J-STEM

Journal of Research in STEM Education

Vol 8, No 2, December 2022

j-stem.net

EDITORIAL BOARD

Contact

All queries related to manuscript submissions can be directed to Dr. Aydeniz, the Editor-in-Chief, editorial@j-stem.net and jstem.editorialoffice@gmail.com

Mehmet Aydeniz, PhD.
Professor of Science Education.
Program Coordinator, Science Education.
Department of Theory and Practice in Teacher Education
The University of Tennessee, Knoxville
A 408 Jane & David Bailey Education Complex
Knoxville, TN 37996-3442
USA
Phone: +1-865-974-0885

Publisher / Founder : i-STEM / Gultekin Cakmakci
ISSN:2149-8504

Contents

ISSN:2149-8504 (online)

Journal of Research in STEM Education
Vol 8, No 2, December 2022, 35-60
https://doi.org/10.51355/jstem.2022.110

RESEARCH REPORT

Effectiveness of an Inquiry Focused Professional Development: Secondary Mathematics and Science Teachers' Beliefs and Instruction

Jennifer Cribbs[a1], Lisa Duffin[b], Martha Day[b]

[a]*Oklahoma State University, USA;* [b]*Western Kentucky University, USA*

Abstract:
Secondary (grades 6th-12th), mathematics and science teachers participated in a two-year inquiry-based professional development (PD) program focused on inquiry-based instruction. This study draws from surveys and classroom observations to assess potential changes in teacher beliefs (Teaching Philosophy, Openness to Change, Job Satisfaction, Professional Commitment, and Inquiry) and instructional practices using the electronic quality of inquiry protocol (EQUIP). Results of a one-way repeated-measures ANOVA found significant increases in participating teachers' Teaching Philosophy, Openness to Change, Confidence toward Inquiry, and Intentions toward Inquiry. Findings also indicate significant changes in teachers' instructional practice with teachers participating in the PD implementing higher levels of inquiry instruction in their classroom. Finally, a two-way repeated-measures ANOVA found statistically significant differences in participating teachers' Teaching Philosophy, Openness to Change, Confidence toward Inquiry, and Intentions toward Inquiry when evaluated with a comparison group of teachers. Overall, results indicate changes in teachers' beliefs and use of inquiry in their classroom due to their participation in the PD.

Keywords: Teacher professional development, inquiry instruction, mathematics education, science education, secondary education

1 Corresponding Author: 237 Willard, Stillwater, OK 74078, Email: jennifer.cribbs@okstate.edu

To cite this article: Cribbs, J., Duffin, L., & Day, M. (2022). Effectiveness of an inquiry focused professional development: Secondary mathematics and science teachers' beliefs and instruction. *Journal of Research in STEM Education, 8*(2), 35-60. https://doi.org/10.51355/jstem.2022.110

ISSN: 2149-8504 (online)

Initiatives in mathematics and science education – including the transition to new standards – call for teachers' use of reform-based instruction in the classroom (National Council of Teachers of Mathematics [NCTM], 1989; 2000; 2014; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010; National Research Council [NRC], 2000; 2012; NGSS Lead States, 2013). Research supports these calls indicating improved learning when teachers are using reformed-based practices in both mathematics (Cain, 2002; Mac Iver & Mac Iver, 2009) and science classrooms (Furtak et al., 2012; Minner, Levy, & Century, 2010). Students in these settings also have increased interest and motivation for learning (Brown et al., 2013; Cichon & Ellis, 2003; Jiang & McComas, 2015).

With these initiatives and supporting research, there is a need to explore effective models of professional development (PD) for how to assist teachers in transitioning from a more traditional/direct instruction approach to a more reformed/inquiry approach in their classrooms. In this study, we highlight a model for PD with secondary mathematics and science teachers and examine its effectiveness for changing teachers' beliefs and classroom practice. Prior research exploring PD programs and beliefs related to inquiry-based instruction provide evidence to support how PD programs can lead to changes in teacher beliefs and their use of inquiry-based instruction (Yow & Lotter, 2014). However, many of these efforts do not consider these changes with respect to a comparison group of teachers, possibly accounting for changes that may evolve as teachers acquire more experience in the classroom. Additionally, while research notes the benefit of PD on transitioning beliefs related to reform- or inquiry-based instruction (Carney, 2016) and there is evidence to support teacher beliefs influencing teacher practice (Lloyd, 2002), this study examines the influence of the PD program on teachers' beliefs and classroom practice.

Literature Review

Inquiry Instruction

Inquiry instruction is a broad description of a variety of practices that support a student-centered method of instruction with a focus on conceptual understanding of content. The National Research Council (NRC, 1996) describes scientific inquiry as:

> asking questions, planning and conducting investigations, using appropriate tools and techniques to gather data, thinking critically and logically about relationships between evidence and explanations, constructing and analyzing alternative explanations, and communicating scientific arguments (p. 105).

Calls for reformed instruction in both mathematics and science support the use of inquiry in the classroom (NCTM, 1989; 2000; 2014; NGSS Lead States, 2013). Marshall (2013) proposes a continuum of inquiry that ranges from highly teacher-structured to completely student-driven (with strong teacher support). This learner-centered perspective of engagement allows students

ISSN: 2149-8504 (online)

to actively construct their understanding of concepts by solving challenging tasks that require executive function, self-regulation, and meaningful discourse – skills that support deep conceptual learning – while being carefully scaffolded by the teacher (Bransford et al., 2000; NCTM, 2014; NRC, 2012). For example, teachers can develop inquiry-based instruction using existing tools like the 5E inquiry model – originally developed for science education (Bybee et al., 2006) or the 4E x 2 inquiry model – developed for both mathematics and science education (Marshal & Horten, 2009). In both instructional models, there are key phases of instruction: Engagement, Exploration, Explanation, Elaboration (5E) or Extension (4E x 2), and Evaluation. In both models, evaluation is iterative and occurs throughout the lesson to allow students to gain on-going feedback about their learning (see Marshal and Horton for a detailed explanation and comparison of the models). Research indicates a positive connection between inquiry-based instruction and the following: student achievement and understanding (Granger et al., 2012; Kang & Keinonen, 2018; Koksal & Berberoglu, 2012; Lazonder & Harmsen, 2016; Marshall & Horton, 2011; Minner et al., 2010), interest and motivation (Chen et al., 2014; Fielding-Wells et al., 2017), and attitudes toward science and mathematics (Kim, 2016; Riegle-Crumb et al., 2019). Knowing the benefit of inquiry instruction on a variety of student factors and measures, it is important to consider how to support teachers' use of this method of instruction in their classrooms.

Teacher Beliefs and Professional Development

Teacher beliefs – or "inferences made…about underlying states of expectancy (Rokeach, 1972, p. 2)" in teaching or learning – are part of an integrated cognitive system that are integral for enacting behavioral change (Rokeach, 1972). However, not all beliefs are equally important. Those that are most central to a person will be the most resistant to change, but if changed, will have the most profound repercussions on the rest of one's belief systems (Pajares, 1992; Richardson, 1996; Rokeach, 1972). Teachers' beliefs about: 1) themselves as professionals (professional identity), 2) inquiry as an effective pedagogical tool, 3) their perceived capabilities to design and implement inquiry-based instruction, and 4) carrying out inquiry in their future classrooms are known to influence teacher practice (e.g., Canrinus, et al., 2012; Cross, 2009; Hofer & Pintrich, 2004) and are germane to this study. Beliefs and experiences are interwoven (Raths, 2001), and teachers have a wealth of experiences both as a student and as a teacher that influence their beliefs (Lortie, 1975). Thus, teacher beliefs are an important but complex topic to research.

Creating dissonance between teachers' beliefs, practice, knowledge, and experiences is one way that literature recommends creating a space to enact teacher change (Woolfolk et al., 2009). However, even this idea can be complex as too much dissonance may lead to teachers' rejecting changes (Timperley & Alton-Lee, 2008). Drawing on a large body of literature related to teacher professional development is helpful for designing programs that effectively facilitate teacher change. For example, research repeatedly cites the importance of professional development being content focused (Desimone, 2009; Taylor et al., 2017). Meaning that mathematics and science educators should be experiencing discipline specific PD rather than general PD, which is often the case with district and school provided PD. In fact, Moyer-Pakenham et al. (2010) noted that PD for mathematics and science teachers often only included a few measures assessing the effectiveness of PD activities and do not connect back to teachers' classrooms. Current literature focused on PD and teachers' use of reform- or inquiry-based instruction also highlights the importance of job-embedded experiences in order to transition beliefs to changes in practice (Shirrell et al., 2019). This focus aligns with other literature noting the positive effect that school-based initiatives, such as a professional learning community, can have on teacher change (Tam, 2015).

