025). Proof in mathematics education: Lost in state standards. In R. M. Zbiek, X. Yao, A. McCloskey, & F. Arbaugh (Eds.), *Proceedings of the forty-seventh annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Pennsylvania State University*. https://doi.org/10.51272/pmena.47.2025

Abbaspour Tazehkand, S., Karimi, M., & Safi, F. (2024). Mathematical proof in standards and practices: A contradiction. In K. W. Kosko, J. Caniglia, S. Courtney, M. Zolfaghari, & G. A. Morris (Eds.), *Proceedings of the forty-sixth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Kent State University.

Bieda, K., Ji, X., Drwencke, J., & Picard, A. (2013). Reasoning-and-proving opportunities in elementary mathematics textbooks. *International Journal of Educational Research, 64*, 71–80. https://doi.org/10.1016/j.ijer.2013.06.005

CadwalladerOlsker, T. (2011). What do we mean by mathematical proof? *Journal of Humanistic Mathematics, 1*(1), 33–60. https://doi.org/10.5642/jhummath.201101.04

De Villiers, M. (1999). The role and function of proof with Sketchpad. In M. D. De Villiers (Ed.), Rethinking proof: with the Geometer's Sketchpad (pp. 3–10). *Key Curriculum Press*.

Eldred, B. A. (2018). Proof and discourse in mathematics: Teaching for competency. *Learning to Teach Language Arts, Mathematics, Science, and Social Studies Through Research and Practice*, *7*(1), 47–53.

Gibbs, B. C. (2014). Reconfiguring Bruner: Compressing the spiral curriculum. *Phi Delta Kappan*, *95*(7), 41–44. https://doi.org/10.1177/003172171409500710

Knuth, E., Zaslavsky, O., & Ellis, A. B. (2019). *The role and use of examples in learning to prove. The Journal of Mathematical Behavior, 53*, 256–262. https://doi.org/10.1016/j.jmathb.2017.06.002

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.*

Nirode, W., & Boyd, B. (2021). High school geometry textbooks' proving opportunities of Common Core theorems. *School Science and Mathematics, 121*(6) 345–356.

Rocha, H. (2019). Mathematical proof: From mathematics to school mathematics. *Philosophical Transactions of the Royal Society A, 377*(2140), Article 20180045. https://doi.org/10.1098/rsta.2018.0045

Stylianides, A. J. (2016). *Proving in the elementary mathematics classroom.* Oxford University Press.

Stylianides, A., Bieda, K., & Morselli, F. (2016). Proof and argumentation in mathematics education research. In Á. Gutiérrez, G. C. Leder, & P. Boero (Eds.), *The second handbook of research on the psychology of mathematics education* (pp. 315–351). https://doi.org/10.1007/978-94-6300-561-6_9

Tall, D., Yevdokimov, O., Koichu, B., Whiteley, W., Kondratieva, M., & Cheng, Y.-H. (2012). Cognitive development of proof. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education (The 19th ICMI Study, pp. 13–51)*. Springer.

Zaslavsky, O., Nickerson, S. D., Stylianides, A. J., Kidron, I., & Winicki-Landman, G. (2012). The need for proof and proving: Mathematical and pedagogical perspectives. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education: The 19th ICMI Study* (pp. 215–229). Springer Netherlands.

CHILDREN'S MAP DRAWING: ANALYZING VISUOSPATIAL REASONING ELEMENTS IN A SUMMER LITERACY PROGRAM

Maria Eloisa Nuguid
Purdue University
mnuguid@purdue.edu

Comfort Aje
Purdue University
caje@purdue.edu

Abigail Erskine
Purdue University
aerskin@purdue.edu

Guili Zhang
Purdue University
zhan4149@purdue.edu

Breanya Hogue
Purdue University
bhogue@purdue.edu

Signe Kastberg
Purdue University
skastber@purdue.edu

Abstract

Visuospatial reasoning pertains to navigating the dynamics of space. We investigated elementary learners' visuospatial reasoning through self-generated maps in an informal learning environment. Utilizing a visuospatial framework, we analyzed learner-created maps drawn by K–5 children participating in a midwestern CDF Freedom School® summer literacy program to explore learners' geometric representation, spatial capabilities, symmetry thinking, and proportional reasoning. Findings revealed visuospatial elements in concert with personal narratives about participants' lived experiences; learners employed top-down and front-view perspectives, symmetry, and intuitive scaling to represent meaningful locations. This study contributes to the growing visuospatial reasoning scholarship in elementary mathematics education.

Keywords: Visuospatial reasoning, mapping, literacy

Introduction

Visuospatial reasoning enables learners to interpret spatial relationships, visualize transformations, and connect abstract geometric ideas to real-world contexts (Battista, 2007; Owens, 2020). In particular, mapping tasks provide a lens into these processes, as children use drawing to externalize how they understand and navigate space (Cohrssen & Pearn, 2021; Richardson et al., 2024; Suazo-Flores et al., 2023). For children, this can include drawing familiar paths, exaggerating meaningful locations, or omitting less relevant features (Jafarpour & Spiers, 2017; Swienton et al., 2023). For example, familiarity with spaces tends to expand perceived spatial areas while contracting the sense of time between places (Jafarpour & Spiers, 2017). Moreover, drawing maps within out-of-school programs such as Freedom Schools (FS) (Hogue, 2025) can foster sociopolitical awareness and identity development, particularly among marginalized youth (Davis et al., 2023). Thus, students' choices about landmarks to include, positioning and emphasis of places, and representations of

routes and symbols are examined as both visuospatial reasoning and expressions of lived experience. This connection is particularly relevant in the FS context, where students' community knowledge, identities, and experiences provide the social context through which their visuospatial representations are produced and interpreted.

Objective of the Study

The purpose of this study is to investigate the research question, "What aspects of visuospatial reasoning are present within various elements of children's map-making?" Using a visuospatial framework defined by Owens (2020) and operationalized by Suazo-Flores et al. (2023), this study aims to identify and interpret the presence of seven visuospatial aspects: 1) recognizing shapes, combinations, and their properties, 2) symmetry thinking, 3) comparing quantities and using ratios, 4) spatial capabilities, 5) locating, 6) intention and attention, and 7) space and place. By analyzing these elements, the study seeks to contribute to a deeper understanding of learners' visuospatial reasoning when representing familiar spaces.

Theoretical Framework

This study is theoretically guided by Owens's (2020) seven aspects of visuospatial reasoning as a theoretical framework through which to analyze participants' maps. The visuospatial reasoning aspects are as follows: 1) recognizing shapes, combinations, and their properties, 2) symmetry thinking, 3) comparing quantities and using ratios, 4) spatial capabilities, 5) locating, 6) intention and attention, and 7) space and place (Owens, 2020). Furthermore, visuospatial reasoning is embedded in peoples' languages and representations of mathematics (Owens, 2020). Therefore, we considered the ways that participants represented elements in their maps. More details of our analysis are in the Methodology section.

Related Literature

Visuospatial Reasoning

Visuospatial reasoning is the mental ability to imagine, manipulate, and interpret spatial relationships. It involves creating and transforming mental images to make sense of both concrete and abstract spatial ideas (Battista, 2007). Owens (2020) adds that this reasoning extends beyond physical spaces to intellectual contexts like interpreting diagrams or using graphs to represent abstract concepts.

One way learners develop visuospatial reasoning is through mapping tasks, which require them to coordinate geometric and environmental features. Cohrssen and Pearn (2021) found that children's maps often include recognizable elements like roads and intersections, showing that they can manage multiple perspectives in space. This view is consistent with Matthews's (1984) finding that children's first attempts at mapmaking reflect clear stages in how they come to understand and represent spatial relationships.

Mapping Tasks

Mapping has been widely studied across various disciplines, including mathematics education, to understand how individuals think spatially (Cohrssen & Pearn, 2021; Jafarpour & Spiers, 2017; Lopez & Bosco, 2022; Matthews, 1984; Meehan, 2018; Nuguid et al., 2024; Richardson et al., 2024; Swienton et al., 2023). One example is asking children to draw a map of the route from home to school, an activity that reveals how they conceptualize and represent their everyday environments (Cohrssen & Pearn, 2021; Richardson et al., 2024; Suazo-Flores et al., 2023).

Learners exhibit visuospatial reasoning when drawing maps by imagining and then drawing paths, trees, landmarks, and people in their typical routines (Nuguid et al., 2024; Richardson et al., 2024; Suazo-Flores et al., 2023). For example, drawing buildings represents 3-D objects in a 2-D space, which taps into learners' "visual intuition" (Cox, 2013, p. 4), critical for children's spatial reasoning development (Richardson et al., 2024; Suazo-Flores et al., 2023). This intuition draws from experiences with symmetry, edges of geometric figures, and "gap lines" or lines between identified points on a figure (e.g., space between the wall of a building and the door) (Richardson et al., 2024).

Methodology

FS is set in a local church where 30 K–5 racially and ethnically diverse scholars (children) gather to learn (Richardson et al., 2023). The FS STEAM (science, technology, engineering, art, and mathematics) workshops happened twice weekly where the scholars engaged in visuospatial activities such as mapping, making shapes with pattern blocks, building the tallest building with LEGOS, creating a representation of their building with Play-Doh, making dough, and building structures with Keva planks (Richardson et al., 2023).

Out of 30 IRB-approved FS scholars who volunteered to be research participants, we chose five listed in Table 1. We intended to demographically represent the particular FS site for this study

and present visually interesting maps. While we attempted to make the best effort to have a representative participant pool, there were additional demographics that were not included.

Table 1

Participants' Demographics

Participant (Pseudonyms)	Gender	Ethnicity	Grade
Jo	M	White	1st
Zeke	M	Black	2nd
Pat	F	Latina	3rd
Mary	F	White	1st
Robby	M	Asian	3rd

Data sources for this study include five maps in Figure 1 and five corresponding semi-structured interviews. Each participant was asked, "Tell me about your map" during mapping activities, and their responses were about two minutes. However, Jo (pseudonym) talked about his map for 10 minutes in an interview after the mapping activities. The interview protocol asked participants to describe their maps. Researchers scaffolded questions for participants to assist in content-specific vocabulary related to visuospatial reasoning that the children may not have used to describe their maps. For example, one of the researchers asked Jo if he drew from a top-view perspective (i.e., spatial capabilities).

Multiple researchers/authors of this study procedurally analyzed (Saldaña, 2021) five scholars' first maps with the prompt, "Draw a map from home to three places that you visit regularly" (Richardson et al., 2023) by using characteristics of visuospatial reasoning. Owens's (2020) seven visuospatial characteristics became qualitative codes. First, we analyzed maps independently by coding our observations of maps and corresponding transcript utterances into an Excel spreadsheet organized by Owens's (2020) seven codes. Then we discussed each map during five 60-minute meetings as a team, as the process of deductive analysis was dialogic (Guilfoyle et al., 2007). All of our codes were in the same spreadsheet, so we discussed our independent codes of participants' maps and utterances with each other such as similarities, alignment with the code, and areas of divergence.

Results and Discussion

This section presents the findings of each of the seven visuospatial codes for the set of five participants' maps, depicted in Figure 1. Participants were given the pseudonyms Mary, Robby, Pat, Zeke, and Jo.

Recognizing Shapes, Combinations, and Their Properties

This code was evidenced by geometric shapes typically taught in elementary school (e.g., squares, rectangles, and triangles). For example, Mary drew a house using a combination of rectangles, squares, and triangles. She also drew rectangles for the ladder on a slide and semi-circles on a swing set. Similarly, all participants' maps included combinations of shapes. For example, Robby used rectangular shapes for stairs. Pat drew triangle and rectangle combinations in her buildings. We observed that Jo's map included rectangles and a semi-circle. All five participants drew shapes such as rectangles, squares, and triangles and their combinations that iconically represented elements on their map.

Symmetry Thinking

Symmetry thinking was evidenced by visually aesthetic elements in maps and elements drawn equidistant from each other. For example, the five short lines on the two sides of the Wendy's® building in Pat's map were equidistantly spaced (with some deviation); however, Pat mentioned, "that's the parking spots" when asked about these lines which we interpret added her intent of symmetric thinking. Mary drew paths with symmetric dashes. Robby drew houses equidistantly spaced on his map, with some deviation. Jo drew two symmetric curved lines at the bottom of his map representing a partial view of a bus from his interview response, "That's the bus." All participants had at least one example of symmetry thinking on their map.

Comparing Quantities and Using Ratios

Comparing quantities and using ratios was evidenced by the size of elements in participants' maps drawn in relation to each other. For example, we noticed that Jo drew his house larger than the bus stop. Robby's pond was bigger than his clouds. Mary also explained how close or far away she drew some elements on her map. For example, one researcher asked her, "How did you decide where to put the house versus the church?" Mary responded, "Because our church isn't far away from the house…but our actual house is far away." The researcher clarified, "So your path kind of shows you how far away it goes?" Mary confirmed, "Ya. Because it's far away." Therefore, we inferred that the length of some of Mary's paths represented elements' proximity to each other. All five participants presented elements of varying sizes, suggesting evidence of this code.

Spatial Capabilities

Spatial capabilities were evidenced by two-dimensional (2-D) perspectives and the viewer's perspective (e.g., overhead, front-facing, or top-down view). For example, Jo drew an overhead view of a path and his house. Mary drew a playground set and a car from a front-facing 2-D view. Robby drew a front-facing view of his friends' houses the same sizes along an overhead view of the road. Zeke's map appeared to be top-down; however, the person and houses were drawn at a front facing view. Four out of five participants' maps had a mix of front-facing and overhead views of elements.

Locating

Locating was evidenced by mental mapping in the form of angles, curves, and contours on participants' maps. For example, we noticed that some participants' paths and roads between buildings were curved or angled which suggests evidence of locating. For example, Mary drew right-angled paths on her map. A researcher asked Mary, "How did you know where to put the house versus the church?" which could potentially elicit evidence of locating; Mary responded by describing the elements' proximity to each other (expanded in the Space and Place section). Jo drew a curved line connecting his house and the bus stop. Zeke drew a circular boundary that contained all the elements on his map. All participants' maps had observable evidence of locating with visible angles and curves drawn on their maps; however, more interviews would be needed to determine the meaning of the angles and curves in participants' maps.

Intention and Attention

Utterances from interview transcripts in addition to details on maps were evidence of participants' intention and attention (Owens, 2020). Based on the prompt, all participants drew maps from home to three places they visited regularly. The park, grocery store, school, church, and friends' houses were some examples of participants' intention and attention to details on their maps. Robby drew more details on his house and those of his friends, suggesting familiarity with houses such as visiting often (Jafarpour & Spiers, 2017). Some buildings were labeled, providing evidence of intention. In interview transcripts, participants identified what they drew. For example, Zeke said, "This is my house, this is the grass, this is the river, and that's my driveway," illustrating his intent to depict his experience. Overall, all five participants' evidence of intention and attention were present in transcripts and maps.

Space and Place

Space and place, evidenced by map details and utterances, conveyed participants' sense of belonging to a place or space (Owens, 2020; Suazo-Flores, 2023). The five participants' maps included places they visited regularly but details of the maps suggested their familiarity. Places that participants visited more often had more details than others. Mary described her map, pointing out places, "So this is our house…We go to church. This is Walmart." The church and Walmart were specifically labeled on Mary's map. Similarly, Pat drew details of Wendy's suggesting a familiarity with that place. She said, "This is Wendy's," and "favorite food" was written next to Wendy's on her map. Pat also pointed to the lines next to Wendy's saying, "that's the parking spots" when asked about them. All five participants discussed spaces and places on their maps.

Figure 1

Mary's Map



Figure 2

Robby's Map



Figure 3

Pat's Map



Figure 4

Jo's Map

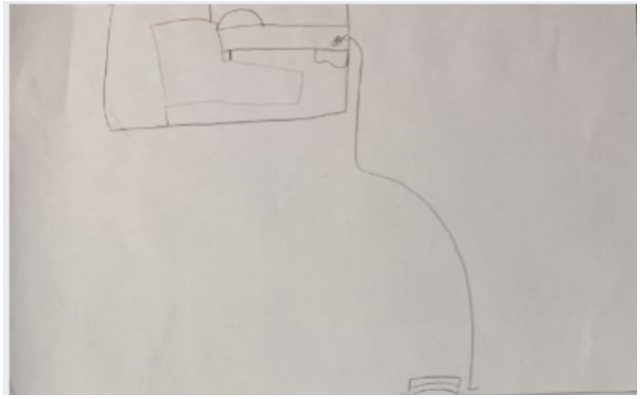
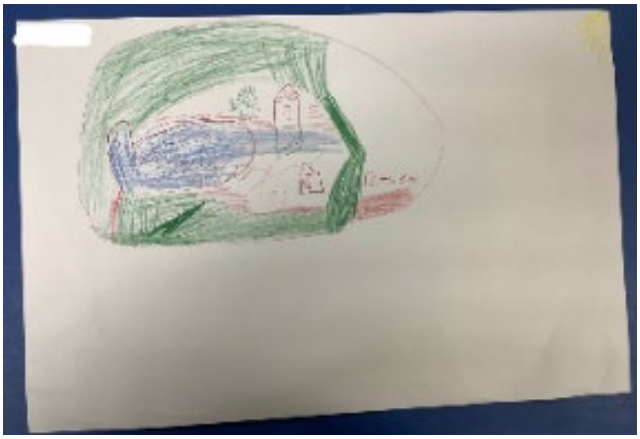


Figure 5

Zeke's Map



Implications

This study examined how children's map drawings reveal elements of visuospatial reasoning, using Owens's (2020) framework as operationalized by Suazo-Flores et al. (2023). The findings indicate that children's maps demonstrated evidence of recognizing shapes, symmetry thinking, comparing quantities and using ratios, spatial capabilities, locating, space and place, and expressing intention and attention through details in their maps and explanations of those representations. These findings demonstrate how mapping tasks can serve as a context for examining learners' visuospatial reasoning across mathematical and spatial ideas. Informal contexts such as FS can provide a context for children to develop and demonstrate visuospatial reasoning through activities such as map-making. These findings highlight learners' spatial and creative contributions as legitimate forms of knowledge. By analyzing their maps through a visuospatial lens, this study recognizes that children are not merely drawing but reasoning, engaging in geometric, proportional, and visuospatial thinking. Findings from this study illustrate the form that visuospatial reasoning takes in informal spaces outside of school contexts where learners are engaged in identity-forming activity.

References

Battista, M. T. (2007). The development of geometric and spatial thinking. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 843–908). National Council of Teachers of Mathematics.

Cohrssen, C., & Pearn, C. (2021). Assessing preschool children's maps against the first four levels of the primary curriculum: Lessons to learn. *Mathematics Education Research Journal, 33*(1), 43–60. https://doi.org/10.1007/s13394-019-00298-7

Cox, D. C. (2013). Similarity in middle school mathematics: At the crossroads of geometry and number. *Mathematical Thinking and Learning, 15*(1), 3–23. https://doi.org/10.1080/10986065.2013.738377

Davis, N. R., Marchand, A. D., Moore, S. S., Greene, D., & Colby, A. (2023). We who believe in freedom: Freedom Schools as a critical context for the positive, sociopolitical development of Black youth. *Race Ethnicity and Education, 26*(1), 34–53. https://doi.org/10.1080/13613324.2021.1969901

Guilfoyle, K., Hamilton, M. L., Pinnegar, S., & Placier, P. (2007). The epistemological dimensions and dynamics of professional dialogue in self-study. In J. J. Loughran, M. L. Hamilton, V. K. LaBoskey, & T. Russell (Eds.), *International handbook of self-study of teaching and teacher education practices* (pp. 1109–1168). Springer. https://doi.org/10.1007/978-1-4020-6545-3_28

Hogue, B. (2025). Teacher authorship as critical self-reflection and engagement in authentic student writing. *Literacy, 59*(2), 268–279. https://doi.org/10.1111/lit.70000

Jafarpour, A., & Spiers, H. (2017). Familiarity expands space and contracts time. *Hippocampus, 27*(1), 12–16. https://doi.org/10.1002/hipo.22672

Lopez, A., & Bosco, A. (2022). Distortions in spatial mental representation affect sketch maps in young adults. *Neurology International*, *14*(4), 771–783. https://doi.org/10.3390/neurolint14040064

Matthews, M. H. (1984). Environmental cognition of young children: Images of journey to school and home area. *Transactions of the Institute of British Geographers*, *9*(1), 89–105. https://doi.org/10.2307/621869

Meehan, P. J. (2018). "*Draw me a map of your town": An investigation of the construction of a free-recall hand-drawn map of a known physical environment by young children* (Publication No. 27775343) [Doctoral dissertation, Canterbury Christ Church University]. ProQuest LLC.

Nuguid, M. E., Aje, C. T., Unlu, M., & Richardson, S. E. (2024). Mapping spatio-visual activities in social contexts. *Indiana Mathematics Teacher* (2), 3–4.

Owens, K. (2020). Transforming the established perceptions of visuospatial reasoning: Integrating an ecocultural perspective. *Mathematics Education Research Journal, 32*(2), 257–283. https://doi.org/10.1007/s13394-020-00332-z

Richardson, S. E., Nuguid, M. E., Kastberg, S. E., Bofferding, L., & Hogue, B. (2024, July 7–14). *Emergence of spatio-visual reasoning in informal STEM* [Conference presentation and proceedings]. 15th International Congress on Mathematics Education (ICME-15), Sydney, Australia.

Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.). SAGE Publishing Inc.

Suazo-Flores, E., Nuguid, M. E., & Kastberg, S. (2023). Urban girls' visuospatial reasoning: Maps as eco-cultural tools to leverage lived experiences and spatial reasoning. In T. Lamberg & D. Moss. (Eds.), *Proceedings of the forty-fifth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 214–215). University of Nevada, Reno.

Swienton, H. A., Giordano, A., & Hagelman, R. R., III (2023). What do children draw when asked to draw a map? Results of a mental map experiment. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *58*(3), 116–143. https://doi.org/10.3138/cart-2022-0016

CENTERING MATHEMATICS: A MATH-FOCUSED INTEGRATED STEM APPROACH THROUGH DIGITAL FLOWERS GENETICS

Hanan Alyami
Purdue University in Fort Wayne
alyamih@pfw.edu

Rebecca Palu
Purdue University in Fort Wayne
palur@pfw.edu

Abstract

Despite various reform efforts toward integrated Science, Technology, Engineering, and Mathematics (STEM) education, meaningfully integrating mathematics into science instruction remains a persistent challenge. In this practitioner paper, we introduce a math-focused integrated STEM approach that foregrounds mathematical reasoning across problem scoping and solution generation. We illustrate the approach through Digital Flowers, an undergraduate genetics lesson in which students use Mendelian inheritance and probability to design multigenerational breeding strategies for producing a target genotype. We conclude with implications for designing interdisciplinary STEM experiences in which mathematics serves as a primary disciplinary anchor.

Keywords: math-focused integrated STEM, probabilistic reasoning, mendelian genetics.

Introduction

While integrated Science, Technology, Engineering, and Mathematics (STEM) emphasize disciplinary connections and has increasingly foregrounded engineering design and interdisciplinary problem solving (e.g., National Academy of Engineering [NAE] & National Research Council [NRC], 2014), mathematics remains underrepresented in such context (English, 2016; Fitzallen, 2015; Li & Schoenfeld, 2019). As a result, several scholars have noted the importance of developing integrated STEM learning experiences in which mathematics is given equal importance to other disciplines (Baker & Galanti, 2017). The joint statement by the National Council of Supervisors of Mathematics [NCSM] and the National Council of Teachers of Mathematics [NCTM] emphasizes the inclusion of "a strong mathematics component, a strong science component, and many opportunities to use mathematical and scientific thinking, reasoning, and modeling across disciplines to tackle real problems that involve any or all of the STEM fields" (NCSM & NCTM, 2018, p. 2). This perspective highlights mathematics as a powerful way of thinking and optimizing design solutions in complex integrated STEM problems.

We take up these calls by introducing a math-focused integrated STEM approach, where core mathematical concepts and practices serve as the primary perspective through which learners engage within an interdisciplinary design problem, while science, technology, and engineering knowledge and practices are embedded within the problem context and used to inform or constrain the mathematical reasoning and decision-making process.

Conceptualizing a Math-Focused Integrated STEM Approach

Integrated STEM education approaches have been conceptualized in various ways across the literature (Nave et al., 2021). One approach involves instructional activities with learning goals for content and/or practices from one or more STEM disciplines serving as anchors for the learning experience, while engineering and/or engineering design practices functioning as integrators to provide the structure through which disciplinary knowledge is applied to authentic problems, which creates opportunities for learners to engage with STEM concepts in meaningful contexts (Bryan & Guzey, 2020). Our math-focused integrated STEM approach builds on this design-oriented structure while intentionally shifting the disciplinary emphasis toward mathematics. Specifically, we draw on the STEM integration framework described by Moore et al. (2014) and the engineering design process employed within Tank et al.'s (2018) model. We retain Tank et al.'s six phases of the design process: define the problem, learn about the problem, plan a solution, try your plan, test your solution, and decide if your solution is good enough. However, while Tank et al. (2018) positioned "engineering design as the interdisciplinary glue" (p. 175), we adapt the disciplinary orientation of the design process by foregrounding mathematical reasoning within and across its phases.