Based on the work of Desimone (2009) and Darling-Hammond and colleagues (2017), there are seven core features of PD that lead to subsequent change. Each of the seven features are supported by research as being effective, and thus, utilized in the design and evaluation of the PD described in this study. Table 1 provides an overview of the seven features along with supporting research citations.

Table 1.
Features of Effective PD

PD Feature	Brief Description	Related Research
Content focused	PD is discipline specific and include opportunities to explore curricula, examine student work, work on or teach lessons, and engage in strategies that are specific to the content of focus.	Gallagher et al. (2017); Johnson & Fargo, (2014)
Active learning	PD involves teachers being active participants in the learning process, which embodies the inquiry approach to learning that is often the focus of current PD. This feature also attends to literature on adult learning (e.g., Knowles, 1990; Trotter, 2006).	Allen et al. (2011); Gallagher et al. (2014)
Collaboration	PD includes teachers collaborating with one another or supporting personnel, such as coaches, and involve authentic experiences (in the school setting, with activities that related directly to the teachers' classroom and practice). These collaborations could be one-on-one meetings or be small groups and could include teachers reflecting, planning, engaging in activities, and solving problems.	Allen et al. (2015); Buczynski & Hansen (2010)
Models and modeling	PD provides models of the specific instructional practices/curriculum that are the topic of focus. These models can be included in a variety of formats such as case studies, model lessons, and student work.	Doppelt et al. (2009); Greenleaf et al. (2011)

ISSN: 2149-8504 (online)

Coaching and expert support	PD includes some type of coaching or expert support, which could include other teachers, content experts, instructional coaches, or university faculty. This support could be included as part of workshop activities, meetings at the school to observe and reflect, or virtual meetings such as videos shared and online discussions.	Kleickmann et al. (2016); Meissel et al. (2016)
Feedback and reflection	PD includes opportunities for feedback and reflection, "often employed during mentoring and coaching" (Darling-Hammond et al., 2017, p. 14).	Kutaka et al. (2017); Landry et al. (2009)
Duration	PD is on-going, multiple activities over time, rather than being a one-time activity/workshop.	Meyers et al. (2016); Polly et al. (2015)

The purpose of this study is to examine how secondary mathematics and science teachers' beliefs and practice changed through their participation in a 2-year PD program. The research questions guiding this study are:

1. Are there changes in participating secondary mathematics and science teachers' beliefs (teaching philosophy, job satisfaction, openness to change, and professional commitment) from the beginning to the end of the PD?
2. Are there differences in beliefs between the participating and comparison group of teachers?
3. Are there changes in participating secondary mathematics and science teachers' beliefs toward, confidence in, and intentions to use inquiry from the beginning to the end of the PD?
4. Are there changes in teachers' beliefs toward, confidence in, and intentions to use inquiry between the participating and comparison group of teachers?
5. Are there changes in participating secondary mathematics and science teachers' level of inquiry instruction in their classroom?

Methods

Professional Development Model

The professional development was designed to highlight key features that research noted to be effective in changing teacher practice and affect student outcomes. Table 2 provides an overview of how features of the PD implemented over the two years aligned with features of effective PD outlined in the research.

ISSN: 2149-8504 (online)

Table 2.
Alignment of PD to Research-Based Practices

PD Feature	Alignment
Content focused	The PD program was designed specifically for mathematics and science teachers at the secondary level with instructional coaches and facilitators who had experience teaching mathematics and science at the secondary level. Instructional practices modeled were specific to mathematics and science education, and were research-based.
Active learning	Each activity during the academic year and summer involved active learning. For example, all teachers actively observed model lessons. After evaluation and discourse about the model lessons, teachers designed their own inquiry-based lessons, presented lesson ideas in workshop form to coaches and peers, received feedback, made revisions, and implemented lessons in the classroom where they were observed and formally evaluated using the observation protocol. Evaluation scores were then provided to the teachers so they could decide on goals they wanted to work on during the year (**job-embedded**). During the summer institutes, teachers learned about new strategies through experiential activities, designed problem- or project-based units, presented ideas, and received extensive feedback.
Collaboration	Teachers were intentionally placed in teams based on the district in which they worked to allow for collaboration at their school sites (**job-embedded**). Additionally, teachers were put into teams during the summer institute and worked together to create/modify lessons and units of instruction for the subsequent academic year (**job-embedded**).
Models and modeling	While model inquiry-based lessons were conducted during PD events, model lessons were also implemented in the teachers' classrooms with their students (**job-embedded**).
Coaching and expert support	An instructional coach partnered with groups of teachers to provide them continual coaching and support. This support often occurred at the teachers' location (**job embedded**), but also took the form of emails, phone calls, and team meetings during PD events.
Feedback and reflection	Teachers were provided feedback through a variety of data sources: during coaching sessions with master teachers, with observational data collected throughout the two years using the inquiry instruction research protocol, through teacher interviews and surveys, and through professional learning community discourse. Reflection occurred throughout the study using the aforementioned data sources.
Duration	This program was designed to be two-years in length and included activities throughout the academic year and a week-long intensive summer session each year.

Research Design

The purpose of the study was to examine the impact of a research-based professional development program (see Table 2) on participating teachers' beliefs in comparison with a group of teachers who did not participate in the PD. Aligned with this purpose, a repeated measure quasi-experimental survey research design was used. In addition, the study explored the impact the PD had on participating teachers' instruction. Aligned with this purpose, a nonexperimental repeated measure research design with an observational protocol was used.

ISSN: 2149-8504 (online)

Participants

There were two primary sources of data for Year 1 and Year 2 of the project including mathematics and science, middle and high school teachers (participating and comparison). It is important to note that during Year 1 of the project, some teachers opted out of participating after beginning project activities and were replaced by other teachers within the same district. This attrition left a sample size of 17 teachers who had completed a pre-, mid-, and post-administration of the survey. Participant selection was dictated by letters of collaboration from school districts agreeing to partner on PD activities as part of a grant-funded project. While some teachers within these districts chose to join the PD program, some were assigned to participate in the PD by their administrators. Additional comparison data were collected from a group of mathematics and science teachers at middle and high schools in adjacent districts in the state, sample size of 14 (with pre and post data). Demographic information for the participating and comparison group of teachers is provided in Table 3 and is representative of the teacher population within the districts.

Table 3.
Demographic Information for Participating and Comparison Groups of Teachers

		Race			Subject Taught		Grade Level Taught		Highest Level of Education			
	F	Cau.	Asian	Other	M	S	6th-8th	9th-12th	BS	MS	PhD/EDD	Mean YearsExp.
P	14	17	0	0	6	11	9	8	8	8	0%	7.2
C	6	12	1	1	3	11	3	11	7%	6	1	8.8

Note. P = participating; C = comparison; F = Female; Cau. = Caucasian; M = mathematics; S = science

Data Collection

Primary sources of data included teacher surveys and teacher observations. A combination of these sources of data provided information on the impact of the program on teacher outcomes. Several steps were taken to ensure the reliability and validity of data collected. The teacher survey was developed using existing instruments with existing reliability and validity information available (Meyer et al., 1993; Starr et al., 2006; Stearns et al., 2014; Vannetta & Nancy, 2004). Reliability scores using Cohen's d was also calculated for each factor.

Professional identity is an umbrella term for one's personal evaluations of self within a profession that can be comprised by the following factors: (1) Teaching Philosophy – 0.82, (2) Openness to Change – 0.73, (3) Job Satisfaction – 0.72, and (4) Professional Commitment – 0.86. Teaching Philosophy and Openness to Change both drew from the Teacher Attribute Survey (TAS; Vannetta & Nancy, 2004). The Teaching Philosophy scale "measured teacher support of a teacher-centered or student-centered instructional environment", while the Openness to Change construct measured teachers' "comfort and excitement when trying new methods of instruction as well as willingness to take risks and make mistakes" (Vannetta & Nancy, 2004, p. 255). Job Satisfaction drew from Stearns, Mickelson, and Moller's (2014) study exploring the construct, which included three items such as "I really enjoy my present teaching job." Occupational Commitment drew from Meyer and colleagues (1993) work, focusing on affective professional commitment with six items such as "Teaching is important to my self-image" and "I dislike being a teacher." It is important to note that Meyer's et al. (1993) was specific to the field of nursing; however, Canrinus et al. (2012) validated the items for the construct specific to the teaching profession.

Teacher's beliefs about inquiry instruction, perceived competence in designing and implementing inquiry-based instruction (confidence), and intention to use inquiry-based practices in their classroom were measured using a modified version of the measures created and implemented by Forbes and Zint (2011). Participants completed 7 items within each measure that were mathematics or science-specific based on their content-area focus. Each measure featured the same seven inquiry-based practices – e.g., "Ask questions and make predictions about [mathematical or scientific] concepts." -- but had a different question for the participant to consider. To assess participants' beliefs about inquiry, for example, the leading question was "When I am teaching [mathematics or science], I should design instruction that requires my students to…" followed by the seven practices and a 7-point Likert scale ranging from "strongly disagree" to "strongly agree". The factors and corresponding reliability scores using Cohen's d are as follows (1) Beliefs toward Inquiry – 0.92, (2) Confidence toward Inquiry – 0.93, and (3) Intentions toward Inquiry – 0.95.

The EQUIP observational protocol was used measure inquiry-based instruction of mathematics and science teachers (Marshall et al., 2008). The factors and corresponding reliability scores for the EQUIP observational instrument are as follows (1) Instructional – 0.94, (2) Discourse – 0.94, (3) Assessment – 0.89, and (4) Curriculum – 0.88. The Instructional factor included five items that relate to the type of strategies being used, sequence of instruction, role of the teacher and student and knowledge acquisition. The Discourse factor included five items that related to the level, complexity, and ecology of questions, communication pattern, and classroom interactions. The Assessment factor included five items including prior knowledge, conceptual development, student reflection, assessment type and the role of assessment. Finally, the Curriculum factor included four items related to content depth, learner centrality, integration of content and investigation, and organization and recording information. The scale on the instrument was 0 (not observed) to 4 (exemplary inquiry).

To ensure consistency in how the observational protocol was administered, every classroom observer was trained on the instrument. This training began by watching informational videos on the Inquiry In Motion website (iim.sites.clemson.edu) followed by watching classroom videos and scoring these videos using the EQUIP protocol. The website provided scores so the raters could compare their rating with the “key” scores on the website. After completing these activities, the rater team met to discuss the EQUIP protocol in detail and go through another series of watching videos as a team and scoring teachers’ use of inquiry using the EQUIP protocol. These meetings included discussions that assisted in clarifying and eventually reaching consensus on components of the protocol. Further, inter-rater assessments occurred during the first implementation of the EQUIP protocol with a pair of observers collecting data on the same lesson. The second implementation involved one member of the research team accompanying two of the raters to verify consistency in scoring. Subsequent rounds of scoring were done independently since the raters had extensive experience with the instrument.

Data Analysis

For the participating teachers, the beliefs survey was implemented three times over the course of the two years: beginning of the spring semester in year 1 (pre-assessment), at the conclusion of the summer institute in year 1 (mid-assessment), and at the conclusion of the summer institute in year 2 (post-assessment). For the comparison group of teachers, the beliefs survey was implemented twice: at the beginning of the spring semester in year 1 (pre-assessment) and at the end of year 2 (post-assessment). To compare differences between the three time intervals (pre, mid, and post) for the participating teachers, a one-way repeated-measures ANOVA was conducted. To compare potential differences in means between the participating and comparison group of teachers, a two-way repeated-measures ANOVA was conducted. In addition, a Wilcoxon signed-rank test was performed to determine if differences existed between the pre- and post-assessment for the comparison group of teachers. Random imputation was conducted to address missing data.

The EQUIP protocol was also implemented four times during the two-year PD: at the end of the fall semester in year 1(pre-observation), at the end of the spring semester in year 1 (mid1-observation), at the end of the subsequent fall semester in year 2 (mid2-observation), and at the end of the spring semester of year 2 (post-observation). Although a one-way repeated ANOVA would have been ideal for determining if differences existed between administrations of the protocol, a Wilcoxon signed-rank test was conducted instead because of the small sample size due to teacher attrition. To better understand potential differences in the program a Wilcoxon signed-rank test was run for the following (1) pre to post (N=14) and (2) mid2 to post (N=27).

Results

Research Question 1

Results of the one-way repeated-measures ANOVA assessing changes in teacher beliefs, pre-, mid-, and post-survey are reported in Table 4 for each of the four factors explored, including: (1) teaching philosophy as it relates to teachers' beliefs about a teacher- versus a student-centered classroom (Teaching Philosophy); (2) Job Satisfaction; (3) Openness to Change; and (4) Professional Commitment.

Results indicated that there was a significant change in teachers' Teaching Philosophy, $F(1, 32) = 14.842$, $p = 0.000$, $\eta^2 = 0.200$, and Openness to Change, $F(1, 32) = 4.565$, $p = 0.018$, $\eta^2 = 0.059$. The Mauchly's test for sphericity for Teaching Philosophy and Openness to Change was met (W = 0.767, $p = 0.136$; W = 0.914, $p = 0.510$ respectively) so the Huynh and Feldt correction was not used. A post hoc Bonferroni test for Teaching Philosophy showed that the pre- and mid-assessment differed significantly at $p = 0.002$ as well as the pre- and post-assessment at $p = 0.002$. A post hoc Bonferroni test for Openness to Change showed a marginally significant difference between the mid- and post-assessment at $p = 0.090$ as well as a statistically significant difference between the pre- and post-assessment at $p = 0.040$. Although the Job Satisfaction and Professional Commitment constructs were not statistically significant, it is worth noting that the means for these factors trended upwards over the course of the two years.

Table 4.
Results for ANOVA (Pre-Mid-Post) – Participating Teachers

Teacher Belief Factors	Pre *M (SE)*	Mid *M (SE)*	Post *M (SE)*	F
Teaching Philosophy[a]	4.17(0.16)	4.76(0.14)	4.76(0.12)	14.84***
Job Satisfaction[b]	4.27(0.17)	4.45(0.13)	4.51(0.15)	0.94
Openness to Change[a]	4.64(0.18)	4.75(0.13)	5.00(0.14)	4.57*
Professional Commitment[c]	6.20(0.18)	6.31(0.15)	6.25(0.17)	0.67

*Note: †p<0.10 *p<0.05 **p<0.01*
[a] – Scale of 1-6, [b] – Scale of 1-5, [c] – Scale of 1-7

Research Question 2

A Wilcoxon signed-rank test indicated there were no differences for the comparison group of teachers from pre- to post-survey. In fact, for three of the four factors, there was a declining trend in the mean from the baseline to post survey for this group of teachers (See Table 5).

Table 5.
Results of Wilcoxon signed-rank test: Comparison Teachers, Pre – and Post-Survey

Teacher Belief Factors	Pre Survey *M (SE)*	Post Survey *M (SE)*	*p-value*
Teaching Philosophy[a]	3.66(0.25)	3.71(0.20)	0.681
Job Satisfaction[b]	4.17(0.19)	4.02(0.21)	0.476
Openness to Change[a]	4.44(0.19)	4.33(0.21)	0.421
Professional Commitment[c]	5.90(0.26)	5.89(0.27)	1.000

[a] – Scale of 1-6, [b] – Scale of 1-5, [c] – Scale of 1-7

Finally, a two-way repeated-measures ANOVA was conducted pre- to post-survey to determine if differences existed for beliefs measured between groups (participant and comparison). Mauchly's test for sphericity is only applied with more than two levels. Since there are only two levels for our ANOVA, Levene's test for equal variance was conducted. This assumption was not violated for any of the variables, allowing us to proceed with the analysis. Results of the ANOVA indicated a significant interaction effect between teacher groups and the time interval for Teaching Philosophy, $F(1, 29) = 10.58$, $p = 0.004$, $\eta^2 = 0.044$. A follow up independent samples t-test found a significant difference between the post-survey responses with the participating teachers reporting a higher mean (M = 4.76) than the comparison group of teachers (M = 3.67). Results of the ANOVA also indicated a marginally significant interaction effect for Job Satisfaction, $F(1, 29) = 3.14$, $p = 0.087$, $\eta^2 = 0.022$. A follow up independent samples t-test found a significant difference between the post-survey responses with the participating teachers reporting a higher mean (M = 4.51) than the comparison group of teachers (M = 4.02). Openness to Change also had a significant interaction effect when conducting a two-way repeated-measures ANOVA, $F(1, 29) = 4.88$, $p = 0.035$, $\eta^2 = 0.030$. The follow up independent samples t-test found a significant difference between the post-survey responses with the participating teachers reporting a higher mean (M=5.00) than the comparison group of teachers (M = 4.33). Professional Commitment was not significant when conducting a two-way repeated-measures ANOVA, $F(1, 29) = 0.15$, $p = 0.703$, $\eta^2 = 0.000$. Table 6 provides a summary of results for the independent samples t-tests and Figure 1 provides a visual representation of changes pre- to post-survey for the participating and comparison groups of teachers.