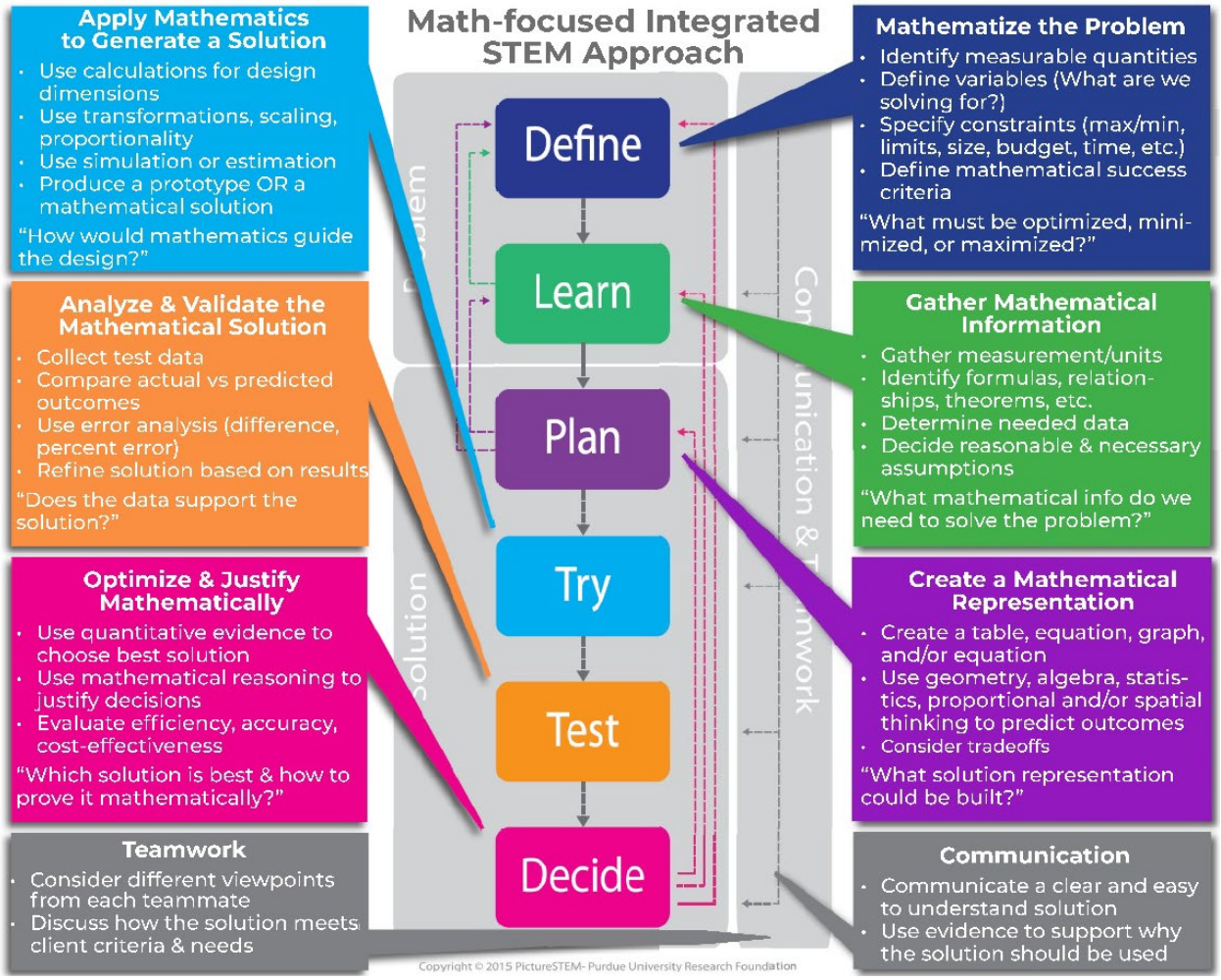
Within our adapted math-focused integrated STEM approach, mathematics plays a central role in each phase of the design process. During the define phase, learners mathematize the problem by identifying relevant measurable quantities, defining variables, specifying constraints, and/or articulating criteria for optimization. The learn phase focuses on gathering disciplinary and mathematical information needed to reason about the problem, including identifying relevant relationships, mathematical concepts, as well as reasonable assumptions needed to represent the problem's context. The plan phase involves constructing or selecting mathematical representations (e.g., tables, equations, probabilistic models) and using them to develop potential solution strategies. During the try phase, learners apply those strategies through mathematical reasoning, calculations, simulations, modeling, or other forms of quantitative analysis to generate solutions. The test phase involves using the mathematical evidence to analyze the performance of potential solutions. Finally,

in the decide phase, learners use mathematical evidence and justification to select and justify the most effective solution based on specified criteria such as efficiency, feasibility, and/or accuracy.

As described, our math-focused integrated STEM approach (Figure 1) positions mathematical reasoning as the primary means for problem scoping and solution generation, informing how potential solutions are generated, compared, and analyzed, and how decisions are justified in relation to specified criteria and constraints.

Figure 1

Math-focused Integrated STEM Approach (adapted from Tank et al., 2018)



A key feature of our math-focused integrated STEM approach is that mathematical reasoning plays a central role in representing the problem, generating and analyzing potential solutions based on specified criteria and constraints. In the following section, we describe how this approach was operationalized to design a multi-stage, constraint-based probabilistic optimization task embedded in an undergraduate biology context (i.e., genetics).

Putting the Approach into Practice: Designing the *Digital Flowers* Lesson

Since mathematics is underrepresented in integrated STEM opportunities, students engage in mathematical procedures without experiencing mathematical reasoning within scientific context. This pattern can be observed in genetics, where students experience Mendelian inheritance procedurally (Newman et al., 2021) despite evidence that conceptual understanding of the underlying genetic mechanisms and the intersection between genetics and probability supports deeper reasoning and knowledge transfer (Borja & Mutya, 2024; Schuchardt & Schunn, 2016). Specifically, students often complete Punnett squares but struggle to apply and connect the mathematical structure of probability rules to generalize and extend to other contexts and practical Mendelian inheritance situations (Avena & Knight, 2019; Cavallo, 1996; Stewart, 1983).

While Mendelian inheritance is usually taught procedurally (e.g., Punnett squares calculations), we designed the *Digital Flowers* lesson to reposition genetics as a multi-step mathematical optimization problem. Rather than working towards isolated answers, students work toward a clear goal (i.e., generating blue roses) by developing and refining strategies across multiple generations. Instead of asking students to complete isolated Punnett squares, students were tasked with designing efficient breeding strategies to produce a target genotype/phenotype given specific constraints. In this context, students designed a breeding strategy using mathematics as the primary means for generating and evaluating solutions. Specifically, students planned multigenerational crosses and optimizing their approach for efficiency and feasibility. Students identified measurable quantities and variables (e.g., genotypes, phenotypes, allele combinations, and associated probabilities) and constructed representations that included Punnett squares and branched line diagrams to model inheritance across multiple generations. These representations were used alongside table summaries to articulate mathematical relationships and predict outcomes.

The task centers around the generation of a single target genotype (*RR YY WW ss*) starting from constrained seed bags with specified genotypes and phenotypes. Unlike other flower colors in this system, which can result from multiple genotypic combinations, generating blue roses requires producing one exact genotypic combination. This constraint creates an optimization problem (probabilistic uncertainty + Mendelian inheritance), in which students must determine not only how to produce the target genotype but also *how efficiently* they can do so. To support this work, students used a searchable genotype–phenotype table (81 genotypes across four genes) to identify outcomes and inform their breeding strategies. Completion of the activity requires students to develop and implement a number of critical skills, including strategic planning with multi-generational modeling

(e.g., F1 through F4), track genotypic and phenotypic distributions, use probability to compare the efficiency of different strategies, predict outcomes, and decide which offspring to keep for future crosses through genetic testing (i.e., a test cross) to confirm a genotype. In the second part of the activity, students developed their own independent strategy, which allows them to explore multiple solution paths and tradeoff evaluation.

A key feature of the lesson is its emphasis on efficiency. Students compare strategies based on (a) the number of generations required and (b) the probability of producing the desired outcome, where the "best" solution is not always the fastest or most direct option. Instead, students balance the likelihood of success with time and effort. The lesson structure aligns with our math-focused integrated STEM approach as outlined below:

- **Define**: Students are given a specific mathematized problem where the goal is to generate blue roses with a specific genotype (i.e., *RR YY WW ss*) serving as a precise target to be achieved in various ways. Students also identify the constraints imposed by the available starting genotypes.
- **Learn**: Students use genotype & phenotype tables along with their knowledge of Mendelian inheritance and probability to gather necessary biological and mathematical information required to develop a breeding strategy.
- **Plan**: Students use a combination of modeling through Punnett squares or probability calculations in conjunction with their genotype/phenotype table to determine genotypes and phenotypes of offspring from their putative crosses. They are largely allowed to select their preferred method of problem-solving, but in later cases, students need to choose an approach that is both accurate and efficient. For most complex crosses, probability calculations and branched line diagrams are more efficient and less prone to errors than Punnett squares.
- **Try**: The first part of the lesson is a guided phase, where students follow a structured sequence of crosses that will, if followed correctly, allow them to produce the blue rose. By calculating or simulating the outcomes of proposed successive crosses, students determine the probabilities of obtaining genotypes needed for subsequent generations.
- **Test**: The second part of the lesson is an open-ended phase, where students develop and analyze their own strategy to produce the targeted genotype. While there is only one genotype for the blue flower phenotype, there are several paths that can be taken to produce

these flowers from the original "seed bags." Students use an online simulation tool to evaluate their own hypotheses according to criteria such as probability of success and the number of required generations.

- **Decide**: Students reflect on the costs and benefits of reducing the number of steps in the process (i.e., fewer generations to produce the offspring) versus the ultimate percentage of offspring in the final cross that will display the target phenotype (i.e., a higher percentage of blue roses in the final group of offspring). They are asked to consider costs such as generation time as well as the likelihood of producing the desired outcome, considering the number of offspring that must be screened to identify the target (e.g., comparing a 4-generation strategy with a 25% success rate to a 3-generation strategy with a 1.5% success rate). For generation time, they are asked to consider how long the farmer or breeder may be willing to wait or the cost of several rounds of breeding. This gives them additional practice while also placing their choices in a real-world context and application of genetics, Mendel, and probability.

Classroom Examples and Reflections from Implementation

This lesson was implemented in a sophomore-level genetics course during a Spring 2026 lab session following instructions on Mendelian inheritance, gene-by-gene interactions, and probability. It was designed to complement a traditional lab based on the fruit fly *(Drosophila melanogaster*) to provide a parallel context where students could apply inheritance patterns alongside probabilistic reasoning and encounter consistent probability structures across distinct biological situations. To examine students' experience of the lesson and inform future revisions, we collected pre- and post-lesson responses and follow-up voluntary follow-up interviews. The examples presented below illustrate students' experiences and reasoning during and following the implementation.

Students' reflections following the lesson provided evidence of how they engaged in different forms of mathematical reasoning to develop a strategy for reaching a specified outcome. For example, one student explained that prior problems typically provided a starting point and asked students to determine the resulting outcome, while the *Digital Flowers* task reversed that direction by specifying the desired outcome and tasked students with reaching it. Additionally, students highlighted the role of probability in evaluating competing breeding strategies. Some students noted the central challenge as deciding between a shorter pathway with a lower probability of obtaining the

blue roses or a longer pathway with a substantially higher probability while considering the time and expected yield.

Students described how the lesson made the relationship between probability and genetics more explicit and consequential, where probability was used as information to anticipate outcomes and make decisions across multiple generations. One student described a greater awareness of "the mathematical side" of genetic problems and reflected on how the lesson encouraged them to consider the probability associated with successive genetic outcomes. Another student described needing to think several steps ahead toward the final goal rather than completing random crosses and observing outcomes.

The lesson structure included built-in checkpoints where students had their work checked by an instructor to monitor their progression and understanding. Instructors observed students beginning to rely on probability to guide their breeding decisions. One challenge we observed involved the use of the genotype-phenotype table. Students struggled to search it efficiently due to formatting limitations where the genotype *RR* was read as the same as *Rr* and *rr* by the search tool unless students specifically selected "case sensitive" in the search tool. In future iterations, we will revise this resource to improve usability and reduce unnecessary cognitive load.

Implications

This work contributes to ongoing conversations about integrated STEM by demonstrating how mathematics can function as the primary disciplinary anchor rather than a supporting tool (e.g., English, 2016; NCSM & NCTM, 2018). The math-focused integrated STEM approach presented here positions mathematical reasoning as the mechanism through which design decisions are generated, evaluated, and optimized. By framing a genetics context as a constraint-based design problem, the *Digital Flowers* task illustrates how math-focused integrated STEM guides the inquiry process, while maintaining meaningful engagement with disciplinary science content. In practice, this repositioning of mathematics is reflected in the types of reasoning students engaged with throughout the lesson. Students identified and coordinated variables (e.g., genotypes, phenotypes, allele combinations), worked within constraints, and constructed probabilistic models to predict outcomes across multiple generations. They engaged in information gathering through test crosses and evaluated competing strategies using trade-off analysis, particularly in balancing the proportion of desired outcomes with the number of generational steps required. These forms of reasoning extend

beyond procedural applications of Punnett squares, emphasizing mathematics as a tool for optimization and decision-making within a scientific context.

More broadly, this work has implications for biology teaching, integrated STEM design, and the development of interdisciplinary tasks. For biology instruction, the lesson demonstrates how genetics can be reframed to foreground probabilistic reasoning to create opportunities for deeper conceptual reasoning about inheritance beyond relying exclusively on procedural applications of representations such as Punnett squares. For integrated STEM, the approach illustrates how an existing design-process structure can be adapted where mathematics serves a central role in problem representation, solution generation, evaluation, and decision-making, while disciplinary science knowledge provides a meaningful context for application. Additionally, students' reflections suggest that the use of a familiar digital environment highlights the potential of game-based contexts to support engagement without sacrificing disciplinary rigor. This work suggests that designing tasks around constraint-based optimization with clear goals and trade-off opportunities can support more balanced and meaningful math-focused integrated STEM learning experiences.

References

Avena, J. S., & Knight, J. K. (2019). Problem solving in genetics: Content hints can help. *CBE—Life Sciences Education, 18*(2), ar23.

Baker, C. K., Galanti, T. M., Kraft, T., Holincheck, N., Hjalmarson, M., & Nelson, J. K. (2022). Researchers as coaches: Developing mathematics teaching capacity using MEAs for STEM integration. *Investigations in Mathematics Learning, 14*(1), 28–48. https://doi.org/10.1080/19477503.2021.2023966

Borja, E. A., & Mutya, R. C. (2024). Enhancing students' conceptual understanding on the patterns of Mendelian genetics through task-based learning. *Jurnal Pendidikan Biologi Indonesia, 10*(1), 107–114.

Bryan, L. A., & Guzey, S. S. (2020). K–12 STEM education: An overview of perspectives and considerations. *Hellenic Journal of STEM Education, 1*(1), 5–15. https://doi.org/10.51724/hjstemed.v1i1.5

Cavallo, A. M. L. (1996). Meaningful learning, reasoning ability, and students' understanding and problem solving of topics in genetics. *Journal of Research in Science Teaching, 33*(6), 625–656.

English, L. D. (2016). STEM education K–12: Perspectives on integration. *International Journal of STEM Education, 3*(1), 1–8. https://doi.org/10.1186/s40594-016-0036-1

Fitzallen, N. (2015). STEM education: What does mathematics have to offer? In M. Marshman (Ed.), *Mathematics education in the margins: Proceedings of the 38th annual conference of the Mathematics Education Research Group of Australasia* (pp. 237–244). MERGA.

Li, Y., & Schoenfeld, A. H. (2019). Problematizing teaching and learning mathematics as "given" in STEM education. *International Journal of STEM Education, 6*, Article 44. https://doi.org/10.1186/s40594-019-0197-9

Moore, T. J., Stohlmann, M. S., Wang, H.-H., Tank, K. M., Glancy, A. W., & Roehrig, G. H. (2014). Implementation and integration of engineering in K–12 STEM education. In S. Purzer, J. Strobel, & M. Cardella (Eds.), *Engineering in precollege settings: Synthesizing research, policy, and practices*. Purdue University Press.

National Academy of Engineering & National Research Council. (2014). *STEM integration in K–12 education: Status, prospects, and an agenda for research*. National Academies Press.

National Council of Supervisors of Mathematics & National Council of Teachers of Mathematics. (2018). *Building STEM education on a sound mathematical foundation: A joint position statement on STEM.*

Nave, S., Kaya, F., Boone, B., Brewster, C., Calvelage, K., Ferdous, T., Hood, E., Sass, L., & Zimmerman, M. (2021). "Beyond an acronym, STEM is …": Perceptions of STEM. *School Science and Mathematics, 121*(1), 36–45.

Newman, D. L., Coakley, A., Link, A., Mills, K., & Wright, L. K. (2021). Punnett squares or protein production? The expert–novice divide for conceptions of genes and gene expression. *CBE—Life Sciences Education, 20*(4), Article 53. https://doi.org/10.1187/cbe.21-01-0004

Schuchardt, A. M., & Schunn, C. D. (2016). Modeling scientific processes with mathematics equations enhances student qualitative conceptual understanding and quantitative problem solving. *Science Education, 100*(2), 290–320.

Stewart, J. (1983). Student problem solving in high school genetics. *Science Education, 67*(4), 523–540.

Tank, K. M., Moore, T. J., Dorie, B. L., Gajdzik, E., Sanger, M. T., Rynearson, A. M., & Mann, E. F. (2018). Engineering in early elementary classrooms through the integration of high-quality literature, design, and STEM+C content. In L. English & T. Moore (Eds.), *Early engineering learning* (pp. 175–201). Springer.

INTRODUCING AN ONLINE FLIPPED MODEL (OFM) IN UNIVERSITY PRECALCULUS COURSES: A PRACTITIONER'S GUIDE

Nelson L. Carter
University of Houston-Clear Lake
carter@uhcl.edu

Sue Brown
University of Houston-Clear Lake
browns@uhcl.edu

Abstract

This practitioner paper shares the implementation of an Online Flipped Model (OFM) in three undergraduate Precalculus courses. The authors describe how instructors who are new to online teaching can adopt and implement an OFM by outlining the overall course design, the development of instructional videos, the technologies utilized, and the grading approach. The study also reflects on lessons learned, highlighting experiences and perspectives that both align with and challenge existing research and conventional practices in online instruction.

Keywords: precalculus; flipped model; online flipped model; video length; didactic contract

Introduction

According to recent national data, approximately 55% of students who begin college as Science, Technology, Engineering, and Mathematics (STEM) majors actually complete a STEM degree, while 29% do not complete a degree, and 16% complete a degree outside of STEM (National Center for Science and Engineering Statistics, 2026). Research may offer a partial explanation for this low completion rate. In a meta-analysis of 225 studies related to undergraduate STEM courses, students in lecture-based classes were 1.5 times more likely to fail when compared to students taught with active learning models (Freeman et al., 2014). In contrast to lecture-based instruction, one active-learning approach is the Flipped Model (FM), in which instructors report higher levels of student engagement than in traditional lecture-based courses (Unruh et al., 2016).

The Flipped Model (FM) moves direct instruction outside of class through videos, class notes, and other instructional resources, allowing class time to focus on active learning and interaction. This approach raises an important question: What does the FM look like when instruction occurs online through synchronous virtual meetings? The purpose of this paper is to examine the implementation of an Online Flipped Model (OFM) in undergraduate Precalculus courses and to describe practical considerations for instructors interested in adopting this instructional approach.

Instructional Framework

This paper adopts the following definition of the Online Flipped Model (OFM): Students complete online pre-class activities asynchronously, such as watching prerecorded video lectures, responding to discussion questions, or completing quizzes (Swart & Macleod, 2020). Next, students meet with their instructor and classmates online (synchronously), using a communication platform (such as Zoom or Microsoft Teams), rather than meeting in a physical classroom. Thus, the OFM combines the two major components of flipped instruction, asynchronous delivery of instructional content before class and synchronous class time devoted to interaction, active learning, and individualized feedback (Bergmann & Sams, 2012). The present study examines flipped instruction in which both the pre-class and class components occur online.

Theoretical Framework and Related Literature

The rationale for the OFM is grounded in research on flipped instruction. By moving initial content delivery outside of class, flipped instruction creates opportunities for synchronous class time to focus on active learning, interaction, and instructor feedback. Research has shown that flipped instruction can improve student engagement and learning outcomes compared with traditional lecture-based instruction (Elmaadaway, 2018). These findings provide a foundation for the OFM, in which students engage with content asynchronously and use synchronous class time for interaction and active learning. Because the present study takes place entirely online, research on flipped instruction in online environments is also relevant.

The shift from face-to-face to online flipped instruction has been examined in business and education courses. Studies have found similar levels of student satisfaction and achievement in online and face-to-face flipped environments (Swart & Macleod, 2020). These findings support extending the flipped model to online settings and provide a basis for examining its implementation in an online mathematics course.

Research in undergraduate mathematics courses provides a further connection to the present study. Collins (2019), for example, found improved achievement among students enrolled in a flipped Precalculus course. This finding supports the use of flipped instruction in Precalculus while highlighting the need to examine models in which both asynchronous preparation and synchronous class meetings occur entirely online. The present study addresses this context by examining the implementation of an OFM in an undergraduate Precalculus course, with particular attention to instructional practices used in this online environment.

Methodology

This practitioner paper is an adaptation of a quantitative study conducted as part of a doctoral dissertation (Carter, 2022). The corresponding author served as the doctoral student and course instructor, and the coauthor served as the dissertation chair. This paper focuses on implementation of the Online Flipped Model (OFM), student perceptions and achievement, and instructors' observations and reflections.

The study for the doctoral dissertation was conducted at a medium-sized university in the Gulf Coast region of Texas during the 2020–2021 academic year (Carter, 2022). Of the 141 students purposefully selected from all Precalculus sections for development and validation of the Precalculus Self-Efficacy Survey (PCSES), 81 participated in the dissertation study examining self-efficacy, grit, and achievement in OFM courses. Students completed the PCSES and 12-Item Grit Scale (GS) before and after the course, and course grades provided a measure of achievement. Pretest and posttest PCSES and GS scores were analyzed using two-tailed paired t-tests to determine statistically significant mean differences. Pearson correlations determined relationships between precalculus self-efficacy and achievement and between grit and achievement. Multiple regression techniques determined if precalculus self-efficacy or grit could predict achievement.

For this practitioner paper, student evaluations provided anecdotal perceptions of the course, while instructors' written reflections documented their observations of the OFM. These sources were included in this paper to describe the implementation of the OFM, student achievement and perceptions, and the instructors' experiences with the approach.

Assessment of Student Performance

Each student was graded with a weighted average according to the following categories: (a) online classwork problem sets (i.e., WebAssign) (15% of the overall numeric average after the lowest two grades were dropped), (b) participation (15%), (c) challenge problems (10%), (d) midterm examination (25%), and (e) the cumulative final examination (35%). Online classwork problem sets were built in a textbook-accompanied online homework system called WebAssign. Each set consisted of digital questions that were selected from one section of the textbook. Students' responses were automatically graded by WebAssign. Students were encouraged to work on the problems in class and to complete any unfinished work outside of class.

Each student received a participation grade each week. This weekly participation grade was a composite of two other grades: a Weekly Assignment grade, and a weekly attendance grade. The

instructor created and distributed a Weekly Assignment each Monday before class, which consisted of approximately four exercises similar to examples in the prerecorded lecture videos. Regarding attendance, after each class meeting ended, the instructor ran Zoom reports and noted the amount of time each student attended. To allow for potential disconnection issues, a student was defined as being present if they attended at least 25 minutes of a 50-minute class. Each week, a student's participation grade simply inherited the higher grade of their Weekly Assignment grade and their weekly attendance grade.

Practice/Innovation

Prior to using the OFM, the instructor had taught Precalculus in a face-to-face lecture format for more than 8 years and had taught Calculus I in a flipped, face-to-face model five times. The COVID-19 Pandemic prompted the change. The OFM discussed in this section was as similar to a traditional FM as possible, so that the only substantial change was the in-class sessions were conducted synchronously online in Zoom instead of face-to-face in an actual classroom on campus. To accomplish this, the instructor required synchronous online meetings at a formal class time four days per week (Monday through Thursday) for 50 minutes. In each class meeting, the instructor divided each class into two formats: addressing the entire class and addressing breakout rooms. Breakout rooms were virtual meeting places for small groups of 3–4 students in the class. The table below provides a snapshot of how each class was organized.

Table 1

Online Flipped Model Class Structure

Time	Format	Instructor Role	Student Role
Before Class	Individual Students	Post videos; provide notes	Watch videos; take notes
0–10 min	Entire Class	Clarifies misconceptions	Ask questions
10–45 min	Breakout Rooms	Visits groups; provides feedback	Collaborate; solve problems
45–50 min	Entire Class	Address Misconceptions	Listen; Ask Questions

Results

The dissertation study also examined students' self-efficacy and grit. Findings indicated provided evidence that the Precalculus Self-Efficacy Survey (PCSES) was a valid and reliable measure. Results indicated that students' precalculus self-efficacy increased in classes taught with the OFM, but their grit decreased. Furthermore, precalculus self-efficacy significantly predicted achievement, whereas grit did not. Of the 112 students enrolled in the courses, 95 students (84.8%)

earned a grade of C- or higher, while 17 students (15.2%) earned a D, F, or withdrew from the course. Forty-nine students completed both pre- and post-surveys. Results indicated a statistically significant increase in self-efficacy from the beginning to the end of the semester, with a large effect size. Student perceptions of the OFM were mixed. Some students appreciated opportunities to ask questions and collaborate with classmates. Others reported challenges with online group work and indicated that some instructional videos were too long.

Implications: Lessons Learned

Video Length

In our opinion, students will determine strategies to be successful within the first few weeks of a term, deciding the extent to which an instructor's resources will be part of their learning strategy. This agrees with findings from Lage et al. (2000) who found that, when a variety of different types of learning resources (e.g., PowerPoints and videos) are provided to students, they generally coalesce around one type of resource early in the term, having decided which resources work for them. With that in mind, for instructors who plan to start developing resources for any flipped models, the authors strongly recommend polishing the first few videos as much as possible. The course's earliest videos should have good sound quality, cohesive editing, and be of reasonable length.

In terms of video length, we have some opinions about research findings. In a literature review, researchers found a large body of research suggests that instructor-created videos should be short—approximately six minutes in length (Lo et al., 2017). The authors understand the argument that breaking videos into shorter segments decreases students' cognitive load, freeing up mental energy to facilitate learning (Mayer, 2014). However, we are skeptical that six-minute videos are ultimately best for learners in all situations. While some students will balk when they open their assigned video and discover it is two hours long, we believe those students will still balk when they discover they have been assigned 20 videos that are each six minutes long.

One way we tried to navigate the video length problem was to make videos by learning objective instead of by sections or chapters; however, feedback presented here suggests some students still felt that videos were too long. The authors believe an updated grounded theory study is needed to develop a theoretical framework regarding how students utilize instructional video content post-pandemic. We hypothesize research will be able to categorize learners by their video utilization preferences. For example, some students may prefer to watch the videos to take detailed

notes, whereas others may prefer to skip through videos to watch examples. Pragmatically, future research may affirm that short video length is the crucial factor related to students' video viewing habits. However, the authors believe adherence to the six-minute rule may lead to the type of compartmentalized learning Hiebert and Lefevre (1986) warned against 40 years ago—it is difficult for continuously connected instruction to be delivered in discrete, six-minute bits.

Considerations Before Starting to Make Videos

The authors would recommend beginners start slowly by planning to flip only a few lessons at the beginning. Learning objectives that are procedural in nature may be a good place to start because those types of lessons require students to watch demonstrations. Since it can be hard to write notes and watch at the same time, video is a great medium for procedural topics. Furthermore, we recommend starting with especially dense sections that are difficult to pace and adequately cover in a typical face-to-face format. Post-pandemic, the authors find students need more JiT instruction. Unfortunately, this increases demand for instruction time, which is a zero-sum resource.

Unless practitioners frequently remind students about the video pacing schedule, they may receive negative student feedback, such as, "The instructor just shows up to class and answers questions; I'm having to teach myself!" Therefore, from the beginning of the semester, the authors recommend practitioners clearly communicate these differences with their students, so students know what their instructor's expectations are. Along those lines, the authors recommend instructors consider including a learning agreement as an attachment to the course syllabus, such that students must sign the agreement and turn it in for their first assignment. In addition, we recommend instructors explain the possibility of backlash to their administrators in advance.

Student Participation in the Model

Practitioners may be hesitant to adopt an OFM due to concerns that students may not complete pre-class activities, which might discourage students' attendance and engagement. When students do not watch the videos before class, we recommend instructors encourage those students to watch the videos during the scheduled class time. When students see other classmates working ahead on homework, those students may become extrinsically motivated to watch the videos before class so they can join their classmates in the future.

Another lesson learned is that webcam usage can be awkward. Gherheş et al. (2021) found that students preferred to keep their web cameras off. Since the social mores about webcams are delicate, we recommend that instructors try icebreaker activities and establish *classroom norms*

regarding webcam expectations. For example, instructors could recommend students turn their microphones on as needed but encourage students to check-in with their groupmates every 5–10 minutes.