ISSN: 2149-8504 (online)

Table 6.
Independent Samples T-test for Post-survey between Groups of Teachers

Factor	Part. M(SD)	Comp. M(SD)	t	*p-value*	d
Teaching Philosophy[a]	4.76(.50)	3.67(.72)	4.987	0.000***	1.79
Job Satisfaction[b]	4.51(.61)	4.02(0.78)	1.945	0.062†	0.71
Openness to Change[a]	5.00(.58)	4.33(0.78)	2.753	0.010**	0.99

*Note: †p<0.10 **p<0.01 ***p<0.001*
[a] – Scale of 1-6, [b] – Scale of 1-5

Figure 1. Change in Beliefs Pre- to Post-Survey for Participating and Comparison Groups of Teachers

Research Question 3

The researchers asked participants to respond to a series of questions designed to assess their beliefs about, confidence in their abilities to design, and intentions for implementing instruction that is inquiry based. Results of the one-way repeated-measures ANOVA assessing changes in teacher beliefs, pre-, mid-, and post-survey are reported in Table 7 for each of the three factors explored, including: (1) Beliefs toward Inquiry, (2) Confidence toward Inquiry, and (3) Intentions toward Inquiry.

Results indicated there was not a significant change in teachers' Beliefs toward Inquiry, $F(1, 32) = 1.480$, p = 0.243, $\eta^2 = 0.031$. However, it is worth noting that the mean for this factor trended upwards over the course of the two years. A significant change was found in teachers' Confidence toward Inquiry, $F(1, 32) = 15.546$, $p = 0.000$, $\eta^2 = 0.161$, and Intentions toward Inquiry, $F(1, 32) = 27.994$, $p = 0.000$, $\eta^2 = 0.311$. The Mauchly's test for sphericity for Confidence toward Inquiry was not violated (W = 0.809, p = 0.205); however this assumption was violated for Intentions toward Inquiry (W = 0.527, $p = 0.008$), so the more conservative Greenhouse-Geisser correction was used. This correction still indicated a significant change in teachers' Intentions toward Inquiry, $p = 0.000$. A post hoc Bonferroni test for Confidence toward Inquiry showed that the mid- and post-assessment differed significantly at $p = 0.006$ as well as the pre- and post-assessment at $p = 0.000$. A post hoc Bonferroni test for Intentions toward Inquiry showed that the pre- and mid-assessment differed significantly at $p = 0.001$, mid- and post-assessment differed significantly at $p = 0.013$, and the pre- and post-assessment differed significantly at $p = 0.000$.

Table 7.
Results for ANOVA (Pre-Mid-Post) – Participating Teachers

Teacher Belief Factors	Pre *M (SE)*	Mid *M (SE)*	Post *M (SE)*	F
Beliefs toward Inquiry	6.24(0.17)	6.25(0.17)	6.49(0.13)	1.48
Confidence toward Inquiry	4.82(0.24)	5.11(0.21)	5.74(0.21)	15.55***
Intentions toward Inquiry	5.00(0.27)	5.97(0.18)	6.32(0.15)	27.994***

*Note: ***p<0.001*
Scales: (Strongly Disagree; Not at all confident; Very unlikely) – 7 (Strongly Agree; Completely confident; Completely likely)

Research Question 4

A Wilcoxon signed-rank test indicated there were no differences for the comparison group of teachers from pre- to post-survey. In fact, there was a declining trend in the mean for Intentions toward Inquiry from the pre- to post-survey for this group of teachers (See Table 8).

ISSN: 2149-8504 (online)

Table 8.
Results of Wilcoxon signed-rank test: Comparison Teachers, Pre – and Post-Survey

Teacher Belief Factors	Pre Survey *M (SE)*	Post Survey *M (SE)*	*p-value*
Beliefs toward Inquiry	6.07(0.22)	6.14(0.22)	0.753
Confidence toward Inquiry	5.34(0.22)	5.41(0.24)	0.779
Intentions toward Inquiry	5.67(0.26)	5.13(0.28)	0.209

Scales: (Strongly Disagree; Not at all confident; Very unlikely) – 7 (Strongly Agree; Completely confident; Completely likely)

Finally, a two-way repeated-measures ANOVA was conducted pre- to post-survey to determine if differences existed for beliefs measured between groups (participant and comparison). Levene's test for equal variance was conducted. This assumption was not violated for any of the variables, allowing us to proceed with the analysis. Results of the ANOVA indicated there was not a significant main effect for Beliefs toward Inquiry, $F(1, 29) = 9.19$, $p = 0.297$, $\eta^2 = 0.012$. However, a significant interaction effect for Confidence toward Inquiry was found, $F(1, 29) = 12.27$, $p = 0.002$, $\eta^2 = 0.072$. A follow up independent samples t-test did not find a significant difference between the post-survey responses between the participating and comparison group of teachers. Finally, a significant interaction effect was found for Intentions toward Inquiry, $F(1, 29) = 20.09$, $p = 0.000$, $\eta^2 = 0.200$. A follow up independent samples t-test found a significant difference between the post-survey responses with the participating teachers reporting a higher mean (M = 6.32) than the comparison group of teachers (M = 5.13). Table 9 provides a summary of results for the independent samples t-tests and Figure 1 provides a visual representation of changes pre- to post-survey for the participating and comparison groups of teachers.

Table 9.
Independent Samples T-test for Confidence and Intentions toward Inquiry, Post-survey between Groups of Teachers

Factor	Part. M(SD)	Comp. M(SD)	t	*p-value*	d
Confidence toward Inquiry	5.74(.86)	5.29(0.94)	1.243	0.224	0.50
Intentions toward Inquiry	6.32(.63)	5.13(1.05)	2.753	0.001**	1.42

Note: **p<0.01
Scales: (Strongly Disagree; Not at all confident; Very unlikely) – 7 (Strongly Agree; Completely confident; Completely likely)

ISSN: 2149-8504 (online)

Figure 2. Change in Beliefs, Confidence, and Intentions toward Inquiry Pre- to Post-Survey for Participating and Comparison Groups of Teachers

Research Question 5

A Wilcoxon signed-rank test was performed to determine if differences existed for the pre and post administration of the EQUIP observational protocol. Only 14 teachers were included in this analysis as teachers who did not participate all four semesters of the program were excluded. Results indicated that Discourse and Assessment were statistically significant and Instruction was marginally significant pre to post. Teachers scored higher on the post-observation for Discourse (Mdn = 2.8) than on the pre-observation (Mdn = 2.2), $p = 0.011$, r = 0.48. Teachers also scored higher on the post-observation for Assessment (Mdn = 2.4) than on the pre-observation (Mdn = 2.0), $p = 0.030$, r = 0.41. Additionally, teachers scored higher on the post-observation for Discourse (Mdn = 2.8) than on the pre-observation (Mdn = 2.2), $p = 0.055$, r = 0.36. These results indicate that teachers were using a slightly higher level of inquiry in their classrooms at the end of the two years as reported by means and levels of significance in Table 10.

Table 10.
Results of Wilcoxon Signed-Rank Test for Participating Teachers' Use of Inquiry

EQUIP Factors	Baseline Fall 2015 *M (SE)*	Post Spring 2017 *M (SE)*	*p-value*
Instructional	2.25(0.21)	2.77(0.25)	0.011*
Discourse	2.09(0.20)	2.66(0.24)	0.030*
Assessment	1.92(0.15)	2.48(0.23)	0.055†
Curriculum	2.13(0.18)	2.81(0.28)	0.118

*Note: †p<0.10 *p<0.05*
Scale: 1 (Pre-Inquiry), 2 (Developing Inquiry), 3 (Proficient Inquiry), 4 (Exemplary inquiry)

A Wilcoxon signed-rank test was also performed to determine if differences existed for the mid2 and post administration of the EQUIP observational protocol. Twenty-seven teachers were included in this analysis. Results indicated that Instruction and Discourse were statistically significant, and Curriculum was marginally significant mid2 to post. Teachers scored higher on the post-observation for Instruction (Mdn = 2.8) than on the mid2-observation (Mdn = 2.4), p = 0.047, r = 0.271. Teachers also scored higher on the post-observation for Discourse (Mdn = 2.2) than on the mid2-observation (Mdn = 2.0), p = 0.050, r = 0.267. Additionally, teachers scored higher on the post-observation for Curriculum (Mdn = 2.25) than on the pre-observation (Mdn = 2.00), p = 0.089, r = 0.232. These are shown in Table 11.

Table 11.
Results of Wilcoxon signed-rank test for Participating Teachers' Use of Inquiry (Year 2 only)

EQUIP Factors	Mid2 Fall 2016 *M (SE)*	Post Spring 2017 *M (SE)*	*p-value*
Instructional	2.33(0.19)	2.75(0.14)	0.047*
Discourse	2.11(0.29)	2.50(0.15)	0.050*
Assessment	2.07(0.18)	2.36(0.14)	0.266
Curriculum	2.07(0.18)	2.40(0.16)	0.089†

*Note: †p<0.10 *p<0.05*
Scale: 1 (Pre-Inquiry), 2 (Developing Inquiry), 3 (Proficient Inquiry), 4 (Exemplary inquiry)

Discussion

Findings from this study provide evidence that the professional development program was effective in initiating the transitioning of teachers' beliefs and classroom practice. As discussed by Desimone (2009), results indicate a connection between professional development, changes in teachers' beliefs, and changes in teachers' practice. In particular, a focus on PD

ISSN: 2149-8504 (online)

elements as discussed in literature (Darling-Hammond et al., 2017) could account for changes noted over the course of the two-year PD.

Teacher Beliefs

Given the connection between individuals' experiences and beliefs (Raths, 2001), exploring teachers' beliefs based on their experiences in the PD was a way we could assess the effectiveness of the PD program. Drawing on prior work pertaining to the variety of beliefs that are relevant to better understand teachers' instructional choices and potential changes in practice (e.g., Canrinus, et al., 2012; Cross, 2009; Hofer & Pintrich, 2004), we measured teacher beliefs through the constructs Teaching Philosophy, Job Satisfaction, and Openness to Change. Results indicated that these beliefs increased over the course of the two-years teachers participated in the PD. Teaching Philosophy in particular indicated the most significant difference with teachers reporting a more student-centered belief toward teaching. This finding may be in response to the PD focusing on inquiry-based practices, which include student-centered methods. Teachers were actively engaged in reflecting on their own practice throughout the program along with experiencing inquiry-based instruction through a variety of PD activities. As Levitt (2002) indicates, unconventional professional development can lead to a transitioning teacher philosophy and that these changes in teachers' beliefs directly align with the focus of the PD teachers are experiencing. Further, Brand and Moore (2011) noted the importance of teachers being "active participants in both goal-setting and ongoing work of the professional development process" (p. 908). Rather than having data collection being a by-product of the PD, teachers became active participants of their growth, setting goals, revising existing lesson plans, and reflecting on their beliefs and practices as it related to their classroom. Teachers' beliefs related to Job Satisfaction and Openness to Change also increased over the course of the PD. Research related to Openness to Change is linked to teachers' use of innovative technologies and risk-taking (Baylor & Ritchie, 2002; Howard & Gigliotti, 2016). Openness to change and a willingness to take risks is important when teachers are testing different strategies and considering changes to their practice. Additionally, there is some evidence that attitudes about inquiry and teachers' openness to ideas and actions may be correlated (Meijer et al., 2016).

Job Satisfaction, related to teacher burnout (Wang et al., 2015) and professional identity (Canrinus et al., 2012), is of particular importance in light of a large percentage of teacher attrition reported for the field (Ingersoll et al., 2018). The evaluation of change between the participating and comparison group of teachers makes clear the importance of effective PD on transitioning teachers' beliefs. Our results indicate participating teachers' beliefs transitioned over the course of the biennium but the comparison group of teachers did not significantly change. In fact, the means of the comparison group were trending down. However, no changes were evident with Professional Commitment. Given that these constructs are closely connected with professional identity, which is a more stable construct (Beijaard, 1995), this finding may not be surprising. However, as noted in the current study, professional identity is not rigid and factors such as a teachers' work environment, sense of agency, and personal and professional experiences could lead to change (Day et al., 2006). Perhaps results from this study make evident Rokeach's (1972) concept of belief systems where peripheral beliefs such as beliefs about inquiry are easier to change than beliefs that are core to a teacher's sense of self, professional identity.

Inquiry-Based Beliefs

It is interesting that teacher beliefs related to the seven inquiry practices did not change over the course of the two-year PD. Teachers' initial responses were generally positive and stayed the same throughout the two years. However, changes became evident when exploring their level of confidence in designing instruction and implementing instruction that is inquiry based with a significant difference pre to post for the participating group of teachers. Changes were more significant for the intent to implement inquiry with statistically significant differences found between the pre- to post-survey and for the participating group of teachers and with results indicating a significant difference in the post-survey means between the participating and comparison group of teachers. It is possible that teachers observing these practices being implemented and practicing implementing the strategies in their own classrooms made them more likely to respond that they intended to implement the inquiry practices. Given that beliefs are theorized and have been shown to be predictive of intentions and intentions is predictive of actions (Aelterman et al., 2016; Ajzen, 1996), the change in belief may indicate that teachers will be more likely to enact the list of inquiry practices. Again, no significant difference was found for the comparison group of teachers and more notable shifts were noted for the comparison group of teachers for decreasing intentions between pre- to post-surveys.

Inquiry-Based Instruction

Results indicate changes in teachers' implementation of inquiry-based practices in their classroom. In particular, significant changes were found with the instructional and discourse factors. The Instructional factor corresponded with teachers' promoting conceptual understanding, student engagement, teachers allowing students to explore content before explaining, teachers acting as a facilitator, and students applying concepts being learned to new situations (depth of knowledge demonstrated). The Discourse factor also increased significantly over the course of the PD. Discourse corresponded with teachers' level of questioning, the complexity of questions, ability to engage students with questions, discourse, investigations or reflection, and how discussions were facilitated (types of support and reasoning evident). Even for experienced teachers, transitioning to use these practices effectively can be very difficult as they require the implementation of rich tasks. In addition, teachers may need to transition their role in the classroom, where they may have previously directed the learning experiences rather

ISSN: 2149-8504 (online)

than facilitated the experiences for students. Assessment was found to be marginally significant and corresponded to teachers assessing and modifying instruction based on prior knowledge, being process focused with learning activities, encouraging reflection and higher-level thinking with students, using authentic assessment measures consistently and effectively, and encouraging explanations with support. Research indicates that PD involving ongoing coaching, with particular focus on including teacher reflection, leads to change in teacher practice (Teemant, 2014). The instructional coach could be one of the features of the PD program that accounts for the changes found in teachers' practice. These coaching sessions in addition to the modeled lessons and instructional strategies during workshop sessions and over the summer provided teachers with a picture of what inquiry looks like in practice. In addition, teachers had personal goals, which related directly to their own classroom practice. Since goals help direct attention and energy, influence strategy selection and planning, and ultimately affect performance (Locke & Latham, 2002), it is possible that these goals helped the teachers grow through the challenge of developing new skills with inquiry-instructional practices. With specific goals that aligned to a rubric describing the target pedagogical outcomes, teachers could attend to an attainable practice, use feedback to monitor their progress towards goal attainment, and plan steps for improving their inquiry-based instructional practices. It is possible that despite only focusing on one goal at a time, teachers were influencing other areas of inquiry since many of the ideas related to one another. For example, meaningful questioning (Discourse) would likely lead to students being more active (Instructional) and soliciting explanations (Assessment). It is also important to note that while teachers use of inquiry increased over the course of the program, many of them were still at the developing inquiry level. The time it takes to transition practice seems to be extensive and may indicate the need for continued PD and support to maintain and further transition practice.