Regarding the weekly participation component of the course, it was not difficult to automate in the LMS gradebook, but we concede that it would not be preferred for an instructor that manually calculates their grades. We understand some criticism that the participation grade was convoluted. However, convoluted as it may be, it provided students with flexibility when flexibility was needed. Some students may have misunderstood and believed that each grade was explicitly factoring into their numeric average; therefore, repeated applications of the model may lead to different results in that regard. In that case, the authors recommend adjusting the numeric average weight, where Weekly Assignments and attendance directly count toward each students' grade. Offering to drop the lowest grade or two could provide similar flexibility. Ultimately, we believe students in entry-level college mathematics courses need structure.

The implementation of an Online Flipped Model (OFM) in three undergraduate Precalculus courses provided practical lessons for instructors considering this approach. The model moved instructional content outside the scheduled class period and used synchronous Zoom meetings for questions, collaborative problem solving, and instructor feedback. Students valued these experiences but some faced challenges with online group work and video length. Successful implementation requires more than moving lectures to prerecorded videos. It also requires clear expectations, thoughtfully designed asynchronous materials, and intentional use of synchronous class time to promote interaction, collaboration, and problem solving.

References

Bergmann, J., & Sams, A. (2008). Remixing chemistry class. *Learning & Leading with Technology*, *36*(4), 22–27. Education Source (508021497).

Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. International Society for Technology in Education.

Carter, N. L. (2022). *Measuring precalculus self-efficacy, grit, and achievement in university precalculus courses taught with an online flipped model* [Doctoral Dissertation, University of Houston-Clear Lake]. https://doi.org/10.13140/rg.2.2.20870.24643

Collins, B. V. C. (2019). Flipping the precalculus classroom. *International Journal of Mathematical Education in Science and Technology*, *50*(5), 728–746. https://doi.org/10.1080/0020739X.2018.1535098

Elmaadaway, M. A. N. (2018). The effects of a flipped classroom approach on class engagement and skill performance in a Blackboard course. *British Journal of Educational Technology*, *49*(3), 479–491. https://doi.org/10.1111/bjet.12553

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Gherheş, V., Şimon, S., & Para, I. (2021). Analysing students' reasons for keeping their webcams on or off during online classes. *Sustainability*, *13*(6), Article 3203. https://doi.org/10.3390/su13063203

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Lawrence Erlbaum Associates, Inc.

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, *31*(1), 30–43.

Lo, C. K., Hew, K. F., & Chen, G. (2017). Toward a set of design principles for mathematics flipped classrooms: A synthesis of research in mathematics education. *Educational Research Review*, *22*, 50–73. https://doi.org/10.1016/j.edurev.2017.08.002

Mayer, R. E. (Ed.). (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369

Natarajan, R., & Bennett, A. (2014). Improving student learning of calculus topics via modified just-in-time teaching methods. *PRIMUS*, *24*(2), 149–159. https://doi.org/10.1080/10511970.2013.854853

National Center for Science and Engineering Statistics. (2026). *The state of U.S. science and engineering 2026: STEM talent, education, training, and workforce.* National Science Foundation.

Swart, W. W., & Macleod, K. R. (2020). Flipping online analytics classes: Achieving parity with their face-to-face counterparts. *Decision Sciences Journal of Innovative Education*, *18*(1), 119–137. https://doi.org/10.1111/dsji.12200

Unruh, T., Peters, M. L., & Willis, J. (2016). Flip this classroom: A comparative study. *Computers in the Schools*, *33*(1), 38–58. https://doi.org/10.1080/07380569.2016.1139988

OPERATIONALIZATION OF PARTNERSHIP IN MATHEMATICS EDUCATION RESEARCH: A SCOPING REVIEW

S. Megan Che
Clemson University
sche@clemson.edu

Lianne Jones
Appalachian State University
joneslm15@appstate.edu

Amy James
Baylor University
amy_b_james@baylor.edu

Trena L. Wilkerson
Baylor University
Trena_Wilkerson@Baylor.edu

Colleen M. Eddy
University of North Texas
Colleen.Eddy@unt.edu

Abstract

This scoping review focuses on the contexts, configurations, and scale of research partnerships in mathematics education from 1989 to the present. We begin in 1989 because we reason that the 1989 NCTM Standards motivated many research partnerships, particularly in North America, as stakeholders responded to these new expectations. The review process includes 4 phases: Literature Search and Retrieval, Document Screening, Data Collection, and Data Extraction. Some salient aspects of the data are that 50% of research partnerships rely on external funding sources; that none of the studies partnered with teachers of pre-primary aged pupils; and that more than half of the studies partnered with in-service rather than prospective teachers.

Keywords: research partnership, mathematics education, institutes of higher education, schools, in/preservice teachers

Introduction

The work of deeply exploring mathematics teaching and learning is fundamentally collaborative, drawing from the experiences and expertise of students, teachers, school leaders, community members, scholars, and other stakeholders. Meaningful change in the educational space is rarely accomplished in isolation; rather, it tends to occur through sustained collaborative work that draws upon the strengths, resources, and perspectives of different individuals and entities. Such partnerships can take many forms, including lesson studies, action research, and school-university partnerships, to name a few. Research conducted within these partnerships is central to understanding the teaching and learning of mathematics, as underscored by prominent scholars in the field (e.g., Ruiz, 2021; Wilkerson, 2022), as well as the National Council of Teachers of Mathematics (NCTM, 2014, 2021). The continued use of partnerships to explore problems of

practice necessitates an understanding of the contexts, configurations, and supports that characterize those partnerships. In doing so, we can illuminate the contexts in which we engage (or not) in research partnerships in mathematics education. Scholars frequently share findings generated through partnerships; however, less is known about the configurations, durations, and scale of the partnerships, as well as how they are supported. To address these questions, we present a scoping review of research partnerships in mathematics education from 1989 - present, contributing to an initial mapping (Peters et al., 2015) of the work we undertake with fellow interested partners, which opens up new, relatively understudied, potential contexts for future such partnerships.

Objectives of the Study

In mathematics education, as a field, we as yet do not have a view on contexts of research partnerships, including to what extent such partnerships are externally funded, who we tend to collaborate with in this work, and typical durations of partnerships. To begin addressing these unknowns, with the aim of uncovering potential 'blind' or underrepresented partnership contexts, we conducted a scoping review to map out the landscape of research partnerships in mathematics education (Peters et al., 2015). This study focuses on the following research question and subquestions: What are the contexts of mathematics research conducted in partnership? Specifically, who engages in mathematics research partnerships? How long do partnerships last? What is the scale of each partnership? What types of approaches, or partnership ecologies, do partnerships employ to guide their work? What sources of support facilitate research in partnership? The insights gained from this field-level view can uncover important but perhaps overlooked contexts of research partnership, and can provide our field with a deeper sense of the resource-intensive nature of conducting research in partnership.

Theoretical Framework and Related Literature

As we have been investigating the term *partnership*, it has become clear that its meaning can be nebulous. At times, for example, the term partnership is used interchangeably with other terms such as collaboration. Although related, we define these terms distinctly. For this reason, we feel as though it is important for us to define what we mean by partnership, as well as how it differs from other frequently-utilized characterizations such as relationship and collaboration.

Relationships are the core of partnerships; however, it is possible to be in relationship with others without being in a partnership. We conceptualize relationships as entities connected by shared interest. When relationships are extended such that parties are working together to achieve a

shared task or goal, we consider this a collaboration. Partnerships, then, extend the nature of collaboration with a commitment to sustained engagement (Casey, 2008). For the purposes of this study, we define a *research partnership* as a specific context of relationships or collaborations such that two or more entities engage in conducting research together. For these endeavors to be considered a research partnership, all members of the partnership must be involved, in some way, in the research itself (e.g., decision-making, member-checking, data collection and analysis).

Interest in better understanding partnership processes in research intensified following the advent of the external funding structure of the Mathematics Science Partnership (MSP) program, which, for instance, led to Scherer's (2008) review of instruments for evaluating partnerships in MSPs. Currently, Research Practice Partnerships (RPPs) are prioritized by funding agencies such as the National Science Foundation (NSF) in the U.S. and have provided impetus for examining partnerships between researchers and practitioners. For instance, early childhood education practitioners and scholars devoted a special issue of *the Future of Children* in the Spring of 2021 to the potential of RPPs to benefit and strengthen early childhood education. Penuel's work, often with Coburn, has greatly contributed to our understandings of dynamics, tensions, negotiations, and equity in partnerships involving researchers and practitioners (see Coburn & Penuel, 2016; Penuel & Gallagher, 2017). Informed by Penuel's work, the NSF has prioritized funding RPP projects in its Computer Science For All program; in a recent solicitation, approximately 75% of funded projects were anticipated to involve RPPs (NSF, 2024). Though not every RPP studies itself formally, the existence of an RPP provides rich opportunities to do so, as Penuel et al. (2015) and Henrick et al. (2017) illustrate in the field of education and as Superfine et al. (2023) indicate specifically within mathematics education.

As yet, we can find no scoping or systematic reviews of research partnership contexts or processes focusing primarily on mathematics education and considering a range of potential partnership configurations or frames, such as activity theory or sociocultural theory, among additional approaches. The absence of such reviews restricts our understandings of contexts in which we undertake research partnership as well as our understandings of ways we negotiate a variety of processes as we engage in research in partnership. This restriction in our understanding of mathematics education research partnerships is important because such understandings have the potential to inform the choices we make at every phase of a research partnership, bolstering its efficacy and functioning. Additionally, knowing more about what we currently do in our research

partnerships can simultaneously uncover what and who we might be missing in our research partnerships.

Methodology

Following the guidance of Peters et al. (2015), and relying on a SPIDER framework (Sample, Phenomenon of Interest, Design, Evaluation, Research Type), to bound our search (Methley et al., 2014), we stipulate these inclusion criteria: peer-reviewed, empirical articles published in academic journals from 1989 until 2023, available in English, and focusing primarily on mathematics education—which excludes studies of STEM research partnership. We focus our time frame from 1989 because the first version of the NCTM Standards was published in 1989 in North America and spurred efforts to examine and reform mathematics teaching and learning. We reason that the publication of the 1989 NCTM Standards motivated many research partnerships, particularly in North America, as stakeholders responded to these new expectations.

Our review process includes 4 phases: (1) Literature Search and Retrieval, (2) Document Screening, (3) Data Collection, and (4) Data Extraction. To search and retrieve relevant studies (phase one), we teamed with a research librarian to conduct an electronic database search of Academic Search Complete, Education Research Complete, and Education Resource Information Center (ERIC-EbscoHost) using the following title/abstract keywords: "research partnership" OR "research collaboration" OR "research association" AND math* OR geomet* OR trigonometry OR algebra* OR operation* OR measurement OR "data analysis" OR probability OR number. For each database, relevant subject headings/descriptors were also employed to guard against overlooking potential articles. This database search was completed in July 2023 and returned a total of 7492 distinct articles across the three databases. With the expertise of our research librarian partner, we are also conducting a re-run of this search and retrieval process to include studies published since July 2023. None of these top-up studies are included in the results and discussion below, as the review process is ongoing for those studies.

Phase 2: Document Screening--in conjunction with our research librarian team member, we used Covidence software to review and analyze articles. Each abstract was screened for relevance by two members of our research team, working independently. When we encountered split decisions for inclusion/exclusions, we employed team consultations with a third team member to resolve the discrepancy. Of the 7492 articles, 6979 were not relevant to our research questions, leaving 513 articles to move forward. During the second phase of document screening, each of these 513 full-

text articles were screened by two independent team members, with conflicts being resolved by consultation with a third team member. At this point, 372 studies were excluded, leaving 141 studies for data extraction. Exclusion reasons included: articles or reports lacking research questions, methods, or analysis (not empirical, n = 146); articles not primarily situated in the field of mathematics education (n = 98); articles whose configuration was not a partnership (study on participants rather than with partners, n = 88); articles whose only mention or evidence of partnership was as a description of the context of the study (n = 29), in the recommendations section (n=4), or the literature review (n=3), rather than in the study research processes; and articles not published in peer-reviewed sources (n = 4).

Phases 3 and 4: Data Collection and Extraction—two of our team members, now using a freely available and shareable database, independently extracted and coded the following information from the 141 included studies: year of publication, publication outlet, scale of partnership, duration of partnership, members of the partnership, partnership ecology/structure, sources of support for the partnership, grade ranges for the partnership, and theoretical framing of the partnership work. Following this process, our two team members resolved any coding discrepancies by consulting with each other until full agreement was reached on all codes for all 141 articles. Coding themes and clusters were constructed inductively and iteratively during the early processes of this phase, as the two coding team members regularly discussed and revised themes.

Results and Discussion

Publication context: Our data extraction revealed that the earliest published paper appeared in 1995, with only two additional papers being published between 1995 and 2002. 44 additional papers that engaged a research partnership in mathematics education were published from 2002 up until 2013. Between 2013 and 2023, 94 more studies were published that met our inclusion criteria. The studies were published in 92 different academic journals; the 17 (18% of the 92 different outlets) of these journals that are specific to mathematics education included: Journal of Mathematics Teacher Education (n = 12), ZDM (n = 7), Educational Studies in Mathematics (n = 3), Journal of Research in Mathematics Education (n = 3), as well as additional journals with 1 or 2 articles, for a total of 44 of the 141 articles, or 31%, being published in mathematics education-focused journals.

Partnership members, scale, and duration: Our inductive coding uncovered 5 groupings for members, including in-service teachers (ISTs) with mathematics teacher educator scholars (MTEs);

combinations of districts and schools with MTEs; combinations of schools, communities, and MTEs; MTEs with mathematicians; and other (ISTs with ISTs, prospective teachers (PSTs) with ISTs and scholars, PSTs with scholars, and MTEs with other MTEs. See Table 1 for the distribution of these 5 groupings. Partnerships between ISTs and MTEs accounted for more than half of all partnerships.

Table 1

Partnership Member Configurations

Configuration	Number of Studies (141)	Percent of Studies
IST/MTE	77	55
School or District & MTE	22	16
School/Community/MTE	18	13
MTE and Mathematician	14	10
Other	10	7

Table 2 provides our findings of the scale or reach of each research partnership. Our inductive coding uncovered 6 increasing geographic spans for the research in which the partnerships engaged: classroom or departmental, single institution, district/community, state/province/region level (to include multiple agencies), national (including multiple agencies across regions), and international. Two studies were unclear about the scope of their partnership. Two-thirds of the research partnerships were contained within a classroom, a single institution or a single district/community.

Table 2

Partnership Scale

Scale	Number of Studies (141)	Percent of Studies
Classroom	16	11
Single Institution	40	28
District/Community	38	27
State/Province/Region	22	16
National	18	13
International	5	4
Unclear	2	1

Table 3 provides findings of the duration of each partnership. 42% of all included studies had a duration of 1 year or less, and 7 of these studies had a duration of one unit of instruction.

Table 3

Partnership Duration

Duration Span	Number of Studies (141)	Percent of Studies
< 6 months	33	23
6 months to 1 year	27	19
1-2 years	19	14
2-4 years	32	23
5-10 years	10	7
Unstated or unclear	20	14

Table 4 shows the grade-level focus of the research that was undertaken during the partnership. Note that this is not the grade bands for the partnership members, but it is the grade of focus for the actual study that was conducted. There were no included studies at the pre-primary level, but it may be important to note that those grade bands were not included in the 1989 NCTM Standards.

Table 4

Grade Level Focus of the Study

Grade Band	Number of Studies	Percent of Studies
Pre-Primary	0	0
Primary/El	41	29
Middle	18	13
Secondary	33	23
Post-secondary (e.g., PSTs)	7	5
Multiple Levels	18	13
Practicing Professional (e.g., IST, faculty)	18	13
Parents/Guardians	1	1
Unclear	5	4

Table 5 shows the sources of support that were cited in the study as a resource for the research partnership. The total is more than 141 because a few partnerships utilized multiple sources of support. Approximately half of the research partnerships (n=70) relied on competitive external funding sources. More than a third of studies either did not cite sources of support for their work or were unclear about the sources of the support the study utilized.

Table 5

Sources of Support

Configuration	Number of Studies (n=147)	Percent of Studies
Funded Grant (External = 70)	74	50
District Sources (funds, time, schedules)	10	7
Data sharing MOU	1	1
Non-specific support statement	8	5
No supports cited or unclear	54	37

In addition to the information conveyed above, we found that 58, or 41%, of the 141 studies conducted research on aspects of the research partnership. Also, there was little to no consensus around a theoretical or conceptual approach to research partnership. 44 (31%) of the studies relied on one of an array of content-focused theoretical frames rather than frames that supported research partnership processes, although one study did rely on frames from organizational psychology to guide their work. The studies that utilized partnership-focused frames included a wide variety of conceptual and theoretical guidance, including Activity Theory, Frame Analysis/Communities of Practice, and Sociocultural learning theories. The most frequent theoretical approach was to not explicitly state one at all, which was the case for 21 studies, or 15%.

Implications

Some of the salient aspects of these data are that 50% of research partnerships rely on external funding sources; that none of the studies partnered with teachers of pre-primary aged pupils; and that more than half of the studies partnered with in-service rather than prospective teachers. These findings point to a dearth of research partnerships engaging pre-primary ISTs as well as with PSTs across grade levels. These findings also point to the importance of external funding to support research partnerships; in an era where such funding for education is diminishing, we as a field may face the necessity of rethinking the scale and scope of the work we do in partnership in order to economize. During our screening process in particular, we were struck by the frequency with which scholars employ terms like partnership unaccompanied by an explanation or rationale for why they perceive their configuration as a partnership. We are heartened by the finding that 41% of the studies conducted research on their partnerships, because those studies can further deepen our understandings of dynamics that undergird how we conduct and negotiate research in partnership in our field.

Resources are important for most partnerships in this study, as 62% of partnerships articulated a variety of supports that sustained them. Given that external funding is not always or equally available, the fact that a half of partnerships received external grant funding to support the partnership functioning indicates a potential impediment for many scholars' access to conduct partnership in research because of the resource-heavy nature of such work. Though districts rarely contribute funds, our study indicates that districts provide other important resources to partnerships, notably substitute teachers for release time, extra planning time, and carefully crafted scheduling to support teachers' work in partnership during a school day.

One final implication is from the finding that scholars employ disparate frames to support their research in partnership, and that almost a half of the research studies conducted in partnership did not employ theoretical or conceptual frames specific to such work; 15% of included studies did not specifically articulate a theoretical frame, and 31% of included studies used content-specific frames rather than also relying on partnership-focused frames. This means almost half of the included studies that conducted research in partnership did not specifically include a framing to support that partnership. The implication we take from this finding is that our field is in the process of seeding the ground with an understanding of the importance of purposeful, thoughtful consideration of partnership processes. Further, there is still a large variety of potential partnership frames which may take root. So, if processes of frame coalescence occur, they will likely follow some time of growth and ripening in our familiarity with conceptions of how, why, and to what extent we engage in research in partnership with fellow interested and engaged persons in mathematics teaching and learning.

References

Arbaugh, F., Herbel-Eisenmann, B., Ramirez, N., Knuth, E., Kranendonk, H., & Quander, J. R. (2010). Linking *research* and *practice*: The NCTM Research Agenda Conference *report*. *National Council of Teachers of Mathematics*.

Casey, M. (2008). Partnership–success factors of interorganizational relationships. *Journal of Nursing Management*, 16(1), 72–83. http://dx.doi.org/10.1111/j.1365-2934.2007.00771.x

Coburn, C. E., & Penuel, W. R. (2016). Research–practice partnerships in education: Outcomes, dynamics, and open questions. *Educational Researcher*, *45*(1), 48–54.

Giles, D., Smythe, E., & Spence, D. (2012). Exploring relationships in education: A phenomenological inquiry. *Australian Journal of Adult Learning*, 52(2), 214–236.

Henrick, E. C., Cobb, P., Penuel, W. R., Jackson, K., & Clark, T. (2017). *Assessing research-practice partnerships: Five dimensions of effectiveness*. William T. Grant Foundation.

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: A comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Services Research*, *14*(1), Article 579. http://dx.doi.org/10.1186/s12913-014-0579-0

National Council of Teachers of Mathematics. (2018). *Catalyzing change in high school mathematics: Initiating critical conversations*. NCTM.

National Science Foundation. (2024, September). NSF 24-555: *Computer Science for All Program Solicitation.* https://new.nsf.gov/funding/opportunities/csforall-research-rpps-computer-science-all/nsf24-555/solicitation

Peters, M. D., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *JBI Evidence Implementation*, *13*(3), 141–146.

Penuel, W. R., & Gallagher, D. J. (2017). *Creating research practice partnerships in education*. Harvard Education Press.

Ruiz, A. (2021). Plenary panel 1: Actors for math teacher education: Joint actions versus conflicts. Paper presented at the *14th International Congress on Mathematics Education*, Beijing, China.

Scherer, J. (2008). A review of instruments to evaluate partnerships in math and science education. *Peabody Journal of Education*, *83*(4), 611–636.

Superfine, B. M., De Voto, C., Superfine, A. C., Cosner, S., & Han, A. (2023). How research-practice partnerships learn to develop goals for math and identify local problems of practice. In E. Anderson & S. D. Hayes (Eds.), *Continuous improvement: A leadership process for school improvement* (pp. 315–335). Information Age Publishing.

SECTION III: INTEGRATED STEM

LEARNING SCIENCE IN ACTION: USING EXAMPLES AND NON-EXAMPLES TO DEEPEN MATH AND SCIENCE UNDERSTANDING

Andrea Foster
Sam Houston State University
asf004@shsu.edu

Kristi Martin
Ball State University
kristi.martin@bsu.edu

Abstract

Strengthening math and science methods courses with learning science principles starts with a simple idea: future teachers learn best when the course itself models the kind of instruction we want them to deliver. The most effective redesigns blend cognitive science, discipline specific pedagogy, and authentic practice so that teacher candidates internalize both why strategies work and how to use them. In this paper we describe how we embed the practice of using Examples and Non-Examples within our math and science methods courses to deepen teacher candidates' and P–12 learners' understanding of math and science principles.

Keywords: cognitive science, examples/non-examples, math & science, preservice teacher preparation

Introduction

How can teachers help students clearly see what a mathematical or scientific concept is—and what it is not? Developing deep conceptual understanding in mathematics and science requires students to build schema that helps them recognize patterns, identify structures, and apply ideas in new contexts. Evidence from cognitive science shows that students must address key features of a concept, process information in working memory, and encode it in long-term memory so it can later be retrieved and applied (Deans for Impact, 2026; Weinstein et al., 2019). Instruction that includes examples and non-examples supports this process by helping students identify both the defining features and conceptual boundaries of an idea. Intentionally varied examples highlight the underlying structure of a concept and help students avoid undergeneralizing when encountering new situations. For example, when learning to add three-digit numbers, students might examine problems requiring regrouping across different place values (e.g., 128 + 287, 356 + 429, 753 + 262), which allows them to examine when to regroup and when regrouping is not necessary. In contrast, non-examples clarify conceptual boundaries and prevent overgeneralization. Examining an incorrectly solved problem,

such as regrouping unnecessarily in the ones place when solving 325 + 682, can prompt students to analyze errors, refine their reasoning, and explain why the solution is incorrect.

Objectives

We will share how we used the cognitive science principle of Deepening Meaning & Learning by teaching with examples and non-examples to support our own teacher candidates' (TCs) learning. We also modeled how they could use the principle in their future classrooms. TCs learned to use this strategy in elementary mathematics and science methods courses. Guided by the Deans for Impact (DfI) Learning by Scientific Design (LbSD) framework (Deans for Impact, 2026), candidates explore three criteria for designing instruction with Examples and Non-Examples: *varied examples highlighting conceptual structure*, c*ontrasting non-examples revealing conceptual boundaries*, and *prompts to elaborate requiring students to compare and explain similarities and differences*. Through content- and pedagogy-based activities—including sorting fraction representations and analyzing classroom vignettes—teacher candidates experience how learning through both examples and non-examples strengthens conceptual understanding and supports equitable learning opportunities by supporting all students to build stronger schema.

Instructional Framework

Our work is grounded in the science of learning and the premise that effective teaching begins with understanding how people learn. Advances in cognitive science have demonstrated that learning is not the passive accumulation of information but the active construction of mental models through attention, working memory, retrieval, comparison, and meaningful practice (Darling-Hammond et al., 2020; Goodwin et al., 2020; Willingham, 2009). Yet, despite decades of research describing how learning occurs, these principles remain underrepresented in many teacher preparation programs. As a result, novice teachers often enter classrooms with limited understanding of the cognitive processes that support deep and durable learning.

The conceptual foundation for this study is the DfI Learning by Scientific Design (LbSD) framework, which translates findings from cognitive science into specific teacher actions that improve instruction. Rather than viewing learning science as abstract theory, the LbSD framework operationalizes research into instructional practices that teacher candidates can learn, apply, and refine (Deans for Impact, 2026). We will focus on the teacher action of Examples and Non-Examples (ENE).

The Examples and Non-Examples teacher action is grounded in research demonstrating that concepts are learned most effectively when learners encounter multiple, carefully selected examples that illuminate a concept's defining features while simultaneously examining contrasting cases that clarify conceptual boundaries (Chen et al., 2023; Große & Renkl, 2007). Rather than relying on a single stereotypical example, teachers intentionally vary examples to highlight the deep structure of a concept while minimizing attention to irrelevant surface characteristics. Contrasting non-examples further strengthen conceptual understanding by helping learners recognize what a concept is *not*, reducing misconceptions and overgeneralizations (Große & Renkl, 2007).

Within the LbSD framework, effective instruction with Examples and Non-Examples includes three essential teacher actions:

1. Provide varied examples that highlight the deep conceptual structure rather than superficial features of the concept.
2. Use carefully selected contrasting non-examples to make conceptual boundaries explicit and surface common misconceptions.
3. Prompt learners to compare, contrast, explain, and justify similarities and differences between examples and non-examples, promoting elaboration and deeper processing.

These instructional actions are consistent with cognitive science principles regarding attention, schema development, retrieval, discrimination learning, and cognitive engagement (Deans for Impact, 2026). Comparing multiple representations encourages learners to identify invariant features across examples while suppressing misleading surface cues. Asking students to explain why an example belongs—or does not belong—to a category requires effortful retrieval and elaborative processing, both of which strengthen long-term memory and transfer (Große & Renkl, 2007).