Differences were also noted for the 27 teachers who participated in the second year of the PD. The Assessment factor was not significant for this group, but the Curriculum factor was marginally significant. The Curriculum factor corresponded with the depth of content in the lesson as well as connections to the big picture, flexibility during investigations, the connectedness of the content and the investigations, and students being able to record information in non-prescriptive ways and communicate effectively. It is possible that the focus of the summer institute during the second year might account for this change. In the second year, the PD focused on project-based instruction with teachers designing a unit of instruction aligned to this method. Rather than focusing on individual lessons, a focus on a unit of instruction might have helped teachers see overarching concepts with the content they were teaching and how to make these more evident in their instruction. It is also likely that the job-embedded approach of the PD as recommended in current literature (Shirrell et al., 2019) provided teachers with the support needed to transition their practice.

ISSN: 2149-8504 (online)

Limitations and Conclusions

A limitation for this study was the attrition of teachers, which reduced the sample size. The small sample size increased the potential for Type II error with the ANOVAs conducted. However, it was preferable to keep the significance level alpha at 0.05 rather than potentially conflate results. It is possible that additional significance might have been found with a larger sample size of teachers. In addition to reducing the sample size, this attrition made some of the PD efforts more challenging as the facilitators needed time to build rapport with the teachers and teachers needed time to develop a sense of community with each other.

Meaningful PD can be challenging. Participating teachers in our study were required to be active rather than passive in the process. They were challenged to think deeply about their own practice, create professional goals that would lead to changes in their practice, and modify their instruction based on these goals. The intensity of the PD may be one reason we experienced a lot of transition in the program despite getting very positive feedback from the participating teachers. Being engaged in this type of PD is different than what is traditionally experienced by many teachers and calls for a level of commitment and engagement that may be taxing on their time and mental energy. We do not make this statement to be disparaging toward teachers who may have dropped the PD, but to share a challenge that PD developers and implementers may face when engaging in these activities. Despite challenges faced by teachers and facilitators, PD can result in changes in beliefs and practice as was evident in this study. It is also important to note that changes in teachers' beliefs and practices were potentially informing one another. Teaching using inquiry might have changed teachers' beliefs and self-efficacy and vice versa as literature indicates that the relationship between teacher beliefs and practice is "reciprocal, but complex" (Levin, 2015 p. 70). The job-embedded approach and coaching aspects of the PD could also inform educators and administrators as they consider future PD activities. A way that these aspects of PD could be included is through instructional coaches. While requiring a significant investment by a school or district, instructional coaches could provide the type of PD that builds on successful elements included as part of the current study as well as providing the on-going, context-specific experiences for teachers. It is also important to have the support from administrators, without including an evaluative component. This may include peer mentor programs, which would include many elements of the PD reported on in the current study, such as being job-embedded, collaboration with other teachers, and instructional support.

McGee and colleagues (2013) noted that future research regarding PD needed to examine "how teachers are able to translate their new learning into classroom experiences for their students" (p. 25). This study is illustrative of how PD can affect changes in beliefs and practices, which could ultimately influence student learning.

ISSN: 2149-8504 (online)

References

Aelterman, N., Vansteenkiste, M., Van Keer, H., & Haerens, L. (2016). Changing teachers' beliefs regarding autonomy support and structure: The role of experienced psychological need satisfaction in teacher training. *Psychology of Sport and Exercise*, *23*, 64-72. https://doi.org/10.1016/j.psychsport.2015.10.007

Ajzen, I. (1996). The directive influence of attitudes on behavior. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior*: 385-403. Guilford Press.

Allen, J. P., Hafen, C. A., Gregory, A. C., Mikami, A. Y., & Pianta, R. (2015). Enhancing secondary school instruction and student achievement: Replication and extension of the My Teaching Partner-Secondary intervention. *Journal of Research on Educational Effectiveness*, *8*(4), 475-489. https://doi.org/10.1080/19345747.2015.1017680

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, *333*(6045), 1034-1037. https://doi.org/10.1126/science.1207998

Baylor, A. L., & Ritchie, D. (2002). What factors facilitate teacher skill, teacher morale, and perceived student learning in technology-using classrooms?. *Computers & Education*, *39*(4), 395-414. https://doi.org/10.1016/S0360-1315(02)00075-1

Beijaard, D. (1995) Teachers' prior experiences and actual perceptions of professional identity, *Teachers and Teaching*, *1*(2), 281–294. https://doi.org/10.1080/1354060950010209

Brand, B. R., & Moore, S. J. (2011). Enhancing teachers' application of inquiry-based strategies using a constructivist sociocultural professional development model. *International Journal of Science Education*, *33*(7), 889-913. https://doi.org/10.1080/09500691003739374

Bransford, J. D., Brown, A. L. & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school*. National Academies Press.

Brown, S. W., Lawless, K. A., & Boyer, M. A. (2013). Promoting positive academic dispositions using a web-based PBL environment: The GlobalEd 2 project. *Interdisciplinary Journal of Problem-Based Learning*, *7*(1), 67–90. https://doi.org/10.7771/1541-5015.1389

Buczynski, S., & Hansen, C. B. (2010). Impact of professional development on teacher practice: Uncovering connections. *Teaching and Teacher Education*, *26*(3), 599-607. http://dx.doi.org/10.1016/j.tate.2009.09.006

Bybee, R. W., Taylor, J. A., Gardner, A., Van Scotter, P., Powell, J. C., Westbrook, A., & Landes, N. (2006). *The BSCS 5E instructional model: Origins and effectiveness.* Colorado Springs, Co: BSCS.http://www.fremonths.org/ourpages/auto/2008/5/11/1210522036057/bscs5efullreport2006.pdf

Cain, J. S. (2002). An evaluation of the connected mathematics project. *Journal of Educational Research*, *95*, 224–233. https://doi.org/10.1080/00220670209596595

Canrinus, E. T., Helms-Lorenz, M., Beijaard, D., Buitink, J., & Hofman, A. (2012). Self-efficacy, job satisfaction, motivation and commitment: Exploring the relationships between indicators of teachers' professional identity. *European Journal of Psychology of Education*, *27*(1), 115-132. https://doi.org/10.1007/s10212-011-0069-2

Carney, M. B., Brendefur, J. L., Thiede, K., Hughes, G., & Sutton, J. (2016). Statewide mathematics professional development: Teacher knowledge, self-efficacy, and beliefs. *Educational Policy*, *30*(4), 539-572. https://doi.org/10.1177/0895904814550075

Chen, H. T., Wang, H. H., Lin, H. S., P. Lawrenz, F., & Hong, Z. R. (2014). Longitudinal study of an after-school, inquiry-based science intervention on low-achieving children's affective perceptions of learning science. *International Journal of Science Education*, *36*(13), 2133-2156. https://doi.org/10.1080/09500693.2014.910630

Cichon, D., & Ellis, J. (2003). The effects of MATH connections on student achievement, confidence, and perception. In S. L. Senk, & D. R. Thompson (Eds.), *Standards-based school mathematics curricula. What are they? What do students learn?* (pp. 345–374). Lawrence Erlbaum.

Cross, D. I. (2009). Alignment, cohesion, and change: Examining mathematics teachers' belief structures and their influence on instructional practices. *Journal of Mathematics Teacher Education*, *12*(5), 325-346. https://doi.org/10.1007/s10857-009-9120-5

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.

Day, C., Kington, A., Stobart, G., & Sammons, P. (2006). The personal and professional selves of teachers: Stable and unstable identities. *British Educational Research Journal*, *32*(4), 601-616. https://doi.org/10.1080/01411920600775316

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*, 181-199. https://doi.org/10.3102/0013189X08331140

Doppelt, Y., Schunn, C. D., Silk, E. M., Mehalik, M. M., Reynolds, B., & Ward, E. (2009). Evaluating the impact of a facilitated learning community approach to professional development on teacher practice and student achievement. *Research in Science & Technological Education*, 27(3), 339-354. https://doi.org/10.1080/02635140903166026

Fielding-Wells, J., O'Brien, M., & Makar, K. (2017). Using expectancy-value theory to explore aspects of motivation and engagement in inquiry-based learning in primary mathematics. *Mathematics Education Research Journal*, *29*(2), 237-254. https://doi.org/10.1007/s13394-017-0201-y

Fischer, C., Fishman, B., Dede, C., Eisenkraft, A., Frumin, K., Foster, B, … McCoy, A. (2018). Investigating relationships between school context, teacher professional development, teaching practices, and student achievement in response to a nationwide science reform. *Teaching and Teacher Education*, *72*, 107-121. https://doi.org/10.1016/j.tate.2018.02.011

Forbes, C.T., & Zint, M. (2011). Elementary teachers' beliefs about, perceived competencies for, and reported use of scientific inquiry to promote student learning about and for the environment. *The Journal of Environmental Education*, *42*, 30-42. https://doi.org/10.1080/00958961003674673

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, *82*, 300–329. https://doi.org/10.3102/0034654312457206

Gallagher, H. A., Arshan, N., & Woodworth, K. (2017). Impact of the National Writing Project's College-Ready Writers Program in high-need rural districts. *Journal of Research on Educational Effectiveness*, *10*(3), 570-595. https://doi.org/10.1080/19345747.2017.1300361

Gallagher, H. A., Woodworth, K., Park, C., & McCaffrey, T. (2014). *Evaluating the impact of professional development to meet challenging writing standards in high-need elementary schools.* SRI International.

Granger, E. M., Bevis, T. H., Saka, Y., Southerland, S. A., Sampson, V., & Tate, R. L. (2012). The efficacy of student-centered instruction in supporting science learning. *Science*, *338*(6103), 105-108. http://doi.org/10.1126/science.1223709

Greenleaf, C. L., Litman, C., Hanson, T. L., Rosen, R., Boscardin, C. K., Herman, J., Schneider, S. A., Madden, S, & Jones, B. (2011). Integrating literacy and science in biology: Teaching and learning impacts of reading apprenticeship professional development. *American Educational Research Journal*, *48*(3), 647-717. https://doi.org/10.3102/0002831210384839

Hofer, B. K., & Pintrich, P. R. (Eds.). (2004). *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Psychology Press.

Howard, S. K., & Gigliotti, A. (2016). Having a go: Looking at teachers' experience of risk-taking in technology integration. *Education and Information Technologies*, *21*(5), 1351-1366.

Ingersoll, R. M., Merrill, E., Stuckey, D., & Collins, G. (2018). *Seven trends: The transformation of the teaching force–updated October 2018.* Consortium for Policy Research in Education.

Jiang, F., & McComas, W. F. (2015). The effects of inquiry teaching on student science achievement and attitudes: Evidence from propensity score analysis of PISA data. *International Journal of Science Education*, *37*, 554–576. https://doi.org/10.1080/09500693.2014.1000426

Johnson, C. C., & Fargo, J. D. (2014). A study of the impact of transformative professional development on Hispanic student performance on state mandated assessments of science in elementary school. *Journal of Science Teacher Education*, *25*(7), 845-859. https://doi.org/10.1007/s10972-014-9396-x

Kang, J., & Keinonen, T. (2018). The effect of student-centered approaches on students' interest and achievement in science: Relevant topic-based, open and guided inquiry-based, and discussion-based approaches. *Research in Science Education*, *48*(4), 865-885. https://doi.org/10.1007/s11165-016-9590-2

Kim, H. (2016). Inquiry-based science and technology enrichment program for middle school-aged female students. *Journal of Science Education and Technology*, *25*(2), 174-186. https://doi.org/10.1007/s10956-015-9584-2

Kleickmann, T., Tröbst, S., Jonen, A., Vehmeyer, J., & Möller, K. (2016). The effects of expert scaffolding in elementary science professional development on teachers' beliefs and motivations, instructional practices, and student achievement. *Journal of Educational Psychology*, *108*(1), 21-42. https://doi.org/10.1037/edu0000041

Knowles, M. (1990). *The adult learner: A neglected species.* Gulf Publishing.

ISSN: 2149-8504 (online)

Koksal, E. A., & Berberoglu, G. (2014). The effect of guided-inquiry instruction on 6th grade Turkish students' achievement, science process skills, and attitudes toward science. *International Journal of Science Education*, *36*(1), 66-78. https://doi.org/10.1080/09500693.2012.721942

Kutaka, T. S., Smith, W. M., Albano, A. D., Edwards, C. P., Ren, L., Beattie, H. L., Lewis, J., Heaton, R. M., & Stroup, W. W. (2017). Connecting teacher professional development and student mathematics achievement: A 4-year study of an elementary mathematics specialist program. *Journal of Teacher Education*, *68*(2), 140-154. https://doi.org/10.1177/0022487116687551

Landry, S. H., Anthony, J. L., Swank, P. R., & Monseque-Bailey, P. (2009). Effectiveness of comprehensive professional development for teachers of at-risk preschoolers. *Journal of Educational Psychology*, *101*(2), 448-465. https://doi.org/10.1037/a0013842

Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research*, *86*(3), 681-718. https://doi.org/10.3102/0034654315627366

Levitt, K. E. (2002). An analysis of elementary teachers' beliefs regarding the teaching and learning of science. *Science Education*, *86*(1), 1-22. https://doi.org/10.1002/sce.1042

Lloyd, G. (2002). Mathematics teachers' beliefs and experiences with innovative curriculum materials. In G. Leder, E. Pehkonen, & G. To¨rner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 149–159). Kluwer.

Lock, E.A., & Latham, G.P. (2002). Building a practically useful theory of goal setting and task motivation. *American Psychologist, 57*, 705-717. https://doi.org/10.1037/0003-066X.57.9.705

Lortie, D. (1975). *Schoolteacher: A Sociological Study.* University of Chicago Press

Mac Iver, M. A., & Mac Iver, D. J. (2009). Urban middle-grade student mathematics achievement growth under comprehensive school reform. *Journal of Educational Research*, *102*, 223–236. https://doi.org/10.3200/JOER.102.3.223-236

Marshall, J. C. (2013). *Succeeding with inquiry in science and math classrooms*. ASCD & NSTA.

Marshall, J.C. & Horton, R. M. (2009). *Developing, assessing, and sustaining inquiry-based instruction: A guide for math and science teachers and leaders*. VDM Publishing House Ltd.

Marshall, J. C., & Horton, R. M. (2011). The relationship of teacher-facilitated, inquiry-based instruction to student higher-order thinking. *School Science and Mathematics*, *111*(3), 93-101. https://doi.org/10.1111/j.1949-8594.2010.00066.x

Marshall, J. C., Horton, B., Smart, J., & Llewellyn, D. (2008). EQUIP: Electronic Quality of Inquiry Protocol: Retrieved from Clemson University's Inquiry in Motion Institute.

McGee, J. R., Wang, C., & Polly, D. (2013). Guiding teachers in the use of a standards-based mathematics curriculum: Teacher perceptions and subsequent instructional practices after an intensive professional development program. *School Science and Mathematics*, *113*(1), 16-28. https://doi.org/10.1111/j.1949-8594.2012.00172.x

Meijer, M. J., Geijsel, F., Kuijpers, M., Boei, F., & Vrieling, E. (2016). Exploring teachers' inquiry-based attitude. *Teaching in Higher Education*, *21*(1), 64-78. https://doi.org/10.1080/13562517.2015.1115970

Meissel, K., Parr, J. M., & Timperley, H. S. (2016). Can professional development of teachers reduce disparity in student achievement?. *Teaching and Teacher Education*, *58*, 163-173. https://doi.org/10.1016/j.tate.2016.05.013

Meyer, J. P., Allen, N. J., & Smith, C. A. (1993). Commitment to organizations and occupations: Extension and test of a three-component conceptualization. *Journal of Applied Psychology*, *78*(4), 538-551. https://doi.org/10.1037/0021-9010.78.4.538

Meyers, C. V., Molefe, A., Brandt, W. C., Zhu, B., & Dhillon, S. (2016). Impact results of the eMINTS professional development validation study. *Educational Evaluation and Policy Analysis*, *38*(3), 455-476. https://doi.org/10.3102/0162373716638446

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *47*(4), 474-496. https://doi.org/10.1002/tea.20347

Moyer-Packenham, P. S., Bolyard, J. J., Oh, H., & Cerar, N. I. (2011). Common features of professional development activities for mathematics and science teachers. *Professional development in education*, *37*(4), 571-589. https://doi.org/10.1080/19415257.2010.531597

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

National Research Council. (1996). *National science education standards*. Washington, DC: National Academies Press.

National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. National Academies Press.

National Research Council. 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. The National Academies Press.

NCTM. (1989). *Curriculum and evaluation standards for school mathematics*. NCTM.