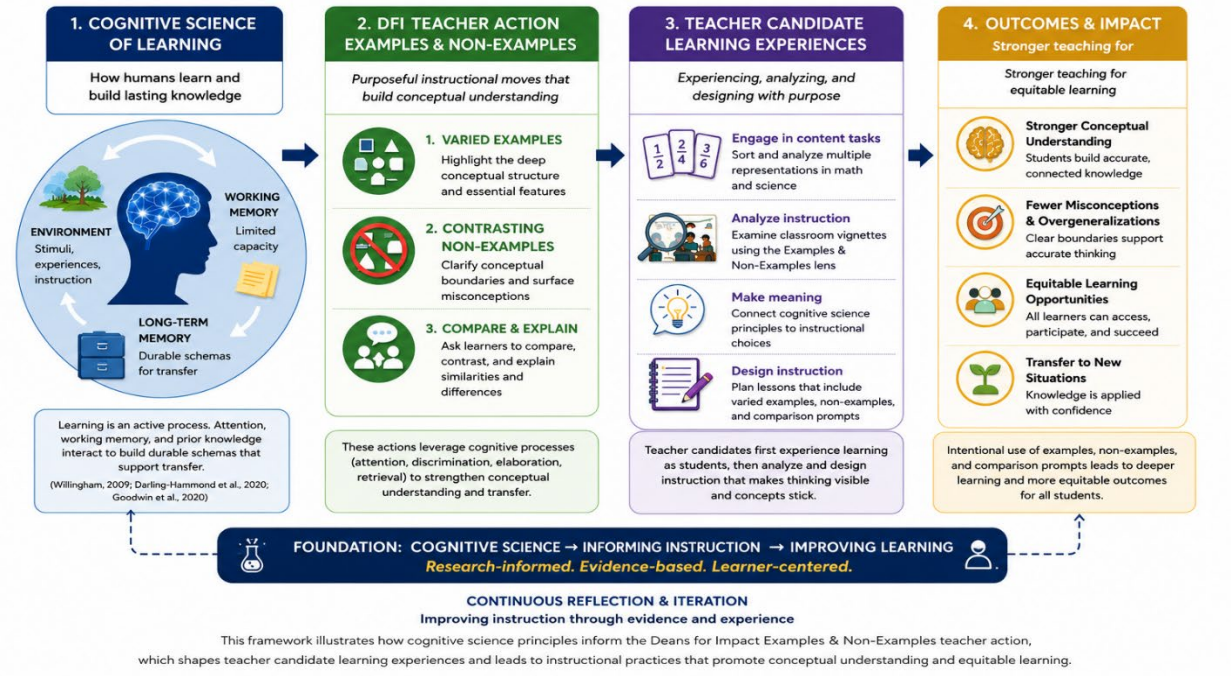
The study also draws upon Willingham's (2009) Model of the Mind, which illustrates the interaction among the environment, working memory, and long-term memory during learning. Because working memory is limited, instructional design must reduce unnecessary cognitive load while directing learners' attention toward the most important conceptual features. Carefully selected examples and non-examples support this process by helping learners encode accurate schemas into long-term memory. Once these schemas are established, learners are better able to retrieve and apply knowledge in novel situations.

Accordingly, this study conceptualizes TCs as learners who first experience Examples and Non-Examples from the perspective of students before analyzing the instructional decisions that produced those learning opportunities. Through mathematics and science methods activities,

candidates engage in sorting tasks, compare multiple representations, analyze classroom vignettes, and design their own examples and non-examples aligned with elementary learning goals. These experiences are intended to deepen candidates' understanding of both the cognitive mechanisms underlying conceptual learning and the pedagogical decisions that support equitable instruction.

Figure 1

Conceptual Framework for Using Examples and Non-Examples to Strengthen Teacher Candidate Learning



Note. Conceptual framework developed by the authors. Visual representation generated with ChatGPT (OpenAI, 2026) from an author-developed prompt requesting a four-part conceptual model connecting cognitive science of learning, Deans for Impact Examples and Non-Examples teacher actions, teacher candidate learning experiences, and outcomes and impact.

The conceptual framework therefore proposes that cognitive science principles provide the theoretical foundation, the DfI Examples and Non-Examples teacher action serves as the instructional mechanism, and teacher candidate instructional design represents the intended outcome. As teacher candidates learn to intentionally select varied examples, design meaningful non-examples, and facilitate opportunities for students to discuss both, they develop instructional practices that foster stronger conceptual understanding and more equitable learning opportunities for elementary students.

Practice/Innovation

When we first introduced Examples and Non-Examples as an instructional strategy, TCs' written exit tickets indicated that they recognized the value of selecting varied examples. Most TCs, however, expressed concern that non-examples might confuse students or interfere with their learning. We used constant comparison (Glaser, 1965) to examine responses across exit tickets and identify recurring patterns in TCs' thinking. We documented candidates' comments from class discussions through instructor observation and notes, since classes were not recorded. The concern evident in the exit tickets prompted us to incorporate activities in which candidates experienced learning from both examples and non-examples before considering how to use the strategy in their own classrooms.

Classroom Examples

For the math methods class, embedding examples and non-examples happened within both content and pedagogy portions of the course. We will share an example of both a content-based activity and a pedagogy-based activity, though many others were embedded throughout the course.

Fractions Card Sort & Vignettes

One example of a content-based activity for TCs to experience learning through examples and non-examples was a card sort activity with potential depictions of ¾. See a sample of the cards in Figure 2. They were asked within their groups to determine whether each card did or did not represent ¾, as well as to be able to explain why they had made that conclusion. Then they were asked to think about how elementary students might think about each card and how it would help them to understand fractions. Finally, they were asked to consider why each card was likely included in the sorting activity, especially for cards that were tricky or difficult for them. TCs recognized that including common misconceptions, such as Model A in Figure 2, would help students to see that pieces need to be the same size, but not shape, and would increase their understanding of how to identify fractions in models. This discussion included thinking about what they learned from the sorting activity and what a similar activity might look like for elementary students. TCs used this knowledge later in the semester by including common misconceptions when they planned lessons.

Figure 2

Sample of Example and Non-Example Sorting Cards



For a pedagogy-based activity, TCs would learn about a teaching practice from Principles to Action (NCTM, 2014), such as supporting productive struggle in learning mathematics. As part of this learning, they would read two vignettes about similar classrooms. In one vignette, the teacher would use the teaching practice of letting students experience productive struggle, and in the other vignette, the teacher would not use the practice. The TCs were first asked to identify which vignette was in their groups and then to discuss how using or not using the practice impacted the learning that happened in each vignette. Discussions with both vignettes were stronger and TCs felt more confident about what each practice was and what it was not than when they had just seen the example vignette in previous semester.

Rocks Rock!

In the science methods course, teacher candidates (TCs) experienced, analyzed, and modified a common fourth-grade lesson on rock classification and formation. They first participated as learners in a *Rock Detective* investigation, adapted from *How Rocks Rock! Processes That Shape Our World* (Harrison, 2023). TCs examined rock samples for observable "clues"—such as layers, crystals, texture, and other physical characteristics—and used evidence to classify each sample as igneous, sedimentary, or metamorphic.

Following the investigation, TCs completed a Modify Examples/Non-Examples Exit Ticket. The targeted learning goal was, *Students will classify rocks based on observable characteristics and explain how those characteristics reflect the processes that form igneous, sedimentary, and metamorphic rocks.* The lesson was situated within ESS2.A: Earth Materials and Systems and NGSS 4-ESS1-1, emphasizing observation, classification, interpretation of evidence, and connections between rock properties and formation processes. Rather than simply evaluating whether the lesson was engaging, TCs examined it through an ENE lens. They were asked to identify: (a) the deep structure students needed to understand; (b) potentially misleading surface features or misconceptions; (c) opportunities to strengthen the

selection and framing of examples and contrasting non-examples; and (d) elaborative prompts that would help all students compare, distinguish, and explain their thinking.

A central deep-structure idea was that rocks are classified by how they formed, not simply by what they look like. Observable properties provide evidence of those formation processes, but individual surface features can be misleading. For example, granite and sandstone may both be light-colored, yet granite contains interlocking crystals while sandstone consists of cemented grains. Without attention to these distinctions, students may overgeneralize rules such as "light-colored rocks are one type" or "rocks with holes are volcanic. Teacher candidates therefore considered purposeful contrasts that could make conceptual boundaries more visible:

Table 1

Examples and Contrasting Non-Examples for Rock Classification

Example	Contrasting non-example	Conceptual Distinction
Granite	Sandstone	Similar color does not indicate the same rock type; crystals distinguish granite from cemented grains.
Pumice	Limestone	Holes alone do not determine rock classification.
Marble	Limestone	Marble forms when limestone is changed by heat and pressure.
Slate	Shale	Metamorphic rocks can form from sedimentary parent rock.
Gneiss	Granite	Both may contain visible crystals, but banding provides evidence of metamorphism.

Note. Contrasting examples and non-examples are used to draw attention to defining characteristics and help learners distinguish the deep structure of rock classification from potentially misleading surface features.

These contrasts encourage learners to move beyond superficial similarities and address the defining attributes and formation processes that distinguish rock types. TCs then reflected on how their modified lesson might produce different learning outcomes than the original lesson. The larger goal of this experience was to prepare future mathematics and science teachers to intentionally select, sequence, and discuss examples and non-examples to support conceptual understanding. By

first experiencing the strategy as learners and then analyzing and applying it as teachers, TCs had an opportunity to connect cognitive science principles with instructional decision-making. Evidence of transfer can subsequently be examined in their own lesson plans, particularly in the purposeful inclusion of examples, non-examples, and comparison prompts designed to make important concepts visible and accessible to learners.

Implications

Experiencing the learning science principle of Examples and Non-Examples as learners enabled TCs to recognize the instructional value of this practice. Rather than viewing examples solely as illustrations of correct responses, participants came to appreciate how varied examples revealed the deep conceptual structure of a concept, while contrasting non-examples made conceptual boundaries explicit and exposed common misconceptions. Because teacher candidates experienced these benefits firsthand, they were better able to envision how intentionally pairing examples with non-examples could support their own future students in developing more accurate and transferable conceptual understanding.

The implications of this work extend beyond mathematics and science methods courses. As teacher educators, we can intentionally model the use of examples and non-examples across a wide range of learning contexts. For example, in a graduate-level research course, one of the authors asked students to analyze both an exemplar and a non-example of a final research paper before beginning their own projects. Comparing the two prompted students to identify the defining characteristics of high-quality scholarly writing while recognizing common weaknesses to avoid. This same instructional approach can be applied across disciplines and educational levels whenever learners are developing new concepts, skills, or professional practices. Examples and non-examples serve as a powerful, research-informed instructional strategy because they help learners construct richer mental models. Varied examples promote abstraction by highlighting the essential features shared across different contexts, while contrasting non-examples clarify the limits of a concept and challenge misconceptions. Together, these complementary instructional moves support deeper conceptual understanding, stronger transfer of learning, and more durable knowledge.

References

Chen, O., Retnowati, E., Chan, B. K. Y., & Kalyuga, S. (2023). The effect of worked examples on learning solution steps and knowledge transfer. *Educational Psychology, 43*(8), 914–928. https://doi.org/10.1080/01443410.2023.2273762

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science, 24*(2), 97–140. https://doi.org/10.1080/10888691.2018.1537791

Deans for Impact. (2026). *The science of learning*. https://www.deansforimpact.org/tools-and-resources/the-science-of-learning

Glaser, B. G. (1965). The constant comparative method of qualitative analysis. *Social Problems, 12*(4), 436–445. https://doi.org/10.2307/798843

Goodwin, B., Gibson, T., & Rouleau, K. (2020). *Learning that sticks: A brain-based model for K–12 instructional design and delivery*. ASCD; McREL.

Große, C. S., & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and Instruction, 17*(6), 612–634. https://doi.org/10.1016/j.learninstruc.2007.09.008

Harrison, J. (2023). *How rocks rock: The processes that shape our world*. Old Dominion University Digital Commons. https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1000&context=k12_sciences_lessonplans

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

Willingham, D. (2009). *Why don't students like school? A cognitive scientist answers questions about how the mind works and what it means for your classroom*. Jossey-Bass.

Weinstein, Y., Sumeracki, M., & Caviglioli, O. (2019). *Understanding how we learn: A visual guide*. Routledge.

Acknowledgements

The researchers would like to acknowledge Deans for Impact and the Learning by Scientific Design Network for supporting educator-preparation programs in preparing novice teachers with a deep understanding of learning science principles and the ability to apply that knowledge to help students learn. For more information, see https://deansforimpact.org/.

INCREDIBLE DECISIONS: INVESTIGATING A TEACHER'S CHANGES TO THEIR LESSON PLAN IN ENACTMENT

Alisa H. Chung
Purdue University
chung443@purdue.edu

Jenna Gist
Purdue University
clingerm@purdue.edu

Rebekah Hammack
Purdue University
rhammack@purdue.edu

Abstract

Teachers must make in-the-moment decisions (IMDs) when teaching as it is virtually impossible for them to predict all possible events that can occur in the classroom. These IMDs can cause deviations from a teacher's lesson plan in both positive and negative ways. By understanding what motivates teachers' responses to unexpected events, teacher educators may help teachers better realize their intended enactments and capitalize on unexpected learning opportunities. This single case study investigates the decision making of one rural elementary teacher. Results showed lack of resources, seizing learning opportunities, and teacher insight to be this teacher's primary decision-making motivations.

Keywords: engineering education, professional development, decision making

Introduction

The classroom is a highly dynamic space where learning manifests as students engage in meaningful dialogues (Coskun et al., 2021a). Teachers can often anticipate the most common ways students may have difficulties or make errors. However, with the many different ideas, questions, suggestions, and reactions from students, it is nigh impossible for teachers to anticipate or plan for all possible events that occur in classrooms (Coskun et al., 2021b; Remillard, 2018). Thus, teachers must make in-the-moment decisions (IMDs) to address these unexpected events which may cause deviations from curricula or their lesson plan (Lloyd, 2019; Remillard, 2018). These divergences are often the result of teachers responding to students' needs and capitalizing on fruitful learning opportunities (Coskun et al., 2021b; Kim, 2018; Remillard & Kaye, 2002). However, for reasons beyond teachers' control, divergences can result in difficulty realizing their intended goals (Clark-Wilson & Hoyles, 2019). By gaining a better understanding of how teachers react to unexpected events and the motivations behind these responses, teacher educators may be able to help teachers better realize their intended enactments and make the most of unexpected opportunities for learning through teacher training and professional learning (PL).

Objectives of the Study

This single case study centers on the enactment and IMDs of one rural elementary teacher, Adam (pseudonym), and seeks to answer the following research questions: (1) What are the sources of discrepancy between Adam's lesson plan and his actual classroom implementation? (2) Which factors influenced Adam's responses to unexpected events in the classroom?

Theoretical Framework and Related Literature

The present study was influenced by the Culturally Relevant Engineering Design (CRED) framework (Bowman et al., 2024), as well as literature on curriculum versus teachers' enactment (Remillard, 2018) and teachers' responses during contingent moments (Coskun et al., 2021a; 2021b). This study seeks to add to the growing body of literature investigating teacher decision-making (Clark-Wilson & Hoyles, 2019; Coskun et al., 2021a; Coskun et al., 2021b; Remillard, 2018).

The CRED Framework

One purpose of the CRED framework is to help educators teach engineering lessons situated in place and local community. The CRED framework situates engineering problem solving within students' experiences and local communities. Working within the framework, students are guided through an iterative process that involves five main phases. After first working to **identify** a community-based need to center their design around, students **describe** the social, cultural, environmental, and historical dimensions of the problem. Informed by community knowledge and values, the students then work to **generate** multiple ideas for solving the problem. Next, students **embody** one of those solutions through a process of building, testing, and refining prototypes. Students then **finalize** their solutions by sharing them with community stakeholders.

CRED lessons are designed to allow students to connect their own interests, observations of their community, and relevant experiences from their own lives to problem-solving using engineering practices. They also encourage inviting community experts into the classroom and involving cultural community as part of engineering learning and teaching. The framework was developed with enough flexibility for teachers to not only adapt their lessons based on place and community but also accommodate for specific classroom contexts such as multi-grade classrooms that require differentiated instruction, or classrooms with multilingual learners. Additionally, the CRED framework was designed for easy integration with the Science and Engineering Practices (SEPs) of the Next Generation Science Standards (NGSS Lead States, 2013).

Curriculum and Enactment

The use of the same curriculum materials does not necessarily result in the same actions or results in the classroom. Remillard (2018) indicated that while different lessons from the same teacher tended to be similar, lessons taught by different teachers given the same curriculum materials varied widely. Ultimately, Remillard (2018) concluded that understanding more about teachers' decisions can provide key insight into factors influencing the quality and fidelity of enacted lessons—a conclusion echoed by other studies examining curriculum versus enactment (Clark-Wilson & Hoyles, 2019; Earnest & Amador, 2019; Kim, 2018; Lloyd, 2019). Adding another layer of complication, the classroom brings many unexpected moments that teachers must respond to. Studies on teacher responses in these unexpected moments conclude that (1) decision making is highly complex and (2) *how* a teacher responds in these moments can have a great effect, positively or negatively, on students' learning and the trajectory of a lesson (Coskun et al. 2021a, Coskun et al., 2021b; Lloyd, 2019; Mason, 2015; Remillard & Kaye, 2002; Wendell & Watkins, 2016). Additionally, they propose PL or teacher training surrounding IMDs may help teachers better notice learning opportunities and respond effectively. Lloyd (2019) found teachers tend to resist self-reflective decision-making practices typically promoted in pre- and in-service trainings regardless of experience level, further highlighting the need for new PL surrounding decision making. Though these previous studies have examined the effects of IMDs and the relationship between curriculum and enactment, literature diving deeper into which specific pressures influence teachers' IMD making is still lacking.

Methodology

Study Context

This study centers Adam, a rural elementary teacher who leads a multi-grade (3^{rd}–5^{th}) gifted and talented class. Adam was part of a larger, multi-year, nationally funded study that examined how ongoing participation in PL affected rural elementary teachers' science and engineering teaching practices. In Year 1, teachers attended a weeklong virtual summer intensive PL and 7 synchronous, ~90-minute online meetings throughout the following school year. The PL introduced teachers to the CRED framework and helped them integrate it with the NGSS and SEPs. Adam was part of a subset of 20 teachers who also participated in an engineering learning community (ELC) focused specifically on how to implement CRED lessons. ELC teachers participated in two additional PL meetings in the spring to provide teachers opportunities to debrief, discuss, and share ideas

regarding the planning and implementation of their CRED lessons. These additional meetings were recorded and transcribed for future reference.

During the PL, teachers were given adaptable lesson plan templates that included sections for state standards, student learning goals, assessment, and guides for the segmenting and flow of the instructional sequence. These lesson segments corresponded to the five phases of the CRED framework: Identify, Describe, Generate, Create, Finalize. This CRED lesson plan served as the basis of comparison for determining where a teacher's classroom enactment was aligned or misaligned with their intended enactment.

Adam's CRED lesson focused on snow removal. The lesson's aim was for students to develop a device that would aid in removing heavy snow that often covers their school and local community. Students were asked to reflect on the previous year's snow days and local snowfall data for the previous decade. Then, in small groups, they brainstormed design ideas on paper. After choosing a design, they built models of their devices from recyclables and craft materials. Then, students presented their models to classmates, peers from a different class, parents, and the school's maintenance team, who typically handles snow removal on campus. After presentations, students were asked to reflect on their model and the feedback they received when sharing about their design.

Data Analysis

Data for this study included Adam's lesson plan; 45 minutes of classroom video footage covering the Create and Finalize stages of the CRED; a 40-minute end-of-year interview; and ELC meeting transcripts. The data was analyzed using thematic analysis (Braun & Clarke, 2008). First, the first author thoroughly examined the lesson plan, as this would serve as the basis of comparison for deviations in enactment shown in classroom videos. Next, classroom video data were open coded for any discrepancies between the teacher's lesson plan and enactment. This included unplanned misalignments (omissions, adaptations, or improvised additions to the CRED lesson plan); contingent moments sparked by classroom situations or student needs/responses; and any explicit mentions in the interview or meeting transcript of a teacher's change in plans or reasoning for a certain action/response. Our definitions for unplanned misalignments and contingent moments drew upon the work of Remillard (2018) and Coskun et al. (2021a). The first author then coded the interview and ELC meeting transcripts for any discussion about or explicit mention of misalignments, the pressures felt in a particular moment, and reasonings behind their response or

approach taken. The entirety of the data set was examined multiple times before the first author began to extract and identify patterns and organize data into themes and sub-themes.

The criteria established by Lincoln and Guba (1985) were utilized to ensure trustworthiness, including the use of a detailed audit trail that comprehensively documented all methodological decisions and data analysis, as well as data triangulation across participant interview, lesson plan, and classroom teaching video.

Results and Discussion

Lack of Resources

The first theme, *lack of resources*, refers to the lack of available resources affecting how Adam guided his students through the lesson and responded to students' questions. This theme was broken into two sub-themes: *lack of materials* and *lack of time*. The sub-theme *lack of materials* refers to the availability of physical materials and monetary funds. During the portion of the lesson where students were building their prototypes, one team informed Adam they had run out of a particular craft material needed to complete their model. In response, Adam suggested an alternative for the students to use instead. Thankfully, Adam's resourcefulness in that moment allowed the lesson to continue smoothly despite the unexpected lack of materials.

The other sub-theme, *lack of time*, was arguably the strongest theme seen throughout Adam's lesson. Available time for this lesson was interpreted as a kind of resource because teachers have limited time in the school day and over the school year to ensure their students achieve their learning goals. Additionally, teachers may not be afforded time to teach engineering during a normal school day given how their school day is structured and must find alternative ways to attain this resource.

In his interview, Adam shared that one of his goals for this lesson was "trying to focus on limiting that time so it doesn't take forever…If we could do one in a couple weeks, then we'd be able to get a lot more opportunities to practice engineering." His lesson plan reflected this, showing an allotment of 20 minutes for each lesson segment. However, Adam had difficulty adhering to this plan in the classroom despite his best efforts. Throughout the lesson, Adam would frequently remind students of the time, even calling out in class saying, "I know how many of you struggle with the amount of time", "it does not need to be perfect; you do not need to spend a ton of time on this", and "we only have 15 minutes for the next activity and *it will go fast*." Even with these reminders, Adam's students struggled to complete some of the tasks within the original 20 minutes, leading Adam to extend some lesson segments at the expense of others. This showed tension that

Adam experienced in wanting to offer his students enough time to fully explore and record their thoughts but having a limited amount of time to complete his lesson. We also saw this tension reflected in a particular moment when a group of students were struggling to add the final details to their model. When the group came to Adam for a solution, Adam, instead of troubleshooting or posing possible alternatives, asked them to consider omitting that design detail since there were only a few minutes of work time left. In his interview, Adam mentioned repeatedly feeling immense time pressure, even stating, "It was hard to limit the time that it took. That was my biggest challenge." The lack of time Adam felt significantly impacted his response to students and the overall flow of the lesson. Without time pressure, Adam may not have placed so much emphasis on working quickly, and instead, taken more opportunities to explore different ideas in depth with his students.

Removal of this time pressure may have also allowed Adam the capacity to fully engage his students and incorporate community relevance into his lesson in a more meaningful way. The CRED framework emphasizes strong community connection and stresses the importance of generating design solutions that both address specific needs and uphold the community's values. The focus on getting things done quickly may have detracted from students' ability to fully engage in meaning-making during the design process and fully considering impacts to community on a deep level. Confining the lesson to only a few days also removed opportunities for students to incorporate knowledge and feedback from community experts while finalizing their designs.

Capitalizing Learning Opportunities

The second theme was *capitalizing on learning opportunities*. Adam utilized certain questions from students or contingent moments in the classroom and turned them into opportunities for his students to connect to different subjects or practice 'soft skills' such as conflict resolution and communication. During part of the lesson, a group of students got into an argument about whose design ideas they were going to incorporate into their model. Adam responded to the situation by asking the students to give each other a moment to calm down, reconvene, and try to work out a compromise. Adam explained his reasoning behind this response in his ELC meeting:

"So, I teach gifted and talented students. They're used to being right a lot, and that comes to group activities as well. So, you know, they always think their idea's the best, and they want what they want. In that first video, the three of them were having a disagreement, and they had kind of come to a resolution that kind of included everybody's ideas in that. And so, that was good to see

because that's a skill we've worked on. It's a skill we have to teach especially to my students, and so it was good to have that there in the engineering process as well."

Given this explanation, it is clear Adam's response was intentional despite the unexpected situation. Instead of mediating directly, he capitalized on the opportunity for his students to practice communication, conflict resolution, and teamwork skills. In a different moment, a group of students expressed having difficulty constructing a more complicated piece of their model. Instead of telling the students what to do or asking students to omit that part of the design, Adam asked the students to generate and consider possible alternatives. Here, Adam saw and capitalized on an opportunity for his students to practice adaptability, resourcefulness, and independence.

Teacher Insight

Lastly, the theme of *teacher insight* came from situations where Adam, drawing on his professional experience, realized that he needed to modify or correct something he had originally planned. In enactment, Adam made two major in-the-moment improvisations to his intended curriculum. The first was when a student asked whether they would be choosing who presents first. Adam explained that his initial plan was to have everyone share their models at once, which was reflected in his written lesson plan. However, after a quick moment to think, Adam said that he changed his mind and he wanted each group to present individually. The second improvisation came when Adam asked his students to write a brief reflection about what knowledge they have gained from getting feedback from group members. A student asked Adam what to do, as he did not work with a group. Adam accommodated this student by adapting the prompt about group work to instead focus on the feedback he gained from asking members from other groups for critiques. These two examples were categorized under the *teacher insight* theme, since both relied on Adam's experience as a teacher to recognize and find effective solutions.

Limitations

Though this investigation into teacher decision making uncovered some important insights, it is limited in two major ways. First, many teachers do not necessarily write every detail into their written lesson plans. As such, we cannot determine with full certainty what enactment was planned for or what student questions Adam anticipated without being able to consult him directly. Similarly, our insight into IMD making is limited as we did not have the opportunity to ask Adam directly about the full rationales behind his responses to students in the classroom. If this study were to be

expanded, further analysis utilizing an additional data set, or conducting additional interviews, including video stimulated recall, with Adam may help to curb these limitations.

Implications

Adam's IMDs reflect intentions of wanting to accommodate students' needs and provide a rich learning environment. He made decisions that made the lesson run more smoothly, addressed lack of materials, and allowed his students to fully engage and practice important 'soft skills'. However, Adam's struggle with time shows how teachers' decision-making is complicated by having limited resources. Curriculum developers should keep in mind factors such as limited class time when creating curricula or PL, noting the approximate time it would take to complete a lesson or lesson segment. Including a list of viable, affordable alternatives alongside an ideal list of materials for each lesson may also be helpful for some teachers. Additionally, teacher educators should help teachers budget time for each lesson, both in terms of the school day and in relation to other lessons over the course of the school year, and in such a way that accounts for their specific classroom contexts. Keeping factors like time constraints in mind during teacher training may also help teacher educators better support teachers in navigating competing pressures when decision-making. Notably, while Adam's lesson is engineering-focused, the tensions uncovered as part of this study are not specific to teaching engineering and could impact any lesson. As such, understanding teacher decision making is particularly important for curriculum developers, and teacher educators alike.

References

Bowman, F., Klemetsrud, B. J., Ozturk, E., Robinson, J., & Lacina, E. (2024). Impact of professional development in culturally relevant engineering design for elementary and middle school teachers (RTP, Diversity). In *2024 ASEE Annual Conference & Exposition*.

Braun, V., & Clarke, V. (2008). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Clark-Wilson, A., & Hoyles, C. (2019). From curriculum design to enactment in technology enhanced mathematics instruction—Mind the gap! *International Journal of Educational Research, 94,* 66–76.

Coskun, S. D., Botsan, M. I., & Rowland, T. (2021a). An in-service teacher's responses to unexpected moments in the mathematics classroom. *International Journal of Science and Mathematics Education, 19,* 193–213.

Coskun, S. D., Botsan, M. I., & Rowland, T. (2021b). Surprises in the mathematics classroom: Some in-the-moment responses of a primary teacher. *Mathematics Teacher Education and Development, 23*(1), 91–112.

Earnest, D., & Amador, J. M. (2019). Lesson planimation: Prospective elementary teachers' interactions with mathematics curricula. *Journal of Mathematics Teacher Education, 22,* 37–68.

Kim, O. (2018). Teacher decisions on lesson sequence and their impact on opportunities for students to learn. In *Research on Mathematics Textbooks and Teacher's Resources*. ICME-13 Monographs. 315–339.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.

Lloyd, C. A. (2019). Exploring the real-world decision-making of novice and experienced teachers. *Journal of Further and Higher Education, 43*(2), 166–182.

Mason, J. (2015). Responding in-the-moment: Learning to prepare for the unexpected. *Research in Mathematics Education, 17*(2), 110–127.

Remillard, J. (2018). Mapping the relationship between written and enacted curriculum: Examining teachers' decision making. In G. Kaiser et al. (Eds.), *Invited Lectures from the 13th International Congress on Mathematical Education*, ICME 13 Monographs.

Remillard, J., & Kaye, P. (2002). Supporting teachers' professional learning by navigating openings in the curriculum. *Journal of Mathematics Teacher Education, 5,* 7–34.

NGSS Lead States. (2013). *Next Generation Science Standards: For states, by states*. The National Academies Press.

Wendell, K. B., & Watkins, J. (2016). Noticing, assessing, and responding to students' engineering: Exploring a responsive teaching approach to engineering design. *ASEE's 123rd Annual Conference & Exposition*. New Orleans, LA.

Acknowledgements

Research presented is based on work supported by the National Science Foundation (NSF) under DRK12 Award #2201249. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF.

ENHANCING THE VALIDITY OF PROBABILITY PROMPTS USING FOUR GENERATIVE ARTIFICIAL INTELLIGENCE TOOLS

Ayse Fethiye Gemci
James Madison University
aysefethiyegemci@gmail.com

Amanda Gantt Sawyer
James Madison University
sawyerag@jmu.edu

Haley Smiley
James Madison University
smileyhaleym@gmail.com

Richard Busi
James Madison University
busirp@jmu.edu

Juliana Galano
James Madison University
galanoje@jmu.edu

Abstract

A Generative Artificial Intelligence (genAI) trainer found that genAI tools could more accurately solve specific probability questions with defined context than others (Sawyer et al.,, 2026). However, it is unclear whether that statement is true. Thus, we verified these findings by investigating four genAI tools (ChatGPT, Gemini, CoPilot, and Claude AI) on their ability to complete the identified probability task to determine when genAI can be trusted to perform probability calculations, investigating what we can learn about these genAI models from their responses and the predictions made by the genAI trainer.

Keywords: generative artificial intelligence, probability, accuracy of artificial intelligence

Introduction

With the "improvements" of Generative Artificial Intelligence (genAI), college students tend to believe that the tools are accurate and produce correct reasoning (e.g., Ding et al., 2023; Sawyer, 2024). These students may not consider that genAI can be wrong, which can lead to overconfidence in the tools. However, research has identified that genAI tools can not only make errors but also misdirect their users (Marcus, 2023). Sawyer et al.'s 2026 study of a genAI trainer found that these AI tools are "people pleasers" (p. 28), not providing the best answer, but creating an answer that the user would like. This has little to do with what is accurate and valid and more to do with what a user might think is true. It ultimately comes down to the fundamental fact that genAI might not know the truth and produce biased responses (Sharma, 2024); thus, we investigated when individuals should trust these tools.

In this paper, we answer this question by investigating mathematical probability questions. In Sawyer et al.'s (2026) study, they identified the following:

> I would trust [genAI tool's response to], "What is the probability of drawing the Queen of Hearts from a standard deck of playing cards on the first try?" This is a simple question with a simple answer, which you will likely find in a textbook. The model likely has training data on how to solve this correctly. However, I would NOT trust the response for something along the lines of, "Three cards are lying face down in front of me. Two are face cards, and there are no spades. No card is next to a card of the same color. What are the odds that the card on the left is the Queen of Hearts?" This complex problem combines logical reasoning, spatial understanding, and probability in a unique way that the model is unlikely to have encountered in its training data. Therefore, applying the proper approach to solving it will be difficult. p. 11)

However, the researchers did not provide any evidence to support this statement in that paper. Thus, we verified these findings by investigating four genAI tools on their ability to complete the identified probability task to determine when genAI can be trusted to answer the following research questions:

1. How accurate are the responses from ChatGPT, Gemini, CoPilot, and Claude AI on the two probability questions posed by the genAI trainer in the Sawyer et al. (2026) study?
2. What can be learned about these four genAI models from their responses to the probability problems?

As a group of preservice teachers (PSTs) and two mathematics teacher educators, we aimed to investigate both the accuracy of genAI tools and the explanations they provide for mathematical processes and reasoning. As preservice teachers pursue licensure in math, English language arts, and social studies, we seek to understand how artificial intelligence will be applied in future classrooms.

Objectives of the Study

As educators who recognize genAI's prevalence in the classroom, when should we trust genAI tools? If multiple genAI tools yield different results with different explanations, we assert that at least one is mathematically inaccurate. We assert that one should always verify the answer for themselves before trusting the genAI, which is a field that requires more research. These results are consistent with the NCTM (2024) position statement on genAI use in education "As new answer-generating tools continue to evolve, mathematics teachers need to ensure that their assessments focus students on identifying the reasonableness of the output and the application of the output in specific application-based situations" (NCTM, 2024, para. 3) However, we recognize that this may

not always be an option for our students, in that the very processes being used might be the reason students feel compelled to use genAI in the first place. Therefore, it is imperative that educators make instructional decisions that provide the groundwork for complex mathematical knowledge students will need to verify the problem's accuracy. Regardless, we attest that if there is any uncertainty regarding genAI's output of a mathematical problem, users should always verify that the answer and genAI's are correct before utilizing the output.

Since the genAI models produced varying responses to the same prompt, we conclude that, initially, individuals should look for similarities across genAI's outputs to learn from their responses. Using our data, we found one commonality across the genAI tools: each model began by extracting information from the question, explaining those points, and simplifying the steps to solve the problem. However, it was from this point forward that the models changed their responses. This becomes a teachable moment within a mathematics classroom. Students can deepen their understanding of how to approach a probability problem by replicating the GenAI Triangulation Process cross-checking results using independent methods or alternative tools and justifying whether an answer is mathematically sound. In research, triangulation refers to using multiple methodologies or data collection techniques to enhance the credibility of the results (Jentoft & Olsen, 2019). "Triangulation is supposed to support a finding by showing that independent measures of it agree with it or, at least, do not contradict it" (Miles & Huberman, 1984, p. 235).

Related Literature

Wu (2023) found that, due to the training process of genAI, it can develop its own biases. These massive datasets are pulled from the internet and other sources, allowing genAI to encode and reinforce biases unintentionally. Additionally, Sawyer & Aga (2024) found that genAI tends to lead users to inappropriate and inaccurate suggestions, which do not consider the audience or the user when providing a response. For example, one genAI platform suggested to a middle school student that analyzing weight loss could help understand functional relationships.

Although many studies highlight the potential instructional benefits of genAI, scholars also emphasize substantial risks and limitations that must be considered when integrating these tools into an educational context. A significant and well-documented concern is algorithmic bias, which emerges from the massive internet-scale datasets used to train LLMs (NCTM, 2024). Wu (2023) demonstrates that these datasets embed societal inequities and can lead genAI systems to develop and reproduce their own biases, which particularly disadvantages minority populations. These biased

patterns not only shape outputs but can also subtly reinforce deficit perspectives if teachers and students accept responses with little or no critique, as Sawyer et al. (2026) assert students do.

Methodology

We used a grounded theory approach (Tie et al., 2019) to answer our research questions about the accuracy of responses from ChatGPT, Gemini, Copilot, and Claude AI to two probability questions from a non-expert perspective, and to understand these four genAI models. Grounded theory "sets out to discover or construct theory from data, systematically obtained and analyzed using comparative analysis" (Tie et al., 2019, p. 1). We selected this approach because genAI is changing so rapidly that we wanted to allow the data to speak for itself; we took an unbiased point of view and a willingness to learn from these tools. Thus, on September 17, 2025, we conducted a document analysis (Bowen, 2009) on the responses received from each genAI model (ChatGPT 5.1, Gemini Flash 2.5, CoPilot GPT-5, and Claude Sonnet 4.5 AI) on the same day using the same prompts:

1. What is the probability of drawing the Queen of Hearts from a standard deck of playing cards on the first try?
2. Three cards are lying face down in front of me. Two are face cards, and there are no spades. No card is next to a card of the same color. What are the odds that the card on the left is the Queen of Hearts?

Zamfirescu-Pereira (2023) identified that non-experts would most likely use a free public version of the genAI tool; thus, we selected the free public versions of the genAI in this investigation to mirror that experience with the most commonly used genAI on the market as of the date of the study (Titan Technology, 2025). Also, we conducted a single round of responses for each tool to simulate how a non-expert would use the tools in the real world. We are aware that this approach limits the generated responses, and it should be seen as a limitation of the study; however, it mirrors the non-expert perspective we took as investigators of the probability problems. After submitting the prompt to the genAIs, we analyzed the responses using the constant comparative method (Creswell & Creswell, 2017), first for accuracy and then for themes across the platforms.

ChatGPT, Gemini, CoPilot, and Claude AI were asked to determine if they could solve the problems posed by the genAI trainer. The data for this investigation were the genAI responses to the two probability questions. Asking a genAI tool to answer a question without providing context or other information is referred to as zero-shot prompting (Xu et al., 2022). This is one of the most

common prompts that beginner users or non-AI experts implement (Zamfirescu-Pereira et al., 2023); thus, we wanted to obtain the response of the genAI from this beginner perspective. We organized, analyzed, and derived meaningful insights from unstructured data, including descriptive data from the genAI responses.

To answer the first question on accuracy, we asked three faculty members at a Mid-Atlantic University, all of whom held a doctorate in mathematics, to determine the answer independently. This included a full Professor and two Associate Professors of Mathematics, all of whom had taught for over 10 years at the university, specifically teaching statistics courses. Once they gathered their responses, they compared the results and discussed differences in their own responses until a consensus was reached on the agreed-upon solution. This validated solution was used to compare with the genAI responses to determine accuracy. Then, each genAI response was analyzed to determine how it deviated from the validated solution and in what ways it was mathematically inaccurate.

Next, the responses were reviewed and coded for specific major patterns, meanings, and discrepancies across the genAI responses to answer the second research question. Each open-ended response was analyzed to determine whether its code fit an existing theme. When the data presented new patterns, emerging themes were constructed from participants' responses (Creswell & Creswell, 2017) to address the two research questions posed above. We acknowledge that this study is a single case study, and we utilized free versions of the GenAI platforms, which may have influenced our solution set.

Results and Discussion

Through this investigation, we found that the genAI trainer from Sawyer et al. (2026) study was accurate in its assessment of both probability problems. First, we asked each genAI tool the question "What is the probability of drawing the Queen of Hearts from a standard deck of playing cards on the first try?" All four of the genAIs completed this task with no errors. Finding the value to be 1/52. We speculate that each of these genAI models was able to complete this task because it was a straightforward question that could have been in their training database about probability with cards.

The second question was more difficult for the genAI to solve, as predicted, with all four genAI tools producing different responses, as seen in Table 1. The second question posed by the

genAI trainer required considerable critical thinking and caution; thus, all four genAI platforms provided different answers with varying reasoning.

Table 1

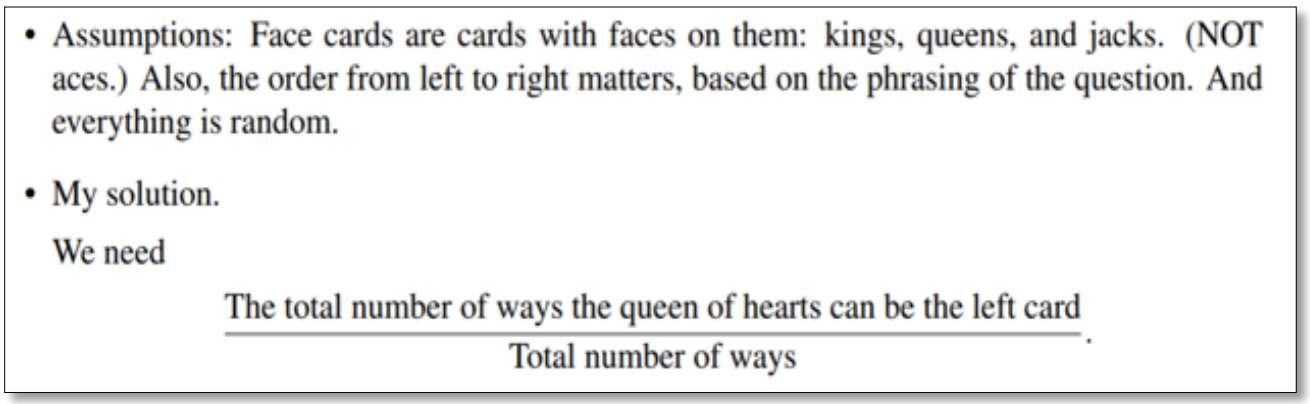
GenAI Numerical Values for Odds of Getting Queen of Hearts

GenAI tool	Answers	Correctness
Gemini	0	✗
Claude AI	1/9	✗
Copilot	Between 12.5% and 16.7%	✗
ChatGPT	11/150	☑

Through this investigation, we found that the genAI trainer from Sawyer et al.'s (2026) was accurate in its assessment. The first problem was easier for genAI, so the result could be trusted more, whereas the second was more difficult for genAI to solve. There was no doubt that the second question posed by the genAI trainer would be tricky, requiring considerable critical thinking and caution. An unintended consequence of this investigation was that we also questioned ourselves more about whether the mathematical problem was even solvable, given that Gemini stated it was a trick problem. Individuals at the top of their field in mathematics questioned their own understanding when provided a solution that did not coincide with their own thinking. This made us stop and realize why so many teachers and students would surrender their authority over mathematics to the computers. For example, the mathematic professors created a framework to solve the problem, however they debated the problem until they came to the same conclusion as ChatGPT.

Figure 1

Math Professor's Response (Van Wyk, personal communication, September 9, 2025)



- Assumptions: Face cards are cards with faces on them: kings, queens, and jacks. (NOT aces.) Also, the order from left to right matters, based on the phrasing of the question. And everything is random.
- My solution.

We need

$$\frac{\text{The total number of ways the queen of hearts can be the left card}}{\text{Total number of ways}}.$$

Less experienced mathematicians may experience even more self-doubt then those who have received PhDs, therefore might have larger impact on their reliabilities of the tools. For the first probability question provided to the genAI, we found all tools were able to answer the question in the same way. The data highlights that mathematically accurate responses can be generated by genAI tools. For the first problem, we also noticed that mathematically accurate responses were created by three or more genAI tools. We speculate that if multiple genAI tools (i.e., 3 or more) yield the same response with the same explanation, which can be verified through mathematical understanding of the problem, then the solution could be trusted. We are calling this GenAI Triangulation Process. For this investigation, our data indicates that the GenAI Triangulation Process could also support the credibility of a response if three or more tools created the same response to a mathematics problem. We are aware that GenAI tools can share similar hallucinations; therefore, we suggest that multiple GenAI tools producing the same response may be more accurate than a single tool's answer.

Mathematics teachers can model this GenAI Triangulation Process for students by inputting the same prompt and examining how multiple genAI tools begin the process of solving. Teachers can note how the models yield different answers as well as how the processes to get to the answers differ, allowing for critical evaluation and discussion of mathematical reasoning. Students then have the unique opportunity to analyze genAI's processes alongside their own methods, providing students with an additional opportunity to evaluate and discuss mathematical reasoning.

Implications

Teachers are being forced to consider genAI's role in classrooms now more than ever. In modern-day classrooms, students have become increasingly familiar with the use of genAI and understand that it can be used in the classroom, whether they agree with it or not. Therefore, this investigation has implications for students and teachers in mathematics classrooms, specifically for determining the validity of solutions.

The use of genAI by students, both inside and outside the classroom, affects how they learn, interact with others, and develop new skills. For some, genAI is a helpful tool to enhance learning, while others view it as a tool used for cheating. As educators, it is important to acknowledge the growth of technology and find ways to ensure students can validate their solutions. In this vein, a major finding of this study is that utilizing one AI response to a mathematical question is problematic. For simple questions, each AI platform will likely get the same response, but as

questions become more complex, the user must find ways to validate responses provided by AI tools. Being a critical consumer of AI-generated mathematical solutions is paramount in teaching and learning.

A growing number of schools have implemented guidelines and policies that strictly prohibit students from accessing genAI technology. Because students have become familiar with genAI and know its abilities, they can find ways around these to continue using it. Some students, for example, use the genAI chatbot on the Snapchat app on their phones to help them complete math problems (Sawyer et al., 2025). Thus, instead of discouraging the use of the tools in classrooms, teachers could critically encourage learners to engage with genAI tools to build their mathematical knowledge.

It is important for teachers to take into consideration how genAI is being rapidly intertwined with education, specifically in math classrooms. Because it has developed so rapidly since 2022, there have been inconsistencies in how educators should, or should not, implement genAI in their classrooms. While some teachers do not use it simply because they are focusing on other resources, others feel as though they are not properly prepared to use it in the classroom with students (Stripling, 2025). Ongoing research is being conducted to find ways to support educators as they consider the use of AI in the classroom. We suggest that teachers be provided with more information about how to use genAI and when to trust the tools to promote responsible use in the classroom.

References

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.

Ding, L., Li, T., Jiang, S., & Gapud, A. (2023). Students' perceptions of using ChatGPT in a physics class as a virtual tutor. *International Journal of Educational Technology in Higher Education, 20*(1), Article 63. https://doi.org/10.1186/s41239-023-00434-1

Jentoft, N., & Olsen, T. S. (2019). Against the flow in data collection: How data triangulation combined with a 'slow' interview technique enriches data. *Qualitative Social Work*, *18*(2), 179–193. https://doi.org/10.1177/1473325017712581

Marcus, G. (2023). *""Math is hard" — if you are an LLM – and why that matters.* Marcus on AI. https://garymarcus.substack.com/p/math-is-hard-if-you-are-an-llm-and

Miles, M. B., & Huberman, A. M. (1984). Drawing valid meaning from qualitative data: Toward a shared craft. *Educational Researcher, 13*(5), 20–30. https://doi.org/10.3102/0013189X013005020

Sawyer, A. G. (2024). Artificial intelligence chatbot as a mathematics curriculum developer: Discovering preservice teachers' overconfidence in ChatGPT, *International Journal on Responsibility. 7*(1), Article 1. https://doi.org/10.62365/2576-0955.1106

Sawyer, A. G., Christian, M., & Sutherland, P. (2026). Educators' discussions with a generative artificial intelligence trainer. *Innovations in Pedagogy and Technology*, *2*(3), 42–53. https://doi.org/10.63385/ipt.v2i3.101084

Sawyer, A., & Aga, Z. (2024). *Counterexamples to demonstrate artificial intelligence chatbots' lack of knowledge in the mathematics education classroom.* AMTE Connections. https://amte.net/connections/2024/05/connections-thematic-articles-artificial-intelligence-mathematics-teacher

Stryker, C., & Kavlakoglu, E. (2024). *What is artificial intelligence?* IBM. https://www.ibm.com/think/topics/artificial-intelligence

Stripling, J. (2025, November 19). *Using AI without (really) cheating.* The Chronicle of Higher Education. https://www.chronicle.com/podcast/college-matters-from-the-chronicle/using-ai-without-really-cheating

Tie, Y. C., Birks, M., & Francis, K. (2019, January 2). Grounded theory research: A design framework for novice researchers. *SAGE Open Medicine, 7*. https://doi.org/10.1177/2050312118822927

Wu, G. (2023, September 17). *8 big problems with OpenAI's ChatGPT*. Make Use Of. https://www.makeuseof.com/openai-chatgpt-biggest-probelms/

Xu, H., Chen, Y., Du, Y., Shao, N., Wang, Y., Li, H., & Yang, Z. (2022). ZeroPrompt: Scaling prompt-based pretraining to 1,000 tasks improves zero-shot generalization. *Findings of the Association for Computational Linguistics: EMNLP 2022*, 4235–4252. https://doi.org/10.18653/v1/2022.findings-emnlp.312

Zamfirescu-Pereira, J. D., Wong, R. Y., Hartmann, B., & Yang, Q. (2023). Why Johnny can't prompt: How non-AI experts try (and fail) to design LLM prompts. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–21. https://doi.org/10.1145/3544548.3581388

FOSTERING CONNECTION: USING AI TO DEVELOP TEACHER CANDIDATES' DISCERNMENT AND REFLECTIVE PRACTICE

Thomas Roberts
Bowling Green State University
otrober@bgsu.edu

Holly M. Plank
Bowling Green State University
plankh@bgsu.edu

Amber Lawson
Bowling Green State University
amberdl@bgsu.edu

Kate Brodeur
Bowling Green State University
brodeuk@bgsu.edu

Karen Wilhelm
Bowling Green State University
wilhelk@bgsu.edu

Stefanie D. Livers
Bowling Green State University
slivers@bgsu.edu

Abstract

As the use of artificial intelligence becomes more prevalent in education, it is crucial for teacher candidates to receive support in developing discernment in its use in teaching and learning in elementary classrooms. This qualitative case study conducted with 27 elementary teacher candidates in the US Midwest examined their interpretations of artificial intelligence-generated feedback they received on their teaching practice. The findings suggest that teacher candidates in the "impact" stage of Fuller's Concern Theory leveraged AI as a pedagogical thought partner, while those in "self/task" stages focused on logistics. Implications suggest that teacher educators should scaffold AI-enhanced reflections to support teacher candidates' professional growth.

Keywords: artificial intelligence, teacher education, reflective practice

Introduction

Artificial Intelligence (AI) is transforming teaching and learning (UNESCO, 2023). Some suggest AI can be used to increase teachers' efficiency (e.g., provide quicker feedback, draft lesson plans), but AI may not be appropriate for student use if it replaces natural struggles that come during learning (Cutler, 2025). Others argue AI has the potential to enhance learning through personalized learning opportunities (Merino-Campos, 2025). Teachers report varying preferences for AI roles depending on the school level (Seo et al., 2024), with many preferring to think of AI as a tool to support the teacher's role, rather than replace it (Cukorova et al., 2019). The role of AI in teacher preparation is an emerging field; therefore, it is imperative that teacher preparation programs build teacher candidates' (TC) discernment with when and how to use AI (Redmond-Sanogo et al., 2024). This is particularly relevant in mathematics and science teacher education, where TCs must attend to content-specific pedagogy and AI tools (Redmond-Sanogo et al., 2026).

Objectives of the Study

The purpose of this study is to understand how elementary TCs view AI-generated feedback they received on teaching episodes of their teaching practice. The two research questions are: (1) How do TCs interpret and discern the quality of AI-generated feedback on their teaching? (2) How are TCs' interpretation of feedback connected to their stage(s) of concern?

Theoretical Framework and Related Literature

Fuller's Concern Theory (Fuller, 1969) frames this study as we explore how TCs integrated AI-generated feedback and how TCs discern the quality of AI-generated feedback. Fuller (1969) identified three stages in which TCs' concerns progress throughout teacher preparation: concern with self, concern with task, and concern with impact. TCs' concerns change over time and can overlap stages as they grow professionally (Livers et al., 2025; Fuller, 1969).

Teacher Candidates' Reflective Practice

Reflecting on practice can transform TCs' development (Juma, 2024) and is a critical component of being an effective teacher (Schön, 1983). Therefore, it is important for teacher preparation programs to help TCs develop their reflective practice (AMTE, 2017; NSTA & ASTE, 2020). TCs can confront their own assumed understandings of teaching and learning through reflection (Livers et al., 2021). Initially, TCs are most likely to be concerned with students liking them and in completing assignments for coursework (Fuller, 1969). As TCs mature and become more prepared for assuming the role of the teacher, these concerns evolve to address the TC's impact on their students (Fuller, 1969; Livers et al., 2021). Structured opportunities to reflect, particularly when connected to fieldwork, are critical to support the progression of TCs to the impact stage of concern (Pino-Fan et al., 2022).

Artificial Intelligence as a Tool for Teacher Reflection

AI can provide academic benefits for students in STEM education (Xu & Ouyang, 2022) and benefit teachers in the instructional cycle (Celik et al., 2022). For TCs, AI can help improve instructional practice (Jamal, 2023). However, teacher educators must help TCs make sense of the benefits and challenges of using AI (Redmond-Sanogo et al., 2024). AI can enhance TCs' reflective practice (Plank et al., 2026; Zawicki-Richter et al., 2019). AI-enhanced reflection can also increase the quality of TCs' reflections, as TCs move beyond lower-level identification in reflection to deeper reasoning of decision-making (Cai et al., 2025). AI feedback has assisted TCs with noticing specific

aspects of their teaching practice (Jacobs et al., 2022). Elementary TCs appreciate the immediacy of feedback from AI-enhanced reflections (Livers et al., 2025). Unfortunately, elementary TCs' use of AI is understudied as elementary teacher preparation programs have not incorporated AI into the curriculum (Grover, 2024).

Methodology

We used a single-case study design (Yin, 2018) to qualitatively explore how a specific group of TCs interpreted and discerned AI feedback.

Participants

TCs from an inclusive elementary education program participated in this study. All students were enrolled in the same program at a large, public regional comprehensive university in the Midwestern United States. The participants, like the institution, were predominantly White. 27 TCs enrolled in a junior-level seminar course gave consent to participate in the study. Twenty-five participants identified as female and two identified as male. The course in which participants were enrolled supports an early field experience focused on small group instruction in reading and mathematics. Through the course and field experience, TCs spend three half-days per week in the field for 10 weeks leading their small group's instruction and assisting the classroom teacher with other tasks.

Data Collection and Analysis

Data for this study consisted of TCs' reflections of AI feedback, midterm reflection, and final reflection. The reflections of AI feedback consisted of four assignments in which TCs recorded themselves during their small group instruction and uploaded a 5–8-minute video to GoReact, a video assessment and skill-mastery platform. All four reflections received AI-generated feedback. The first two focused on relational development and pacing. The final two focused on delivering feedback and assessment. After submitting their videos and receiving feedback, TCs completed a reflection in which they identified strengths and areas for growth related to the skills being measured, described why they did or did not agree with the feedback, and justified their next steps based on the feedback. In addition to the reflections of AI feedback, TCs completed a midterm and final reflection in which they reflected on progress toward their professional goals, identified useful learning experiences and challenges faced, and explained how they use AI. All data were de-identified prior to analysis. TC names used in the results are pseudonyms.

We used a naturalistic inquiry, phenomenological approach to examine TCs' interpretations of discernment of quality of feedback. Naturalistic inquiry allowed for multiple realities to be created by the TCs (Lincoln & Guba, 1985), while the phenomenological approach allowed us to focus on the "lived experience" of TCs receiving and interpreting feedback (Creswell, 2014). Open coding (Saldaña, 2016) was initially used to understand how TCs interpreted and discerned the quality of AI generated feedback. After data reduction through open coding, we engaged in second cycle coding through pattern coding to develop major themes from the data (Miles et al., 2014).

The first author coded midterm and final reflections using a priori coding (Saldaña, 2016) to establish the type of concern expressed by TCs. The three codes used were self, task, and impact taken from Fuller's Concern Theory (1969). The stages of concern were then matched with TCs' feedback interpretations to explore connections between stages of concern and how TCs interpret feedback. Our research team verified these analyses.