NCTM. (2000). *Principles and standards for school mathematics*. NCTM.

NCTM. (2014). *Principles to actions: Ensuring mathematical success for all*. NCTM.

NGSS Lead States (2013). *Next Generation Science Standards: For states, by states*. National Academies Press.

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

Polly, D., McGee, J., Wang, C., Martin, C., Lambert, R., & Pugalee, D. K. (2015). Linking professional development, teacher outcomes, and student achievement: The case of a learner-centered mathematics program for elementary school teachers. *International Journal of Educational Research*, *72*, 26-37. https://doi.org/10.1016/j.ijer.2015.04.002

Raths, J. (2001). Teachers' beliefs and teaching beliefs. *Early Childhood Research and Practice*, *3*(1), 385-393.

Riegle-Crumb, C., Morton, K., Nguyen, U., & Dasgupta, N. (2019). Inquiry-based instruction in science and mathematics in middle school classrooms: Examining its association with students' attitudes by gender and race/ethnicity. *AERA Open*, *5*(3), 2332858419867653.

ISSN: 2149-8504 (online)

Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. *Handbook of Research on Teacher Education*, *2*(102-119), 273-290.

Rokeach, M. (1972). *Beliefs, attitudes and values.* Jossey-Bass Inc.

Shirrell, M., Hopkins, M., & Spillane, J. P. (2019). Educational infrastructure, professional learning, and changes in teachers' instructional practices and beliefs. *Professional Development in Education*, *45*(4), 599-613. https://doi.org/10.1080/19415257.2018.1452784

Starr, S., Haley, H. L., Mazor, K. M., Ferguson, W., Philbin, M., & Quirk, M. (2006). Initial testing of an instrument to measure teacher identity in physicians. *Teaching and Learning in Medicine*, *18*(2), 117-125. https://doi.org/10.1207/s15328015tlm1802_5

Stearns, E., Banerjee, N., Mickelson, R., & Moller, S. (2014). Collective pedagogical teacher culture, teacher–student ethno-racial mismatch, and teacher job satisfaction. *Social Science Research*, *45*, 56-72. https://doi.org/10.1016/j.ssresearch.2013.12.011

Tam, A. C. F. (2012). Teachers' misconceptions and questionable practices when using Putonghua as the medium-of-instruction: A case study of Hong Kong. *Teachers and Teaching*, *18*(6), 655-673. https://doi.org/10.1080/13540602.2012.746500

Tam, A. C. F. (2015). The role of a professional learning community in teacher change: A perspective from beliefs and practices. *Teachers and Teaching*, *21*(1), 22-43. https://doi.org/10.1080/13540602.2014.928122

Taylor, J. A., Roth, K., Wilson, C., Stuhlsatz, M, & Tipton, E. (2017). The Effect of an analysis-of-practice, videocase-based, teacher professional development program on elementary students' science achievement. *Journal of Research on Educational Effectiveness, 10*(2), 1-31. https://doi.org/10.1080/19345747.2016.1147628

Teemant, A. (2014). A mixed-methods investigation of instructional coaching for teachers of diverse learners. *Urban Education*, *49*(5), 574-604. https://doi.org/10.1177/0042085913481362

Timperley, H., & Alton-Lee, A. (2008). Reframing teacher professional learning: An alternative policy approach to strengthening valued outcomes for diverse learners. *Review of research in education*, *32*(1), 328-369. https://doi.org/10.3102/0091732X07308968

Trotter, Y. D. (2006). Adult learning theories: Impacting professional development programs. *Delta Kappa Gamma Bulletin*, *72*(2) 8-13.

Vannatta, R. A., & Nancy, F. (2004). Teacher dispositions as predictors of classroom technology use. *Journal of Research on Technology in Education*, *36*(3), 253-271.

Wang, H., Hall, N. C., & Rahimi, S. (2015). Self-efficacy and causal attributions in teachers: Effects on burnout, job satisfaction, illness, and quitting intentions. *Teaching and Teacher Education*, *47*, 120-130. https://doi.org/10.1016/j.tate.2014.12.005

Woolfolk Hoy, A. W., Hoy, W. K., & Davis, H. A. (2009). *Teachers' self-efficacy beliefs*. In K. Wentzel, & A. Wigfield (Eds.), Handbook of motivation in school (pp. 627-655). Lawrence Erlbaum.

Yow, J. A., & Lotter, C. (2016). Teacher learning in a mathematics and science inquiry professional development program: first steps in emergent teacher leadership. *Professional Development in Education*, *42*(2), 325-351. https://doi.org/10.1080/19415257.2014.960593

ISSN: 2149-8504 (online)

Journal of Research in STEM Education
Vol 8. No 2. December 2022. 61-78
https://doi.org/10.51355/jstem.2022.116

RESEARCH REPORT

Pathways to Teacher STEM Certification in Texas: A Case for Addressing the Minoritized Teacher Shortage

Shetay Ashford-Hanserd[1], Omar S. Lopez, Catherine Cherrstrom, Brett L. Lee

Texas State University, USA

Abstract: *For the United States to remain globally competitive, policymakers, researchers, and administrators emphasize the need for a highly skilled and diverse STEM workforce. As such, the US education system must recruit high-quality, diverse, STEM-certified teachers to improve STEM learning outcomes and career pathways for all students, including historically underrepresented minority students. Increased recruitment and retention of minoritized STEM teachers through alternative certification pathways will dissipate the shortage of qualified STEM teachers. The purpose of this quantitative study was to examine trends in STEM teacher certification by race or ethnicity to address minoritized teacher shortages in Texas, the second largest education authority in the US. The study analyzed 67,629 teacher certification records from the Texas Education Agency's State Board for Educator Certification. Results revealed disparities in Race or Ethnicity among STEM teachers that could dispel the teacher shortage gap if parity were achieved among White, Hispanic, and Black STEM teachers.*

Keywords: *STEM teacher certification, alternative certification programs, teacher shortage, minoritized teacher shortage, teacher certification*

1 Corresponding Author: Pedernales Building, #112 601 University Drive San Marcos, Texas 78666 Email: sashford@txstate.edu

To cite this article: Ashford-Hanserd, S., Lopez, O. S., Cherrstrom, C., & Lee, B. L. (2022). Pathways to teacher STEM certification in Texas: A case for addressing the minority teacher shortage. *Journal of Research in STEM Education, 8*(2), 61-78. https://doi.org/10.51355/jstem.2022.116

ISSN: 2149-8504 (online)

Amid globalization challenges, declines in H1-B visa allocations, rapid increases in technological advances, and demographical shifts, the United States (US) faces severe shortages of qualified science, technology, engineering, and mathematics (STEM) workers (Carnevale et al., 2011; Vincent & Velkoff, 2010). National and state policymakers, researchers, and administrators emphasize the need for a highly skilled and diverse STEM workforce for the US to remain globally competitive and to meet increasing STEM knowledge and skills gaps (Cherrstrom et al., 2022). To meet US STEM workforce demands, and to improve STEM learning outcomes and career pathways for all students, particularly historically underrepresented minority students, the US K-12 education system must address its shortage of high quality, diverse, STEM-certified teachers (Mentzer et al., 2019).

Mainstream news consistently reports on the national teacher shortage (Castro et al., 2018; Cowan et al., 2016; Dee & Goldhaber, 2017; Sutcher et al., 2016, 2019; Wiggan et al., 2020). In 2018, U.S. public schools reached a record high enrollment of 50.7 million students (National Center for Education Statistics, 2018b), warranting 1.5 million new teachers (Wiggan et al., 2020). Since 1985, the number of teachers produced by teacher preparation programs has grown, while the students-to-teacher ratio in public schools has increased (Cowan et al., 2016). Consequently, the national teacher shortage has resulted from staffing challenges in critical subjects (e.g., STEM, special education, bilingual, English language learners), particularly at low-income or high-poverty, high-minority, rural, and urban schools (Cowan et al., 2016; Dee & Goldhaber, 2017; Sutcher et al., 2016, 2019).

Research has revealed a minority or urban teacher shortage represented by the longstanding gap between minority students' percentage concerning minority teachers (Ingersoll et al., 2019; Sutcher et al., 2019). On average, high-poverty and urban schools are “3 to 10 times more likely to have teachers who are uncertified, not fully prepared, or teaching outside their field of preparation than students in predominantly White and more affluent schools” (Castro et al