Results and Discussion

The first theme we determined was TCs' viewed AI-generated feedback as an objective evaluator in the reflective process. In particular, TCs noted the AI feedback removed their own biases from reflections. For example, Kayla highlighted how the AI-generated feedback provides "an opportunity to assess my lessons without my own bias infiltrating the feedback. It makes me rewatch and reflect on my teaching, and determine what ways I can improve" (Final Reflection). Other TCs recognized areas for growth through the AI feedback. Ava noted the feedback "captured my relational development well… I am sometimes impatient with them, I won't explain kindly why something is not okay" (Reflection of Video Feedback 1). The AI feedback also helped TCs recognize their strengths. For example, Katie noted the benefits of the AI feedback because the TC can be "their own worst critic and miss their own strengths" (Final Reflection). With TCs interpreting AI feedback as more objective, it provided opportunities for them to critically reflect on their practice to identify both strengths and areas for growth.

A second theme that we determined was that TCs coupled AI feedback and the use of AI more broadly through the lens of professional expertise. For example, TCs recognized AI as a tool to support their reflection and work, not as a source of authority. Sarah noted that AI feedback is helpful for "extension ideas" that TCs can choose how to integrate (Final Reflection). Chris shared a similar sentiment of using their professional judgment with AI feedback noting, "I do not think the AI accurately captures my relational development… because I really did not know the student

well… yet it did not give me any areas for improvement" (Reflection of Video Feedback 1). TCs acknowledged how AI could be helpful for "brainstorming lesson ideas, rewording explanations, and checking whether my instructions were clear and age appropriate" (Final Reflection). At the same time, TCs did not relinquish their professional knowledge to AI outputs. As Kim explained, "I am with kindergarteners, so it can be hard to ask the types of questions AI suggested… because of their developmental level" (Reflection of Video Feedback 3). Recognizing the benefits and filtering AI-generated feedback and materials through their professional knowledge demonstrated a high level of discernment in the use of AI.

A third theme we determined was frustration and disapproval of AI usage. Some TCs disapproved of the feedback they received due to technical errors. For example, Katie noted, "There were some instances where the AI said a student said something or I said something when that was not the case" (Reflection of Video Feedback 2). Others noted AI's limitations in understanding the contexts in which instruction occurred. Laura commented, "I overall find it challenging to take some of the negative feedback AI gives me because it cannot understand the limitations or the context I am working in" (Reflection of Video Feedback 2). Some TCs also filtered their interpretation of AI feedback through ethical concerns about AI more broadly. Katie expressed concerns about "not overusing [AI] and having it do assignments for me or creating ideas for me" (Midterm Reflection). Jessica, meanwhile, acknowledged how AI is helpful, but disagreed "with using AI on an ethical basis. I believe it's hurting our environment, stealing from writers and artists, and limiting human creativity" (Final Reflection). The negative perspectives about AI feedback reveal TCs' concern for receiving accurate, actionable feedback that is appropriate for their situation. These perceptions also serve as a reminder that beliefs, in this case beliefs about AI, are a powerful mediator in how feedback is received (Thompson, 1992).

When connecting the TCs' interpretations of AI feedback with their stages of concern, we found significant connections. TCs classified in the self and task stages often viewed the AI feedback through a lens of compliance and logistics. Typically, these TCs agreed with AI feedback that validated their practices and disagreed with AI feedback that questioned their practices. For example, Bailey found the feedback very valuable, highlighting logistical improvements like "material preparation" as a way to improve pacing (Reflection of Video Feedback 1). Katie was focused on "being more firm" with students and rejected AI feedback that offered other strategies because it did not understand the "limitations or context I am working in" (Reflection of Video Feedback 2).

While TCs in this stage exhibited discernment of the quality of feedback, that judgment was filtered through their stage of development's focus on self and tasks.

TCs classified in the impact stage of concern more often interpreted AI as a pedagogical thought partner. Their overall focus was using AI feedback to help them better understand student learning. Kim displayed this in her critique of the questions AI feedback suggested not being developmentally appropriate for kindergartners. Chris identified similar issues when he recognized gaps in his knowledge of students but the AI feedback did not identify areas for growth. Julie offered another example of how TCs in the impact stage of concern leveraged AI. She explained, "I would try and use AI to come up with possible responses I would receive for my planned questions, with this I would just practice how I would respond" (Final Reflection). These examples all demonstrate TCs critically engaging with AI outputs with a focus on student learning. This engagement includes using AI to mimic student responses to prepare for lessons and to critique limitations of AI usage if it did not provide relevant information related to student learning.

Discussion

The findings suggest that AI-generated feedback can serve as a tool that mirrors a TC's current stage of concern (Fuller, 1969). Previous research highlights the efficiency of artificial intelligence (Celik et al., 2022). However, this study shows that the utility of AI's efficiency is ultimately mediated by the TC's professional maturity. The results indicate that TCs in the self and task stages largely leveraged AI primarily for logistical validation. When they did this, they often rejected constructive feedback if it was perceived to challenge their ability to survive as a novice teacher in their future classroom. On the other hand, TCs in the impact stage were able to demonstrate "deeper reasoning" described by Cai et al. (2025), using AI to simulate their students' responses and critique any pedagogical gaps. Overall, these findings suggest that AI cannot be a "one-size-fits-all" tool in teacher education. AI's effectiveness in teacher preparation depends on the TC's ability to move from a compliance-based reflection to inquiry about their own impact on students. When TCs described AI's lack of understanding of their context, they were elevating the need for teacher educators to be part of the feedback process (Redmond-Sanogo et al., 2024). This is particularly important for mathematics and science teacher educators who bring the content specific knowledge for teaching to the feedback process (Redmond-Sanogo et al., 2026). AI alone will not work. Therefore, AI is more of a complement rather than a replacement for the teacher educator in the feedback loop.

Implications

TCs' use of AI is more nuanced than generalized headlines expressing "near universal concern" about AI usage among college students (College Board, 2026). Many TCs identified potential challenges with using AI, including technical errors (e.g., misattributing speech), environmental concerns, and ethical concerns (e.g., stealing resources, out-sourcing thinking). TCs also highlighted ways in which they used AI to complement their instruction, noting positive benefits when using AI to brainstorm ideas, predicting student responses to questions, and getting feedback quickly. This aligns with Redmond-Sanogo et al.'s (2024) suggestion that teacher preparation programs must help TCs discern when and how to use AI effectively. The use of structured reflections of AI-generated feedback in this study achieves this goal as TCs consistently engaged with the feedback to agree or disagree based on their professional knowledge and stage of concern. The connection between stage of concern and how TCs viewed the feedback is also informative for supporting TCs' development. Specifically, understanding that TCs in the self and task stages are more likely to agree with positive and disagree with constructive feedback, teacher educators can scaffold reflections to help TCs interrogate their assumptions and beliefs with more evidence.

There are limitations to this study. For example, the study includes a small sample of TCs attending one teacher preparation program in the US Midwest. Since the findings reflect the experiences of TCs who share the same demographics, future research should be conducted to examine the experiences of TCs with varied demographics attending other teacher preparation programs across the country, as the findings may vary in relation to more diverse populations and institutions. Additionally, future research should focus on content-specific practices to inform more targeted instructional implications in mathematics and science teacher education.

References

Association of Mathematics Teacher Educators. (2017). *Standards for preparing teachers of mathematics*. https://amte.net/standards

Cai, J., Lu, L., Han, B., Wong, L. H., & Gu, X. (2024). Exploring pre-service teachers' reflection mediated by an AI-powered teacher dashboard in video-based professional learning: A pilot study. *Educational Technology Research and Development*, *73*, 1129–1154. https://doi.org/10.1007/s11423-024-10442-1

Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends*, *66*(4), 616–630. https://doi.org/10.1007/s11528-022-00715-y

College Board. (2026). *New College Board research: Faculty express near-universal concern that student AI use undermines original writing and critical thinking*. https://newsroom.collegeboard.org/new-college-board-research-faculty-express-near-universal-concern-student-ai-use-undermines

Cukurova, M., Miao, X., & Brooker, R. (2023). Adoption of artificial intelligence in schools: Unveiling factors influencing teachers' engagement. In N. Wang, G. Rebolledo-Mendez, N. Matsuda, O.C. Santos, & V. Dimitrova (Eds.), *Artificial intelligence in education*. AIED 2023. Springer. https://doi.org/10.1007/978-3-031-36272-9_13

Cutler, D. (2025). *Why using AI is fine for teachers but not for students*. Edutopia. https://www.edutopia.org/article/why-students-should-not-use-ai/

Fuller, F. F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal*, *6*(2), 207–226. https://doi.org/10.3102/00028312006002207

Grover, S. (2024). Teaching AI to K–12 learners: Lessons, issues, and guidance. *Proceedings of the 55th ACM Technical Symposium on Computer Science Education Vol. 1 (SIGCSE 2024), March 20–23, 2024*, Portland, OR. https://doi.org/10.1145/3626252.3630937

Jacobs, J., Scornavacco, K., Harty, C., Suresh, A., Lai, V., & Sumner, T. (2022). Promoting rich discussions in mathematics classrooms: Using personalized, automated feedback to support reflection and instructional change. *Teaching and Teacher Education*, 112, Article 103631. https://doi.org/10.1016/j.tate.2022.103631

Jamal, A. (2023). The role of artificial intelligence (AI) in teacher education: Opportunities and challenges. *International Journal of Research and Analytical Reviews*, *10*(1), 139–146.

Juma, A. A. (2024). Self-reflection in teaching: A comprehensive guide to empowering educators and enhancing student learning. *International Journal of Science and Research Archive*, *12*(1), 2835–2844. https://doi.org/10.30574/ijsra.2024.12.1.1113

Lincoln, E., & Guba, I. (1985). *Naturalistic inquiry*. Sage.

Livers, S. D., Plank, H. M., & Roberts, T. (2025). Using artificial intelligence to foster and expand teacher candidates' reflective practice. In K. Anderson-Pence & A. Ray (Eds.), *Proceedings of the 124th annual convention of the School Science and Mathematics Association* (pp. 130–138). Fort Worth, TX: SSMA. http://www.ssma.org/assets/SSMA%202025%20Proceedings.pdf

Livers, S. D., Zhang, S., Davis, T. R., Bolyard, C. S., Syndor, J., & Daley, S. (2021). Examining teacher preparation programs' influence on elementary teacher candidates' sense of preparedness. *Teacher Education Quarterly*, *48*(3), 29–52.

Merino-Campos, C. (2025). The impact of artificial intelligence on personalized learning in higher education: A systematic review. *Trends in Higher Education*, *4*(2), Article 17. https://doi.org/10.3390/higheredu4020017

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.

National Science Teaching Association & Association for Science Teacher Education. (2020). *2020/NSTA/ASTE standards for science teacher preparation*. http://static.nsta.org/pdfs/2020NSTAStandards.pdf

Pino-Fan, L. R., Castro, W. F., & Moll, V. F. (2023). A macro tool to characterize and develop key competencies for the mathematics teachers' practice. *International Journal of Science and Mathematics Education*, *21*(5), 1407–1432. https://doi.org/10.1007/s10763-022-10301-6

Plank. H. M., Livers, S. D., & Roberts, T. (2026). Preservice elementary teachers' perceptions of integrated STEM after participating in an integrated STEAM course. *Education Sciences*, *16*(2), Article 214. https://doi.org/10.3390/educsci16020214

Redmond-Sanogo, A., Maiorca, C., Roberts, T., Ivy, J., & Burton, M. (2024). Navigating the artificial intelligence landscape: Implications for mathematics, science, and STEM teaching and learning. *School Science and Mathematics*, *124*(1), 1–5. https://doi.org/10.1111/ssm.12635

Redmond-Sanogo, A., Burton, M., Ivy, J., & Maiorca, C. (2026). Generative AI in mathematics, science, and STEM education: Research, applications, and emerging themes. *School Science and Mathematics*, *126*(1), 3–8. https://doi.org/10.1111/ssm.70002

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Sage.

Schön, D. A. (1979). *The reflective practitioner*. Harper & Collins.

Seo, K., Yoo, M., Dodson, S., & Jin, S. H. (2025). Augmented teachers: K–12 teachers' needs for artificial intelligence's complementary role in personalized learning. *Journal of Research on Technology in Education*, *57*(4), 876–893. https://doi.org/10.1080/15391523.2024.2330525

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

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

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

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), Article 39. https://doi.org/10.1186/s41239-019-0171-0

STEM Career Changers and Critical Reflection: Identity, Practice, and Equity

Jennifer R. Simons
George Mason University
jsimons2@gmu.edu

Abstract

STEM career-changers increasingly fill hard-to-staff classrooms, yet little is known about how they engage in critical reflection once teaching. This exploratory qualitative study examined three in-service STEM career-changers through semi-structured interviews, analyzed using van Manen's (1997) levels of reflection. All three described reflecting daily, though type and depth varied: some focused on technical and practical concerns, others extended to structural and ethical issues. Participants valued reflection and wanted protected time and structured support. Although limited to a small convenience sample, the findings suggest that career-changers may benefit from reflection embedded as ongoing practice and from scaffolding toward critical reflection.

Keywords: STEM career changers, critical reflection, teacher identity, alternative certification, qualitative case study

Introduction

As education faces a persistent shortage of teachers, schools increasingly recruit STEM professionals as career-changers to fill hard-to-staff classrooms, especially in schools serving low-income students and students of color (Carver-Thomas & Darling-Hammond, 2017). These career-changers can teach students to value reasoning, analyze models, and make real-world connections, provided they maintain their STEM identity while building a teacher identity (Grier & Johnston, 2012). However, little research has studied their ability to critically reflect once in the classroom.

Career-change certification differs from traditional pathways in ways that can hinder the development of career-changers' teacher identities and deemphasize critical reflection, social context, and equity (Grier & Johnston, 2012). While adept in their field, career-changers often struggle to make content accessible to diverse learners (Berry, 2000), which is particularly consequential given that alternatively certified teachers disproportionately staff schools serving Black and Latine students (30 percent of math and science teachers) compared to predominantly White schools (12 percent) (Carver-Thomas & Darling-Hammond, 2017). Despite these difficulties, career-

Anderson-Pence, K., & Shelton, R. N. (Eds.). (2026). *Proceedings of the 125th annual convention of the School Science and Mathematics Association* (Vol. 13). Chicago, IL: SSMA.

changers are motivated to make science and mathematics relevant to students' lives outside the classroom (Grier & Johnston, 2009), often adopting student-centered methodologies to do so (Chambers, 2002), a prerequisite for culturally responsive pedagogy (Howard, 2003).

Objectives of the Study

Because schools serving low-income students and students of color are more likely to have STEM vacancies filled by career-changers (Carver-Thomas & Darling-Hammond, 2017), it is vital to understand how these teachers engage in critical reflection. Existing research on critical reflection has focused largely on pre-service teachers, leaving a gap in our understanding of how in-service STEM career-changers reflect and why. This exploratory study addresses that gap by examining the reflective practices of three in-service STEM career-changers. It contributes (a) a description of the types of reflection these teachers engage in, from technical concerns to structural and ethical issues, (b) the purposes they attach to reflection, and (c) their accounts of what supports or hinders it, including collaborative reflection that one participant experienced as competitive. The study does not aim to generalize to all career-changers. It aims to offer detailed cases that can inform future, larger studies and the design of support for these teachers.

Related Literature

Alternative certification programs have been proposed as a lower-cost, expedited pathway to placing qualified teachers in STEM classrooms (Grier & Johnston, 2012), yet little research exists on the subset known as career-changers. Recruited for their field knowledge and experience, these teachers follow a distinct pathway that can either reinforce the status quo or enact change, aligning with the growing emphasis on equity-centered, social-justice-based STEM education (Carver-Thomas & Darling-Hammond, 2017). While they exhibit some favorable behaviors toward critical reflection, more is needed to understand their preparation, socialization, and construction of teacher identity (Howard, 2003; Liu, 2015; Limbere et al., 2022).

Teacher staffing shortages nationwide have fueled the growth of alternative certification programs (David, 2020; Garcia & Weiss, 2019), though career change pathways remain loosely and inconsistently defined across states (Devier, 2019). These teachers are disproportionately assigned to high-poverty schools (Garcia & Weiss, 2019) and to schools serving students of color (Carver-Thomas & Darling-Hammond, 2017). This is concerning, given that career change programs often offer weaker pedagogy coursework and limited field experience (Whitford et al., 2018), leaving teachers adept in their field yet underprepared to reach diverse learners (Berry, 2000). STEM

positions are especially hard to fill (Carver-Thomas & Darling-Hammond, 2017), making the expedited, lower-cost certification pathways these programs offer particularly attractive to STEM professionals (David, 2020; Matsko et al., 2022). Despite these barriers, career-changers are motivated to make science and math more relevant to students' lives outside the classroom (Grier & Johnston, 2009), often adopting student-centered methods to do so (Chambers, 2002), an orientation that aligns with critical reflection (Limbere et al., 2022).

Reflection has been a core component of teaching since the 1930s (Dewey, 1933), and has been further developed by Schön's (1983) concept of the reflective practitioner. Reflection helps educators to adapt their teaching to particular students within particular contexts (Hammerness et al., 2005), which is incredibly important for teachers whose students may be cultural, ethnic, or racial minorities (Ladson-Billings, 2009). Van Manen (1997) defines three levels of reflection: technical (optimizing teaching-learning methods), practical (considering multiple perspectives without evaluation), and critical (considering the social, moral, and political dimensions of schooling). Critical reflection requires teachers not only to turn their gaze inward but also upward to the overarching structures of the educational system. This is the first step towards addressing unstated assumptions that can negatively impact lower-income students and students of color. Understanding the factors that impact teachers' engagement in critical reflection is key to a more culturally responsive pedagogy.

Research identifies several positive indicators for critical reflection: open-mindedness and diverse life experiences (Liu, 2015), questioning of traditional teacher-centered practices (Howard, 2003), strong teacher identity (Philp-Clark & Grieshaber, 2024), and collaborative reflection (Limbere et al., 2022). Career-changers often bring many of these traits as older entrants. They are also frequently motivated to connect content to students' lives (Grier & Johnston, 2009), typically through student-centered practice (Chambers, 2002). Yet research on whether this predisposition translates into practice is mixed: career-changers reported feeling "annoyed" by required reflection (Grier & Johnston, 2009), and Arend et al. (2021) found that preservice STEM education students and instructors tend to define critical reflection narrowly, often remaining at a technical level and treating reflection as secondary to "straightforward" STEM content. Career-changers may face additional constraints: Grier and Johnston (2012) noted that they often have limited time to collaborate and socialize with their colleagues, which can restrict opportunities for collaborative reflection.

This study draws on van Manen's (1997) three levels of reflection, which I use to describe both the type of reflection participants reported and the purposes they attached to it. This framework informs the study in two ways. First, it provides the a priori codes used in the second round of analysis, which allowed me to describe the type of reflection in each participant's account. Second, it links the research question, which asks about types of reflection and their purposes, to the study's concern with equity: critical reflection is the level at which equity issues most directly come into view. Throughout the paper I use "critical reflection" to refer to reflection at van Manen's critical level.

Methodology

This study asks the following question: What types of reflection do STEM career-changers engage in, and for what purposes? I focus on three career-changers' accounts of reflection in the classroom. Because reflection is introspective and individualistic (Rowley et al., 2023), participants' own words are essential for understanding how they reflect. For this study, STEM career-change teachers were defined as individuals with a bachelor's degree in a STEM field who had worked in a STEM field for at least three years before teaching.

Participants were a convenience sample recruited through my professional networks. Three teachers indicated interest in participating in a study about career-changer experiences and were interviewed using a semi-structured format to elaborate on their classroom reflection experiences, with informal follow-ups conducted about one month later. Interview transcripts were first read individually and in vivo coded to maintain participants' original ideas (Saldaña, 2015), with member checking to confirm accuracy. The second round of a priori coding utilized van Manen's (1997) descriptions of reflection to ascertain the type and level of reflection.

Results and Discussion

True to the personal nature of reflection, each teacher had a distinct approach, using reflection to shape their instructional decisions and relationships with students, and drawing on different types of reflection for different purposes. Technical reflection helped them adjust instruction and materials; practical reflection helped them consider multiple perspectives; and critical reflection helped them interrogate broader issues of equity, identity, and student experience. All names used are pseudonyms.

Purposeful Practitioner

Clara was an environmental engineer with four years of experience teaching grades 2–7. She primarily discussed engaging in a technical level of reflection, noting, "I do kind of go through scenarios like had during the day interactions and think about, would I have handled that better?" As a result, she primarily discussed reflecting on how students understood lessons, different grouping combinations, and what resources to pull to help them better understand. Clara indicated that she reflected individually throughout the day and that her school had also incorporated mandatory collaborative reflection during content-area planning, with a focus on data-driven pedagogical decisions. While she appreciated having dedicated time to reflect and plan, she often felt intimidated during collaborative reflection with her coworkers, stating that it often led to animosity and "like it becomes a competition between the team."

Critical Carer

Gajara was a lab-based microbiologist before teaching life science (seventh grade). She had taught at one school for five years, with a virtual and hybrid first year due to COVID. She expressed interest in teaching to help students remain curious while also meeting high standards. She felt unencumbered by the curriculum and often sought community partnerships for project-based learning, which she found especially helpful for her multilingual learners. She noticed her biggest struggle was keeping content meaningful, reaching students, and still maintaining her identity. She largely engaged in practical and technical reflection: "Was the content meaningful? Was I too harsh on a student? Were my directions not clear on the lab?" Yet she also engaged with structural and ethical issues, participating in committees to address inequities she saw at her school, such as the lack of meaningful science experiences for students with disabilities, and creating community partnerships.

Activist Educator

Rachel had the most teaching experience, having completed Teach For America and taught secondary math and science for over 20 years. She was originally a cell biologist who worked in public health but wanted to make a difference. She often used her reflective experiences to build critical consciousness in her classroom. She noted that in the beginning of her career, reflection was perfunctory, but she later evolved to thinking, "...Am I creating an environment where students are comfortable, asking for help? Where they're comfortable asking each other for help?" She encouraged students to reflect on both their performance in class and on the context surrounding

math and science. She noted the importance of getting students to think about not just how the content relates to their lives, but also how to identify and rectify injustices: "Okay, you're right. So what can we do about it right now? Nothing. But what can we do about things like that moving forward?"

Conclusions and Implications

Career-changers are increasingly filling STEM teaching positions (Garcia & Weiss, 2019) in some of the nation's most diverse classrooms (Ladson-Billings, 2009), making it important to understand how their reflective practices shape their identities, relationships, and instructional decisions. This study adds to a literature that has centered on pre-service teachers by describing how three in-service STEM career-changers reflect. The three participants described reflecting daily and using reflection to navigate instructional challenges, build relationships with students, and make sense of their evolving teacher identities. Their reflections varied in type and purpose: technical reflection helped them adjust instruction, practical reflection helped them consider multiple perspectives, and critical reflection allowed them to interrogate broader issues of equity and student experience. Given the size and nature of the sample, these findings do not show that career-changers as a group reflect in these ways. They do suggest that at least some career-changers bring a meaningful foundation for reflective practice, even if they do not always label it as such, and they raise questions worth testing in larger samples.

At the same time, participants expressed a desire for more intentional time and space to reflect. Reflection was described as "deeply personal," and teachers noted that the fast pace of the school day often limited their ability to engage in sustained, critical thinking that supports culturally responsive pedagogy. This aligns with research indicating that critical reflection "needs to be honed and scaffolded over time" (Yoon & Kim, 2010) and that without structured opportunities, reflection risks remaining at a technical level (Arend et al., 2021). The participants' accounts point to a possible role for capacity building. While all three teachers demonstrated the ability to reflect, they also noted that guided reflections sometimes felt forced, echoing Grier & Johnston's (2009) observation that career-changers may be "annoyed" by required reflection. Yet participants acknowledged that structured prompts can provide a useful starting point, especially for those still developing a teacher identity or learning to navigate diverse classrooms. This is consistent with Gorski & Dalton's (2020) argument that critical reflection often requires scaffolding, dialogue, and intentional practice.

For teacher educators and professional development designers, these cases suggest several possibilities to explore. Reflection may be most useful when it is embedded as a daily, supported practice rather than an isolated assignment. Programs might differentiate types of reflection and offer targeted scaffolds to help career-changers move from technical to critical reflection, particularly around equity and student identity. Collaborative reflection may need careful facilitation, since one participant experienced poorly structured collaboration as intimidating and competitive. Finally, participants described time pressure as a barrier to sustained reflection, suggesting the need to protect time for it. Future research should use larger and more varied samples, including career-changers who do not report regular reflection, and should combine interviews with observations, reflective artifacts, and longitudinal data to examine how reflection develops over time.

References

Arend, B., Archer-Kuhn, B., Hiramatsu, K., Ostrowdun, C., Seeley, J., & Jones, A. (2021). Minding the gap: Comparing student and instructor experiences with critical reflection. *Teaching & Learning Inquiry, 9*(1), 317–332. https://doi.org/10.20343/teachlearninqu.9.1.21

Berry, B. (2000). Quality alternatives in teacher preparation: Dodging the "silver bullet" and doing what is right for students. National Association of State Boards of Education. *The State Education Standard, 1*(4), 21–25.

Carver-Thomas, D., & Darling-Hammond, L. (2017). *Teacher turnover: Why it matters and what we can do about it*. Learning Policy Institute. https://doi.org/10.54300/454.278

Chambers, D. (2002). The real world and the classroom: Second career teachers. *The Clearing House, 75*(4), 212–217. https://doi.org/10.1080/00098650209604935

David, B. G. (2020). *What is the competitive advantage? Exploring the intersection of charter schools and alternative teacher certification upon STEM outcomes* (Publication No. 28763844) [Doctoral dissertation, University of Texas at Austin]. ProQuest Dissertations & Theses Global.

Devier, B. H. (2019). Teacher shortage and alternative licensure solutions for technical educators. *The Journal of Technology Studies, 45*(2), 48–59. https://doi.org/10.21061/jots.v45i2.a.1

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process.* Henry Regnery.

García, E., & Weiss, E. (2019). *The teacher shortage is real, large and growing, and worse than we thought.* Economic Policy Institute.

Gorski, P. C., & Dalton, K. (2020). Striving for critical reflection in multicultural and social justice teacher education: Introducing a typology of reflection approaches. *Journal of Teacher Education, 71*(3), 357–368. https://doi.org/10.1177/0022487119883545

Grier, J. M., & Johnston, C. C. (2009). An inquiry into the development of teacher identities in STEM career changers. *Journal of Science Teacher Education, 20*(1), 57–75. https://doi.org/10.1007/s10972-008-9119-2

Grier, J. M., & Johnston, C. C. (2012). STEM professionals entering teaching: Navigating multiple identities. *Journal of Science Teacher Education, 23*(1), 19–44. https://doi.org/10.1007/s10972-011-9260-1

Hammerness, K. M., Darling-Hammond, L., Bransford, J., et al. (2005). How teachers learn and develop. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 358–389). Jossey-Bass.

Howard, T. C. (2003). Culturally relevant pedagogy: Ingredients for critical teacher reflection. *Theory Into Practice, 42*(3), 195–202. https://doi.org/10.1207/s15430421tip4203_5

Ladson-Billings, G. (2009). Race still matters: Critical race theory in education. In M. W. Apple, A. Au, & L. A. Gandin (Eds.), *The Routledge international handbook of critical education* (pp. 110–122). Routledge.

Limbere, A. M., Munakata, M., Klein, E. J., & Taylor, M. (2022). Exploring the tensions science teachers navigate as they enact their visions for science teaching: What their feedback can tell us. *International Journal of Science Education, 44*(12), 1897–1915. https://doi.org/10.1080/09500693.2022.2105413

Liu, K. (2015). Critical reflection as a framework for transformative learning in teacher education. *Educational Review, 67*(2), 135–157. https://doi.org/10.1080/00131911.2013.839546

Matsko, K. K., Ronfeldt, M., & Greene Nolan, H. (2022). How different are they? Comparing teacher preparation offered by traditional, alternative, and residency pathways. *Journal of Teacher Education, 73*(3), 225–239. https://doi.org/10.1177/00224871211015976

Philp-Clark, C., & Grieshaber, S. (2024). Teacher critical reflection: What can be learned from quality research? *The Australian Educational Researcher, 51*, 697–717. https://doi.org/10.1007/s13384-023-00619-7

Rowley, J., Giles, P., Hammond, A., Hussein, A., Oakey, M., O'Hara, S., Williams, J., & Wood, K. (2023). Applying a framework for critical reflection in educational psychology practice: Views of trainee educational psychologists. *Educational Psychology in Practice, 39*(1), 19–37. https://doi.org/10.1080/02667363.2022.2150603

Schön, D. A. (1983). *The reflective practitioner.* Basic Books.

van Manen, M. (1997). On the epistemology of reflective practice. *Teachers and Teaching: Theory and Practice, 1*(1), 33–50. https://doi.org/10.1080/1354060950010104

Whitford, D. K., Zhang, D., & Katsiyannis, A. (2018). Traditional vs. alternative teacher preparation programs: A meta-analysis. *Journal of Child and Family Studies, 27*(3), 671–685. https://doi.org/10.1007/s10826-017-0932-0

Yoon, H. G., & Kim, M. (2010). Collaborative reflection through dilemma cases of science practical work during practicum. *International Journal of Science Education, 32*(3), 283–301. https://doi.org/10.1080/09500690802516538

INVESTIGATING FACTORS IMPACTING PRESERVICE ELEMENTARY TEACHERS' iSTEM IDENTITY DEVELOPMENT

Azadeh Hassani
University of Nebraska-Lincoln
ahassani4@huskers.unl.edu

Ursula Nguyen
Texas State University
u.nguyen@txstate.edu

Fatemeh Ashrafabadi
University of Nebraska-Lincoln
fashrafabadi2@huskers.unl.edu

Deepika Menon
University of Nebraska-Lincoln
dmenon2@unl.edu

Amanda Henson
University of Nebraska-Lincoln
amanda.thomas@unl.edu

Abstract

This exploratory study investigated factors associated with preservice elementary teachers' integrated STEM (iSTEM) teacher identity development within a STEM methods semester. Written reflections from 40 preservice teachers were analyzed using an abductive coding approach informed by iSTEM teacher identity and teacher identity development frameworks. Findings indicated that participants predominantly described themselves as becoming iSTEM teachers. Their reflections identified two internal factors, self-efficacy and sense of responsibility, and two external factors, field experiences and peer/mentor support, as relevant to their developing iSTEM teacher identities. Findings highlight the role of integrated coursework, authentic teaching experiences, and supportive communities in elementary teacher preparation.

Keywords: integrated STEM; preservice elementary teachers; teacher identity; STEM teacher education; teacher preparation

Introduction

Despite integrated STEM (iSTEM) gaining traction in PK–12 education, recent research suggests that elementary teachers are not well-prepared to plan and enact iSTEM learning experiences (Hammack & Ivey, 2019). To address this challenge, professional development workshops have been implemented to better support in-service elementary teachers (Hammack et al., 2020; Watkins et al., 2021). Yet, these professional development experiences are often one-day occurrences or standalone opportunities for individual teachers. Moreover, this raises the question about the iSTEM learning and teaching experiences and supports available to elementary preservice teachers (PSTs) prior to entering the teaching profession. We conceptualize integrated STEM (iSTEM) as an active and collaborative approach to learning that connects disciplinary content with

authentic, real-world contexts (Kelley & Knowles, 2016; Kennedy & Odell, 2014) and positions scientific inquiry as a foundation for integrating science, engineering, technology, and mathematics (Menon et al., 2023; Moore & Smith, 2014).

Objectives of the Study

This exploratory study examines the experiences and viewpoints of a sample of future elementary teachers who were enrolled in a STEM methods semester that focused on integrated STEM teaching and learning. We describe the innovative design of the STEM methods semester, which couples several STEM content and methods courses with a field-based practicum experience at a local school district located in a rural region. The objective of the study is to examine elementary PSTs' iSTEM teacher identity development as articulated in their written reflections. Further, we identify the internal and external factors related to their iSTEM teacher identity development.

Theoretical Framework and Related Literature

This study is informed by Galanti and Holincheck's (2022) proposed model for conceptualizing iSTEM teacher identity. In their model, iSTEM teacher identity is composed of multiple dimensions, including motivation, self-image, self-efficacy, task perception, and teaching interest. This model is useful as it builds upon prior work on elementary professional teacher identity (Hanna et al., 2020) and STEM learner identity (Hazari et al., 2010). In doing so, it acknowledges past and present STEM learning experiences. For instance, the teaching interest dimension captures the extent to which teachers think about and understand iSTEM teaching while the motivation component describes teachers' reasons for teaching iSTEM. For this study, we focus on the self-efficacy dimension.

At the same time, Galanti and Holincheck's (2022) conceptual model relies solely on teachers' internal beliefs and perceptions. To better understand other possible factors, including external factors, that may also contribute to PSTs' iSTEM teacher identity development, we also frame this study using the Can't-To-Becoming-to-Being progression described by Quaisley et al. (2023). Notably, their model extends prior research on STEM teacher leadership identity by delineating how STEM teachers' perceptions about what they "can" and "should" do are related to their self-efficacy and sense of responsibility, respectively, particularly as their identity develops from "can't" to "becoming" a STEM teacher. Although Quaisley et al. (2023) developed this progression for STEM teacher leadership identity among practicing teachers, we apply its underlying identity-development processes, particularly self-efficacy and sense of responsibility, to PSTs as they begin to

construct their identities as future iSTEM teachers. We also draw on communities of practice, in which learning and identity develop through participation in shared social practices (Lave & Wenger, 1991; Wenger, 1998). Through their coursework and field experiences, PSTs engaged with peers, instructors, and practicing teachers in shared practices of planning and enacting iSTEM instruction; thus, we conceptualize these settings as communities of practice supporting their emerging iSTEM teacher identities. Therefore, building on these insights, we investigate the following research question in this exploratory study: *What internal and external factors contribute to elementary PSTs' iSTEM teacher identity development?*

Methodology

The context for this study involves a STEM methods semester of an elementary teacher education program at a large, land-grant research university located in the Midwestern region of the United States. This STEM methods semester marks the first semester for elementary PSTs accepted into the professional program. Therefore, elementary PSTs are typically sophomores and juniors in college. The courses that comprise the STEM methods semester include a science methods course, a math methods course, an innovative learning technologies (ILT) course, and a course connected to a field-based practicum experience in a neighboring rural school district. Moreover, the STEM methods semester reflects a joint effort established by a group of STEM methods instructors. The STEM methods semester provides elementary PSTs with authentic iSTEM learning and teaching experiences through its cross-cutting themes and shared assignments across the three methods courses (mathematics, science, and ILT). Some of these cross-cutting themes include: Coding & Robotics, STEAM & Creativity, and Engineering Design & Sustainability. To engage elementary PSTs in iSTEM, the shared assignments are a group iSTEM lesson project and an individual written reflection. This latter assignment is the focus of the current study.

In the fall semester of 2023, 78 elementary PSTs were enrolled across three sections of the STEM methods courses. The final STEM growth project included prompts that provided an opportunity for PSTs to reflect on their experiences that supported their growth as learners and future teachers teaching iSTEM. This study utilized elementary PSTs' written reflections, developed, discussed and revised by STEM methods instructors, that asked students to respond to five question prompts to reflect on their developing identities as learners and future teachers (see Table 1). Although these reflections were also part of the shared assignment across courses, the research team analyzed them a semester after the completion of the course. The study was conducted with

institutional IRB approval. Three members of the research team also served as instructors in the STEM methods semester. We prepared a de-identified file (PSTs were identified as participant #s with names removed) before the data were shared among the research team. We also acknowledge that the prompts were framed around growth, which was an intended outcome we expected. However, we did pay close attention to the challenges that participants reported in teaching iSTEM.

Table 1

Integrated STEM Growth Reflection Prompts and Their Use in the Analysis

Reflection Prompt
1. What are two benefits and two challenges of teaching integrated STEM to elementary students?
2. How have you grown with respect to integrating across STEM subjects?
3. How will you continue to grow in your ability to integrate across STEM subjects in the future?
4. From your experience in the STEM Methods Block, what five aspects do you think would be necessary for teaching elementary STEM effectively? Describe why each aspect is important and essential for STEM in particular.
5. What contributed to your knowledge, understanding, and confidence about integrated STEM? Include artifacts from at least two different STEM Block classes to support your response.

As this is an exploratory study, the research team randomly selected 40 individual reflections (about half) for initial analysis, ensuring that a similar number of reflections were included for each section. Future work will consider all 78 reflections. As our focus is on iSTEM teacher identity development, the research team examined the written reflections for mentions of elementary PSTs' confidence, motivation, and interest in relation to iSTEM teaching and learning.

Specifically, we used a hybrid or abductive approach for coding. Specifically, a priori codes were derived from existing literature on sources of teacher identity development (Galanti & Holincheck, 2022; Quaisley et al., 2023). Open coding was utilized to characterize how those sources supported iSTEM teacher identity among elementary PSTs (Saldaña, 2015). A team of three coders conducting the coding, with 100% of the data double-coded by at least one faculty member and one graduate research assistant. Disagreements were resolved through research team meetings. To further ensure trustworthiness, we maintained a spreadsheet with codes and excerpts to discuss codes during peer debriefs.

Results and Discussion

For this study, we operationalized weak, moderate, and strong iSTEM teacher identities that parallel Quaisley et al.'s (2023) framework for can't/become/being progression. Specifically, we considered weak iSTEM teacher identities as those where elementary PSTs do not perceive themselves as an iSTEM teacher. These reflections included hesitation or self-doubt of their ability to teach iSTEM. For example, "*I wouldn't say that my respect per say has changed when it comes to integration around STEM subjects… I do think it is a hard process to do well, STEM itself like I've said before is not an easy thing to just go out and teach,*" (PST, 25). In contrast, moderate iSTEM teacher identities were characterized as having potential to grow into an iSTEM teacher. Oftentimes, reflections included self-perceptions of areas of growth and potential solutions or sources for addressing those areas of weaknesses. Strong iSTEM teacher identities described identities of elementary PSTs where confidence for teaching iSTEM was present as well as a positive attitude toward iSTEM teaching and learning.

Overall, the written reflections suggested that elementary PSTs expressed moderate iSTEM teacher identities or described themselves as becoming STEM teachers. In this study, 16 participants began with a weak iSTEM teacher identity, with 14 shifting to moderate iSTEM teacher identities and 2 remaining weak iSTEM teacher identities. While none began with strong iSTEM teacher identities, 2 of the 24 that began with moderate iSTEM teacher identities shifted to strong identities and the rest (22 participants) still had moderate iSTEM teacher identities. Notably, the 14 elementary PSTs described as a shift from weak toward more moderate in their STEM teacher identities had common threads, such as those shown below, where elementary PSTs reported not only a positive relationship with teaching STEM subjects but also engaging with STEM content during the semester.

> At the beginning of the semester, I was really not a fan of STEM subjects. I have had some bad experiences with teachers and have grown to dislike them. After this semester, I have regained a positive relationship with these subjects. I am excited to be able to teach integrated STEM to my future students and hopefully give them better experiences than I had. (PST 38)

> I also shared at the beginning of the semester that I have never enjoyed science, particularly in elementary school, so I was nervous about teaching science in a way that is both engaging and exciting for me and my students. Now, at the end of the semester, I have a lot of tools

> that I can take into my future classroom to help make teaching integrated STEM seem like more of a realistic possibility. (PST 20)

Similar to these, other elementary PSTs recounted STEM methods semester experiences that contributed to their enthusiasm for teaching iSTEM in the future and becoming a STEM teacher (or having a moderate STEM teacher identity). Therefore, we sought to differentiate between internal and external factors that were related to their positive STEM teacher identity development. We utilize Quaisley et al.'s (2023) model around teacher identity development to focus on the internal factors of self-efficacy and sense of responsibility noted in PSTs' reflections. We also expand on Quaisley et al.'s (2023) work to include multiple external communities of practice including PSTs' field experience and support from peers/mentors.

Internal Factors: Self-Efficacy and Sense of Responsibility

Self-efficacy. Elementary PSTs' confidence grew as they learned the STEM content necessary to teach in the elementary school grades. Indeed, having more STEM content expertise led to more comfortability with teaching iSTEM. The excerpts below exemplify the increase in their iSTEM teaching self-efficacy and their identification as an iSTEM teacher.

> [T]his semester I have grown in my understanding of what STEM integration is as well as my comfort and ability in being able to plan and execute a STEM integrated lesson. (PST 33)

> Toward the end of the semester, I started to actually believe in myself when it comes to teaching all of these subjects and putting them all together in one lesson. (PST 35)

Sense of responsibility. Notably, elementary PSTs expressed their becoming an iSTEM teacher as inherently tied to the responsibility of teaching elementary students. They hoped to provide meaningful iSTEM learning experiences to their future students. To them, their duty of being an iSTEM teacher was for the sake of ensuring that student learning would not be compromised by the omission of any of the STEM areas due to lack of time or planning.

> During my practicum, I firsthand saw the lack of science in elementary classes and how limited time they have to integrate science that they are required to teach. This showed me the importance of integrating STEM into classrooms because otherwise students might not see science as often as I believe they should. (PST 36)

> I know that it takes time and effort to learn as a teacher how these tools can be used in a classroom, but I think it's important that teachers know how to use them confidently. This is important because it can help lessen their workload, and it can help students gain a deeper understanding of the information that they are learning. (PST 40)

Yet, we note that this sense of responsibility (an internal factor) grew out of the elementary PSTs' first-hand experience in their practicum classroom (an external factor). Therefore, it is possible that external and internal factors can influence one another and contribute to PST's iSTEM teacher development.

External factors: Field experience and peer/mentor support

Field experience. Many elementary PSTs attributed their development as an iSTEM teacher to their direct teaching experiences in their field experience (practicum experience). They saw connections between what they were learning as students in the STEM methods semester and as elementary PSTs in their field placements. As recounted below, elementary PSTs had opportunities to plan and teach an iSTEM lesson to elementary students, and their reflections indicated that they perceived these experiences as contributing to their development as iSTEM teachers.

> My practicum experiences were extremely helpful as I was able to experience everything I was learning in the classroom in real life. (PST 32)

> I was also given the opportunity to teach an integrated STEM project to my practicum students. Once our group decided on the project we would be focusing on, finding the aspects to make it integrated was surprisingly easy. (PST 38)

Peer/mentor support. Generally, elementary PSTs whose iSTEM teacher identities shifted from weak to moderate expressed support from peers and mentors (e.g., cooperating teachers, professors, field supervisors). Their reflections associated feelings of empowerment in becoming iSTEM teachers with support from current and future colleagues who embraced teaching iSTEM. As the example below shows, elementary PSTs described being motivated to seek and encounter more mentors/peers teaching iSTEM in their future semesters.

> During my next three semesters, I really want to pay attention to how the cooperating teachers I work with go about STEM integration and teaching those lessons to see how effective different methods are in a real classroom environment. (PST 11)

> [O]ne extremely popular way of teaching Integrated STEM I've seen from my CT in practicum is incorporating literature into science teaching. (PST 5)

Implications

With the increasing adoption of iSTEM learning and teaching across PK–12 education, new teachers need adequate preparation to successfully engage learners. Therefore, the STEM methods semester, with its design to provide meaningful and relevant iSTEM learning and teaching experiences, serves as a purposeful context to explore future elementary teachers' iSTEM teacher identities. In doing so, we found one group of 14 PSTs whose identity development was transformative, evolving from a weak to a more moderate identification as an iSTEM teacher. For these PSTs, their reflections identified both internal (self-efficacy and sense of responsibility) and external (field placement and peer/mentor support) factors as relevant to their STEM teacher identity development. Because we collected and analyzed written reflections at the end of the semester, it is possible that elementary PSTs' iSTEM teacher identities described their experiences that inherently include change or growth. Future work should ask these reflection prompts at the beginning and end of the semester to accurately measure identity shifts and possibly identify additional categories beyond those discussed in this study.

An important implication of our study is that, while many elementary teachers have shared not feeling comfortable or confident in their STEM teaching or having poor past STEM learning experiences (Hammack & Ivey, 2019), elementary teacher education programs can be pivotal for improving future elementary teachers' iSTEM teacher identities (Galanti & Holincheck, 2022; Holincheck & Galanti, 2023). Our findings suggest that PSTs perceived experiences within the STEM methods semester as contributing to their iSTEM teacher identities, including their self-efficacy and enjoyment related to STEM teaching. As we move forward with this work, we plan to analyze PSTs' iSTEM lesson plans, as these were also collected during the STEM methods semester. In particular, we are interested in assessing whether specific factors (internal or external) for iSTEM teacher identity development are related to stronger iSTEM lesson plans.

References

Galanti, T. M., & Holincheck, N. (2022). Beyond content and curriculum in elementary classrooms: Conceptualizing the cultivation of integrated STEM teacher identity. *International Journal of STEM Education*, *9*(1), Article 43.

Hammack, R., Gannon, P., Foreman, C., & Meyer, E. (2020). Impacts of professional development focused on teaching engineering applications of mathematics and science. *School Science and Mathematics, 120*(7), 413–424.

Hammack, R., & Ivey, T. (2019). Elementary teachers' perceptions of K–5 engineering education and perceived barriers to implementation. *Journal of Engineering Education*, *108*(4), 503–522.

Hanna, F., Oostdam, R., Severiens, S. E., & Zijlstra, B. J. (2020). Assessing the professional identity of primary student teachers: Design and validation of the teacher identity measurement scale. *Studies in Educational Evaluation*, *64*, Article 100822.

Hazari, Z., Sonnert, G., Sadler, P. M., & Shanahan, M. C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching*, *47*(8), 978–1003.

Holincheck, N. M., & Galanti, T. M. (2023). Applying a model of integrated STEM teacher identity to understand change in elementary teachers' STEM self-efficacy and career awareness. *School Science and Mathematics, 123*(6), 234–248.

Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education, 3*(1), Article 11.

Kennedy, T. J., & Odell, M. R. (2014). Engaging students in STEM education. *Science Education International*, *25*(3), 246–258.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Menon, D., Bauer, A. S., Johnson, K. L., Hasseler, E. F., Thomas, A., Martinez, R., & Trainin, G. (2023). Greater than the sum of its parts: Centering science within elementary STEM education. In *Reforming science teacher education programs in the STEM era: International and comparative perspectives* (pp. 233–250). Springer International Publishing.

Moore, T. J., & Smith, K. A. (2014). Advancing the state of the art of STEM integration. *Journal of STEM Education: Innovations and Research*, *15*(1), 5–10.

Quaisley, K., Smith, W. M., Criswell, B., Funk, R., & Hutchinson, A. (2023). From becoming to being: How STEM teachers develop leadership identities. *International Journal of Leadership in Education*, 1–28.

Saldaña, J. (2015). *The coding manual for qualitative researchers (3rd ed.)*. Sage.

Watkins, J., Portsmore, M., & Swanson, R. D. (2021). Shifts in elementary teachers' pedagogical reasoning: Studying teacher learning in an online graduate program in engineering education. *Journal of Engineering Education, 110*(1), 252–271.

Wenger, É. (1999). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.

NOT TOO SMALL FOR EMERGING SCIENCE: NAVIGATING SCALE, COMPLEXITY, AND QUANTUM CONCEPTS IN ELEMENTARY CLASSROOMS

Nancy Holincheck
George Mason University
nholinch@gmu.edu

Jennifer Simons
George Mason University
jsimons2@gmu.edu

Tiffany Butler
George Mason University
tbutle5@gmu.edu

Jessica Rosenberg
George Mason University
jrosenb4@gmu.edu

Benjamin W. Dreyfus
George Mason University
bdreyfu2@gmu.edu

Stephanie Dodman
George Mason University
sdodman@gmu.edu

Christina Cameron
George Mason University
ccamer3@gmu.edu

Laura Akesson
George Mason University
lakesson@gmu.edu

Abstract

While quantum information science is increasingly prioritized in national workforce initiatives, the discipline remains virtually untouched in elementary curricula. This study investigates the feasibility of integrating quantum science into elementary education. Analyzing data from 13 elementary educators across two NSF-funded projects, we examined classroom instructional strategies, student sensemaking, and systemic implementation barriers. Grade 5 students demonstrated enthusiasm for quantum concepts and careers when teachers used hands-on, analogy-based pedagogy, but students persistently struggled with scale and tended to echo surface-level jargon. Teachers faced low confidence in content knowledge, rigid pacing schedules, and lack of alignment between content and standards. Overcoming these barriers requires structured curricular scaffolds, targeted professional development, and transdisciplinary integration, showing that elementary classrooms can support early quantum learning.

Keywords: quantum education, elementary STEM, teacher professional learning, career awareness

Introduction

The growth of quantum information science and engineering (QISE) has created an urgent need to build a prepared, diverse workforce (Kelly et al., 2024). Although quantum mechanics is traditionally reserved for higher education due to its mathematical rigor, foundational STEM identities and career aspirations take shape before middle school (Archer et al., 2010). Introducing

quantum concepts in elementary grades can demystify advanced science, cultivate curiosity, and broaden participation in future STEM pathways (Holincheck et al., 2024).

However, quantum topics remain largely absent from K–12 standards (Kelly et al., 2024). As demonstrated in upper grades, successful quantum integration requires supporting educators through targeted professional learning (Matsler & Lopez, 2023). This study investigates the feasibility of integrating quantum concepts into elementary classrooms. Specifically, we explore instructional approaches that make abstract quantum ideas accessible to young learners, examine student sensemaking, and evaluate the systemic barriers educators face during implementation.

Related Literature and Theoretical Framework

Teaching quantum science to young learners requires shifting from mathematical formalisms to visual, tactile, and analogy-based pedagogies (Hasanovic, 2023). Although elementary students demonstrate cognitive readiness for abstract reasoning through interactive models and gamification (Holincheck et al., 2024), implementation is constrained by dual instructional and institutional barriers. Research frequently highlights the low self-efficacy of secondary educators when teaching counterintuitive subatomic concepts (Kelly et al., 2024), a challenge that is heavily compounded for elementary generalists.

This inquiry draws on Conceptual Change Theory (Posner et al., 1982; Vosniadou, 1994), research on cognition of size and scale (Tretter et al., 2006), and early STEM Identity Development (Archer et al., 2010). Conceptual change posits that learners construct intuitive theories from everyday macroscopic experiences (Vosniadou, 1994). When students encounter unobservable phenomena, they often map new vocabulary onto existing macroscopic mental models, producing synthetic misconceptions or repeating surface jargon rather than restructuring their thinking. In the physical sciences, this process is closely tied to scale cognition (Stieff et al., 2013). Upper-elementary students generally possess benchmark intuitions for visible objects, but they frequently collapse micro, atomic, and subatomic worlds into an undifferentiated category of "small" (Tretter et al., 2006). For fifth-grade learners, developmentally appropriate quantum education does not require advanced mathematics. Instead, it entails differentiating scale between macroscopic and subatomic phenomena, contrasting classical and emerging computing logic, and building early career awareness (Holincheck et al., 2024). We also integrate STEM identity development (Archer et al., 2010), recognizing that linking abstract science to societal contributions and practical values can nurture inclusive career aspirations before negative stereotypes take root.

Methodology

This qualitative study examined how elementary teachers introduce foundational quantum concepts and how Grade 5 students respond to these emerging STEM lessons. Two research questions guided our inquiry:

1. How do elementary teachers introduce quantum science concepts, and how do Grade 5 students conceptually interpret and respond to these emerging STEM lessons?
2. What instructional, cognitive, and systemic barriers and opportunities emerge during the implementation of quantum concepts in elementary classrooms?

Data were gathered across two interconnected, National Science Foundation (NSF) funded projects in the mid-Atlantic United States. The elementary-focused initiative engaged teachers in a 30-hour professional development (PD) program (three in-person workshops and nine virtual sessions) centered on a spiraled quantum curriculum, hands-on activities, and teacher reflection, as well as collaborative curriculum development. The K–12 learning progressions initiative engaged teachers in a 15-hour summer institute examining cross-grade learning trajectories. Participants included 13 elementary teachers: nine from the elementary initiative, three from the learning progressions initiative, and one who participated in both projects. All participants were women with 2 to 21 years of teaching experience (mean = 11.2 years), serving culturally, linguistically, and socio-economically diverse urban and suburban public school districts. All names are pseudonyms.

Data sources included teacher perspectives and classroom student work. Teacher data included semi-structured interviews (30 to 45 minutes each) conducted at the beginning, middle, and end of the elementary project, baseline interviews from the learning progressions project, reflective surveys, and instructional design artifacts from PD sessions. Under approved university IRB protocols and district research agreements, teachers collected and shared de-identified student work. Student artifacts comprised several class sets of exit tickets following an integrated cryptography and quantum computing lesson (N = 42), photographs of student question anchor charts, and teacher field logs documenting student discourse. All student artifacts came from Grade 5 classrooms.

Qualitative data were analyzed using reflexive thematic analysis (Braun & Clarke, 2006). The research team reviewed interview transcripts, teacher surveys, and student artifacts to identify recurring patterns regarding instructional moves, student sensemaking, and implementation obstacles. Team members met regularly to compare initial impressions, refine codes, and group related findings into overarching themes. To support trustworthiness, we triangulated findings across data sources, comparing teacher reflections with student exit tickets and anchor charts.

Regular debriefing sessions among the co-authors, who represent science education and physics disciplines, helped examine differing interpretations and ensure that themes captured both teacher and student experiences. All participants provided written informed consent under approved IRB protocols.

Results and Discussion

Qualitative analysis revealed four primary themes addressing our research questions: (1) persistent student confusion with size and scale, (2) the risk of superficial jargon recitation during passive instruction, (3) emerging student career interest linked to altruistic and maker values, and (4) teacher self-efficacy concerns compounded by institutional pacing mandates. Across these themes, student artifacts and teacher reflections demonstrated that while upper-elementary students enthusiastically engaged with quantum ideas, their conceptual progress required addressing fundamental intuitions about the physical world.

Size and Scale as the Primary Conceptual Hurdle

Prior to instruction, fifth-grade pre-assessments indicated that while students possessed little formal knowledge of quantum science, they actively constructed naive theories based on macroscopic observations and lexical associations. Several students linked "quantum" to physical dimensions. One student predicted, "I think quantum is something microscopic," whereas another explained, "I think quantum means big, because I made the connection that quantum and quantity sound the same." Other students conflated subatomic phenomena with everyday macroscopic organisms and consumer electronics, suggesting that quantum behavior could be directly observed in insects or standard laptop screens. These naive conceptions persisted following instruction. In exit tickets collected after a lesson connecting cryptography to quantum computers, students frequently struggled to differentiate between microscopic and subatomic worlds. Rather than conceptualizing discrete, subatomic mechanisms, students tended to lump all unobservable scientific entities into a single, generic category of "tiny."

These findings align with prior research on size and scale cognition, which demonstrates that elementary students lack stable conceptual anchors across extreme orders of magnitude (Tretter et al., 2006). When students attempt to make sense of abstract scientific phenomena, they experience levels confusion, mapping rules from the macroscopic domain directly onto subatomic systems (Stieff et al., 2013). This confusion underscores a vital connection to elementary mathematics.

Conceptual progress in emerging sciences like quantum cannot occur in a vacuum; it requires explicit mathematical grounding in powers of ten, relative scale, and proportional reasoning.

Jargon Versus Conceptual Sensemaking

Analysis of student exit tickets and teacher field logs indicated a clear divide between superficial terminology and conceptual understanding. When instruction relied on passive media, students readily reproduced scientific vocabulary without underlying comprehension. Marley, a fifth-grade teacher, observed this pattern after showing curated introductory quantum videos: "A lot of them wrote something they heard [in the video], but not what they learned... they gave me jargon."

This tendency was visible across student exit tickets following the quantum computing lesson. Several students fixated on isolated technological outcomes, noting that the lesson was "about keeping your password safe" or that "a quantum computer can hack into stuff way faster than normal devices." Only a subset articulated the broader comparative mechanism, such as noting that "quantum can solve very hard problems" that are impossible for standard processors.

Teachers observed a marked shift when introducing physical manipulatives: hands-on models generated the cognitive dissonance required for real conceptual accommodation. These classroom observations support conceptual change literature: passive delivery allows learners to superficially assimilate vocabulary into unchanged macroscopic frameworks, whereas tactile models and discourse create cognitive dissonance, prompting genuine accommodation of new scientific concepts (Posner et al., 1982; Vosniadou, 1994).

Emerging Career Awareness and Motivations

Despite their conceptual struggles with scale and mechanics, students demonstrated remarkable enthusiasm toward quantum career pathways. Analysis of 42 fifth-grade exit tickets collected after the cryptography and computing lesson revealed clear preferences across both technical and applied roles. Students most frequently selected quantum scientist (n = 11, 26.2%), followed closely by quantum cryptographer (n = 10, 23.8%) and quantum hardware technician (n = 9, 21.4%), with the remaining students choosing other emerging or interdisciplinary technical pathways (n = 12, 28.6%).

Qualitative reasoning on the exit tickets showed that student interest was driven by distinct, personal motivations. Students aspiring to be scientists cited curiosity and a desire to conduct original research. Those choosing cryptography were overwhelmingly motivated by helping others, stopping cyber threats, and protecting private information. Students interested in technician roles

consistently connected their choice to physical actions, explaining that they liked to "build" and "fix stuff."

This high level of student enthusiasm, emerging from limited initial knowledge, illustrates how early STEM identities take shape. Upper-elementary children are responsive to STEM careers when roles connect to foundational values such as altruism, creativity, and hands-on problem-solving (Archer et al., 2010). Introducing the mathematical and technical workforce through both theoretical (scientist) and tangible (technician) roles widens access, demonstrating to young learners that the quantum sector extends beyond abstract mathematical physics.

Teacher Self-Efficacy and Systemic Pacing Pressures

While students responded positively to quantum lessons, teachers experienced significant apprehension regarding their own content readiness. Because elementary educators are typically generalists, teaching an emerging, counterintuitive domain created feelings of vulnerability. Brielle noted that while students wanted to learn about quantum careers, "I feel like that's something I would have to research myself first before I can kind of spill that over to them." Zoe similarly shared, "I don't think I understand well enough yet what kinds of careers? What are the practical implications of [quantum] and who would actually use it?"

This self-doubt was not limited to careers. Zoe reflected, "Why is this more uncomfortable to not have all of the answers with quantum than it is with any other topic I teach?" Nicole, a STEM specialist, echoed this unease, noting how introducing counterintuitive ideas continually generates new layers of uncertainty: "It's one of those topics that you get a little piece of information, and it brings out 20 more questions in 20 different directions... I still have lots of questions." Teachers worried that incomplete mastery might lead them to inadvertently reinforce student misconceptions. This confirms previous findings that emerging, non-intuitive science topics often lower teacher self-efficacy unless accompanied by clear curricular boundaries and professional support (Gunning & Mensah, 2011).

Institutional constraints intensified these pedagogical challenges. Teachers reported that rigid school pacing guides severely restricted instructional experimentation. Marley detailed administrative pressures to adhere to strictly synchronized schedules: "The expectation is that we are all teaching the same lesson in the same time so that if central office does a walkthrough, they'll be able to start the lesson in one room, get the middle of a lesson in another, get the end of a lesson in a room, and it will all be perfectly on pace."

Because quantum concepts are absent from state elementary standards, teachers faced a systemic bind: introducing innovative science risked falling out of administrative compliance. To navigate this barrier, several participating teachers devised cross-curricular workarounds, integrating quantum stories and informational texts into scheduled English Language Arts and literacy blocks. This structural tension echoes broader findings in STEM integration, where inflexible pacing guides and standard misalignment represent the primary systemic obstacles to innovative instruction (Margot & Kettler, 2019). Finding legitimate spaces for emerging STEM within standard elementary schedules requires cross-disciplinary alignment that honors both literacy and mathematics learning objectives.

Implications

These findings offer practical guidance for curriculum design, teacher professional development, and systemic integration in elementary STEM education. Rather than treating quantum concepts as isolated advanced physics, elementary integration requires intentional scaffolding across pedagogy, educator support, and policy structures.

Fifth-grade students easily conflate the subatomic with the macroscopic. To prevent empty jargon recitation, curriculum developers must prioritize physical manipulatives and discourse over passive video media. Furthermore, student confusion regarding scale underscores an opportunity for deliberate mathematics integration. Grounding quantum concepts in elementary mathematical standards, including relative scale, powers of ten, and proportional reasoning, provides the cognitive scaffolding students need to distinguish microscopic objects from subatomic phenomena.

Educators' hesitation and concern over inadvertent misinformation demonstrate an urgent need for targeted professional learning. Professional development must provide structured, classroom-ready materials that clearly delineate conceptual boundaries, helping elementary generalists understand which questions are appropriate for fifth graders and which concepts belong in later grades. Additionally, professional learning programs should focus on building teacher self-efficacy, reassuring educators that an inquiry-oriented stance of co-investigation can effectively replace the expectation of teacher omniscience when introducing emerging science.

Rigid pacing schedules and the absence of quantum concepts in state standards represent formidable systemic barriers to classroom implementation. Educational leaders and curriculum developers can address these constraints by advancing transdisciplinary approaches. Linking quantum concepts, biographical narratives, and societal implications directly to existing English

Language Arts and mathematics objectives allows teachers to introduce emerging science without violating administrative pacing expectations or losing essential instructional time. Introducing diverse scientific pathways early also supports inclusive STEM identity development, showing young learners that emerging technological fields value creativity, craftsmanship, and a desire to help others.

Several limitations contextualize these findings. First, the sample was limited to 13 elementary educators within one mid-Atlantic region, restricting broader generalization across varying state policy contexts. Second, participating teachers opted into grant-funded quantum programs, suggesting an openness to innovation that may not reflect the general elementary teaching population. Finally, student data were limited to de-identified classroom artifacts and teacher field reports. Future research should incorporate direct student assessments and longitudinal evaluations across varied school contexts to examine how sustained quantum analogies influence conceptual understanding and STEM identity over time.

References

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2010). "Doing" science versus "being" a scientist: Examining 10/11-year-old schoolchildren's constructions of science through the lens of identity. *Science Education, 94*(4), 617–639. https://doi.org/10.1002/sce.20399

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Gunning, A. M., & Mensah, F. M. (2011). Preservice elementary teachers' development of self-efficacy and confidence to teach science: A case study. *Journal of Science Teacher Education, 22*(2), 171–185. https://doi.org/10.1007/s10972-010-9198-8

Hasanovic, M. (2023, May). Quantum education: How to teach a subject that nobody fully understands. In *Seventeenth Conference on Education and Training in Optics and Photonics,* Article 1272331. Optica Publishing Group. https://opg.optica.org/abstract.cfm?uri=ETOP-2023-1272331

He, Y., Zha, S., & He, W. (2021). A literature review of quantum education in K–12 level. *Innovate Learning Summit,* 418–422. https://www.learntechlib.org/p/220311/

Holincheck, N., Rosenberg, J. L., Zhang, X., Butler, T. N., Colandene, M., & Dreyfus, B. W. (2024). Quantum science and technologies in K–12: Supporting teachers to integrate quantum in STEM classrooms. *Education Sciences*, *14*(3), Article 219. https://doi.org/10.3390/educsci14030219

Kelly, A. M., Darienzo, M., Wei, T. C., & Schneble, D. (2024). Quantum information science and technology professional learning for secondary science, technology, engineering, and mathematics teachers. *Physical Review Physics Education Research, 20*(2), Article 020154. https://doi.org/10.1103/PhysRevPhysEducRes.20.020154

Margot, K. C., & Kettler, T. (2019). Teachers' perception of STEM integration and education: A systematic literature review. *International Journal of STEM Education, 6*(1), Article 2. https://doi.org/10.1186/s40594-018-0151-2

Matsler, K., & Lopez, R. (2023). The quantum for all project: Teacher professional development model. *EDULEARN23 Proceedings,* 3337–3345. https://doi.org/10.21125/edulearn.2023.0919

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*(2), 211–227. https://doi.org/10.1002/sce.3730660207

Stieff, M., Ryu, M., & Yip, J. C. (2013). Speaking across levels – generating and addressing levels confusion in discourse. *Chemistry Education Research and Practice, 14*(4), 376–389. https://doi.org/10.1039/c3rp20158a

Tretter, T. R., Jones, M. G., Andre, T., Negishi, A., & Minogue, J. (2006). Conceptual boundaries and distances: Students' and experts' concepts of the scale of scientific phenomena. *Journal of Research in Science Teaching, 43*(3), 282–319. https://doi.org/10.1002/tea.20123

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*(1), 45–69. https://doi.org/10.1016/0959-4752(94)90018-3

Acknowledgements

Data collected for this study were from programs funded by two U.S. National Science Foundation projects, *EAGER: Quantum Teaching and Learning in Elementary Classrooms* (Grant #2329874) and *Quantum in the Middle: Quantum Progressions from Elementary to High School* (Grant #2508771).

NAVIGATING DISCIPLINARY BOUNDARIES: A QUALITATIVE CASE STUDY OF CROSS-DISCIPLINARY UNDERGRADUATE RESEARCH IN TEACHER EDUCATION

Lacey Fitts
Mississippi State University
llf1@msstate.edu

Abstract

Secondary science and mathematics teachers are typically trained in disciplinary silos, often resulting in superficial classroom STEM integration. To address this gap, this study investigated an undergraduate research partnership between two preservice teachers, one in physics and one in mathematics, facilitated by faculty mentors. Over several months, the pair negotiated disciplinary boundaries to co-develop and pilot a light intensity sensor lab using the Arduino Science Journal. Drawing on Akkerman and Bakker's Boundary Crossing Framework, data from interviews, observational video recordings, and collaborative documentation revision histories were analyzed.

Keywords: STEM integration, undergraduate research, boundary crossing, preservice teacher preparation

Introduction

The idea of STEM integration has been described in the literature over the last two decades, yet meaningful integration remains an elusive goal at the secondary school level. Secondary mathematics and science teachers are largely trained within discipline-specific silos. As a result, classroom STEM integration is often reduced to surface-level connections, such as executing calculations and graphing data in science class or making weak science connections through word problems in math class. To overcome superficial applications, students must engage in active sensemaking and authentic mathematical reasoning embedded in real-world phenomena. Because teacher preparation programs rarely offer cross-disciplinary learning opportunities, preservice science and mathematics teachers seldom develop knowledge of one another's content standards or pedagogical strategies, leaving them ill-equipped for meaningful integrated practice.

To address this gap, this study investigated an undergraduate research partnership between two preservice teachers, one a physics major completing alternate route teacher certification, and the other a secondary mathematics education major, who co-developed and piloted an integrated STEM

Anderson-Pence, K., & Shelton, R. N. (Eds.). (2026). *Proceedings of the 125th annual convention of the School Science and Mathematics Association* (Vol. 13). Chicago, IL: SSMA.

lab as an undergraduate research project, facilitated by two faculty members, one in mathematics education and one in science education. Over several months, the undergraduates designed and piloted a physics-mathematics activity using the free Arduino Science Journal application (Arduino, 2026). The experiment guided students through collecting light intensity data to establish the mathematical relationship between distance and light intensity.

To evaluate cross-disciplinary work, it is helpful to distinguish between the levels of content integration, a distinction clarified by the Interagency Working Group on Convergence (2022). Multidisciplinary integration organizes subjects around a common theme, while the disciplines remain distinct and isolated. Interdisciplinary integration synthesizes concepts and skills from the disciplines to deepen understanding of a shared topic. And transdisciplinary integration moves beyond individual subjects entirely, organizing learning around authentic problems where the disciplinary boundaries disappear completely. This study focuses on interdisciplinary integration, evaluating how core mathematical relationships and scientific inquiry can be mutually reinforced without diluting either.

Objectives of the Study

The purpose of this study was to examine how disciplinary boundaries were negotiated when mathematics education student and a physics education student co-designed a STEM integrated lab, and to identify features of the collaboration that supported authentic disciplinary integration. The study tracked the pair as they reviewed the literature, developed and tested the activity, piloted the lesson, and co-authored scholarly presentations and manuscripts.

The investigation was guided by two central research questions: (RQ1) How do disciplinary differences manifest in the co-design and execution of an integrated lab investigation? (RQ2) In what ways do institutional structures support or constrain collaborative, interdisciplinary undergraduate research in teacher education?

Theoretical Framework and Related Literature

This study is framed using Akkerman and Bakker's (2011) Boundary Crossing Theory within a Communities of Practice (CoP) framework (Wenger, 1998). Boundaries are defined as sociocultural differences that cause discontinuities in action and interaction (Akkerman & Bakker, 2011). Within the context of STEM education, these boundaries are sustained by the traditional disciplinary silos that separate science, mathematics, technology, and engineering in both curriculum design and initial teacher preparation (Bondurant et al., 2025; Nicholas & Scribner, 2021; Siostrom et

al., 2026). Rather than viewing boundaries purely as barriers, Boundary Crossing Theory says that boundary spaces carry rich learning potential. This learning occurs across four mechanisms: *Identification*: defining differences between domains; *Coordination*: establishing a practical and cooperative working relationship across domains; *Reflection*: exposing and expanding perspectives across disciplines; and *Transformation*: developing co-designed artifacts that transcend the boundaries.

Authentic STEM integration is a form of boundary crossing, requiring collaborators to negotiate and combine aspects from multiple disciplines to achieve a new hybrid situation. Cross-disciplinary Communities of Practice (CoP) offer space for this crossing, enabling preservice educators to work alongside peers and mentors. Classroom interaction serves as a primary site for crossing boundaries when students are engaged in project-based learning or authentic inquiry that requires a process of sense-making that synthesizes concepts across disciplinary boundaries.

Despite the promise of cross-disciplinary collaboration, systemic barriers frequently reinforce boundaries as rigid walls rather than learning sites, namely siloed teacher preparation and systemic constraints (Benito, 2022; Johnson, 2013). A recent review of STEM education policy highlights a systemic paradox: that is that teachers cannot be expected to do the work of STEM integration without explicit training, yet teacher preparation programs cannot justify designing coursework for an approach that is not supported or integrated in the K12 schools (Bondurant et al., 2025; Portillo-Blanco et al., 2024; Rimbach-Jones et al., 2025). This undergraduate research project provided an organic environment to observe how novice educators negotiate those boundaries in real time.

Methodology

A qualitative embedded single-case study approach (Yin, 2018) was used to answer RQ1, where the case was defined as the cross-disciplinary undergraduate research dyad. The data corpus included semi-structured interviews (4 transcripts) conducted separately with each subject after piloting the activity with students and again after manuscript preparation and submission, one 45-minute observational video recorded during the piloting activity with high school student aspiring educators, and artifact revision histories documenting 142 distinct revision events of a poster and manuscript.

The participants were "Ms. S" (Science perspective) and "Mr. M" (Math perspective). Ms. S was a senior physics major completing an alternate route teaching certification program with prior informal teaching experience. She brought strong domain expertise in physical phenomena,

experimental setup, and scientific inquiry. Mr. M was a freshman majoring in secondary mathematics education with minimal teaching experience. He brought strong mathematical modeling, graphing, and structural task-design focus.

To ensure analytical credibility and reduce coder bias, member checking was conducted with both preservice teacher participants (Lincoln & Guba, 1985), both of whom were engaged in prolonged project engagement (10 months, multiple phases). Following initial data analysis, draft thematic summaries and the comparative matrix (see Table 1) were shared with both participants. Each participant reviewed the findings to confirm whether the interpretations accurately reflected their lived experience, communication style, and perceived role during the project. Participant feedback confirmed the structure and provided minor refinements, which were incorporated.

The authors served as faculty research advisors initiating the partnership, providing literature review scaffolding, and facilitating research meetings. Written informed consent was secured from both subjects and from all senior high school participants (age 18+) and preservice teachers observed during the piloting exercises. The university's Institutional Review Board determined the project is exempt from further review (IRB#23-510).

Results and Discussion

Context of the Integrated Activity

The pair co-developed a sensor-based lab using the free Arduino Science Journal application. Participants used cell phone sensors accessed by the Arduino app, light bulb stations set up in a darkened room, and meter sticks, plotting the measured intensity of light in lux as the distance from the light source decreased. Student participants analyzed, transformed, and discussed their data and compared it with the expected curve. They were guided in discussing the results and deviations from expected data, and in developing a mathematical model of the Inverse Square Law, $I \propto \frac{1}{d^2}$, where I is the light intensity measured in lux and d is the distance from the source. This activity was designed to anchor abstract mathematical functions in direct physical observation. The project was presented at the university undergraduate research symposium, the state science teachers' conference, and they co-wrote an article for science teachers (in review).

Identification: Navigating Disciplinary Roles and Status

Identification in boundary learning mechanisms involves outlining the domain boundaries and recognizing distinct expertise. Despite a clear status gap (senior physics major versus freshman

math ed major), both participants naturally divided leadership based on content expertise rather than academic standing. Ms. S took ownership of physics explanations and contextualizing phenomena, stating, "I wanted to make sure they didn't just see a number, but understood why the light drops off so fast, like connecting it to astronomy or how light bulbs work in a room." Mr. M spearheaded questioning about data organization, graphed results, and mathematical modeling. Neither student viewed their disciplinary division as limiting the collaboration, and both expressed the other's expertise as being essential to producing an integrated activity.

Coordination: Developing Routines and Managing Conflict

Coordination requires establishing the partnership routines. A difference was noted in their perception of conflict. Both explicitly maintained that there was no conflict or disagreement throughout the project, although differences in their working styles and priorities were noted. Ms. S said, "We never disagreed. We kept the work on safe ground." Despite confirming that they did not disagree, Mr. M said, "We didn't work the same way. She would sometimes see several different routes we could take at once, while I often try to take one idea and think it through fully before trying something different." Their interviews and observational data revealed fundamentally different priorities, roles, and ways of approaching the task. Mr. M noted needing to consciously remind himself not to treat the activity as purely a math exercise, and acknowledged the different terms used in the fields. When working with students, Ms. S says she prioritized alleviating student physics anxiety and fostering engagement, whereas Mr. M prioritized maintaining the mathematical integrity of the graphs and data representations.

Reflection: Exposing and Expanding Perspectives

Reflection occurs when the participants evaluate their home discipline through their peer's lens, and evidence of reflection emerged in their interviews. Ms. S said in the post-project interview, "because we were coming from two sister disciplines, we used our different angles to gain a deeper, more complete understanding. There was no single wrong answer [in what we should do], only different insights that brought us to a stronger final outcome." Observation of the live 52-minute piloting session showed that their speaking times were comparable (~12 minutes each), but the content of their discourse differed. Ms. S spent significantly more time probing student understanding of the physical phenomena, making connections to other topics in physics, and discussing teaching techniques. Mr. M spent more time questioning students about the collected

data, the trends in the graphs, axis labeling, and asking questions about the mathematical relationships.

Transformation: Co-creation of Hybrid Pedagogical Artifacts

Analysis of their document version histories in the presentation and paper further demonstrated true transformation (Akkerman & Bakker, 2011). Rather than writing separate sections, they engaged in cross-editing, adjusting the style and content of each other's slides, whether the slides were more mathematical or scientific in nature. A total of 142 instances of document revision were noted by the duo, happening in multiple sections involving both math and science. Some evidence of transformation was evident in their statements. Ms. S summarized, "Ultimately, science provided the 'why' behind our investigation, while math gave us the rigor to prove 'how' it worked." Mr. M reflected, "My math background shaped this project the most at the beginning. At the start, I really tried to keep that distinction. As time went on, the distinction between the two started to blur. . . other times the integration was so seamless that I forgot it was a science and a math-based topic." Other comments revealed that although evidence of transformation was present, the participants viewed their product as a science project. Mr. M said, "The article we wrote was for a science teacher publication. We really focused more on the science of light and less on the mathematical conversions." Ms. S confirmed, "I think this was primarily a science project, heavily empowered by mathematics. At its core, the project was driven by exploring physical phenomena and testing hypotheses. There is a substantial math component, but the math didn't exist in a vacuum. It served as the language that allowed us to model, prove, and explain the science." Ultimately, the pair together produced a hybrid lesson, presentation, and journal article that neither student could have created independently. A summary of comparative findings is presented in Table 1.

Table 1

Cross-Disciplinary Interview Comparison

Analytical Theme	Ms. S (Physics)	Mr. M (Math)	Boundary Crossing Mechanism
Primary Focus	Making abstract physics concepts tangible and reducing student anxiety	Establishing structural integration where math remains explicit	**Identification –** Science: Phenomenon-first approach Math: Balance and structure
Division of Leadership and Roles	Led physics content explanations, qualitative setup, and slide creation	Led task structure, graphing, data analysis, and technical data display	**Coordination** – Task allocation aligned naturally with their home discipline strengths
Perceptions of Conflict	Explicitly stated zero conflict, attributed to keeping the work "on safe ground."	Noted subtle differences and described negotiation, but reported no conflict	**Coordination** – Perceived harmony, but negotiated many boundaries
Core Values	Affective student connections, scientific literacy, real-life context	Student enjoyment, equal representation of math and science	**Reflection –** Science: Values affective / literacy Math**:** Values structural / Integrity focus
Data Outliers	Interpreted outliers as evidence of students' diverse thinking styles	Expected data errors and viewed them as typical or manipulation of data	**Reflection –** Science: Qualitative insight Math: Expected statistical variance
Mentorship versus Autonomy	High satisfaction with mentor guidance and felt supported	Desired more autonomy, feeling mentors over-assisted in presentation support	**Transformation** – Asserting independent agency in artifact creation; Highlights tension between faculty guidance and student agency.

Institutional structures created both opportunities and obstacles for this team (RQ2). While faculty collaboration provided the necessary physical and conceptual space, an untended barrier arose during the undergraduate research symposium. While the official guidelines explicitly state that group research is encouraged, the registration logistics restricted posters to a single 'principal presenter,' issuing a badge and presentation rights were given to only one student. Unaware of this

operational constraint prior to the event, the team designed the project as a fully shared presentation. So, despite co-designing the lab and analyzing the data together, the physical constraints of the symposium forced one student (the first alphabetically) to act as the sole voice for both disciplines. Even when institutional policies explicitly encourage group research, specific operational procedures, such as designated primary presenter designations, badge allocations, or single-speaker logistics, can inadvertently impose an artificial hierarchy on collaborative student teams.

Implications and Recommendations

To help preservice teachers foster authentic STEM collaboration across disciplinary lines, we propose four concrete facilitation strategies as practical recommendations. The first is intentional boundary pairing of preservice teachers across complementary disciplines early in their programs to normalize boundary-crossing and collaboration before clinical practice. Another is interdependent task design where the activities are structured so that neither discipline functions as a service tool for the other; both mathematical reasoning and scientific inquiry must be necessary to solve the problem. The last is guiding cross-disciplinary preservice partners to explicitly discuss discipline-specific assumptions, terminology, and priorities rather than assuming alignment. An important consideration that we learned the hard way is learning institutional logistics for such processes as symposium guidelines and presentation policies.

Several limitations must be considered when interpreting these findings. This was a single case study that cannot be generalized across other contexts. The participants possessed an inherent status and experience gap that may have influenced deference patterns differently than a peer pairing of equal academic standing. The strong faculty mentor presence throughout the project provided substantial scaffolding, which may have masked conflicts or decision-making struggles that might occur in unguided teams. The boundary crossing dynamics could be quite different in contexts other than the co-design of a single sensor lab.

Breaking the historical silos of secondary teacher preparation requires intentionally designed spaces where preservice educators can cross disciplinary boundaries before entering the classroom. As this case study demonstrates, when mathematics and science preservice teachers engage in authentic, cross-disciplinary co-design within an undergraduate research Community of Practice, boundary crossing transforms tensions into productive pedagogical opportunities. Structured research partnerships can encourage novice teachers to recognize, respect, and synthesize their own

disciplinary perspectives. By aligning teacher education training, task design, and institutional structures to support cross-disciplinary collaboration, teacher preparation programs can better equip future educators to break down disciplinary walls and implement meaningful, integrated STEM learning in K–12 schools.

References

Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research, 81*(2), 132–169. https://doi.org/10.3102/0034654311404435

Arduino Education. (2026). https://www.arduino.cc/education/science-journal/

Benito, R. (2022, October 25). *Guidelines to overcome barriers to STEM integration*. STE(A)M IT: An Interdisciplinary STEM Approach. https://steamit.eun.org/overcome-barriers-guidelines/

Bondurant, L., Fitts, L., Ivy, J., Franz, D. P., & Wan, A. (2025). A review of U.S. education policy on integrating science and mathematics teaching and learning. *Education Sciences, 15*(12), Article 1687. https://doi.org/10.3390/educsci15121687

Interagency Working Group on Convergence. (2022). Convergence education: A guide to transdisciplinary STEM learning and teaching. Executive Office of the President, Office of Science and Technology Policy. https://lemelson.mit.edu/sites/default/files/2025-02/Convergence_Public-Report_Final.pdf

Johnson, C. C. (2013). Editorial: Conceptualizing integrated STEM education. *School Science and Mathematics, 113*(8), 367–368. https://doi.org/10.1111/ssm.12043

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.

Nicholas, C., & Scribner, J. A. (2021). Enhancing PBL authenticity by engaging STEM-professional volunteers. *Interdisciplinary Journal of Problem-Based Learning, 15*(2). https://doi.org/10.14434/ijpbl.v15i2.28734

Portillo-Blanco, A., Deprez, H., De Cock, M., Guisasola, J., & Zuza, K. (2024). A systematic literature review of integrated STEM education: Uncovering consensus and diversity in principles and characteristics. *Education Sciences, 14*(9), Article 1028. https://doi.org/10.3390/educsci14091028

Rimbach-Jones, D., Harper, F. K., & Brown, C. L. (2025). Building primary teachers' capacity for integrated STEM education: A case study of programmatic features and structures. *Education Sciences, 15*(12), Article 1657. https://doi.org/10.3390/educsci15121657

Siostrom, E., Strachan, A., Marshman, M., Scott, J. J., & McMaster, N. (2026). Australian teacher educators' perceptions of integrated STEM practices and barriers in primary initial teacher education. *Teaching and Teacher Education, 176*, Article 105540. https://doi.org/10.1016/j.tate.2026.105540

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage.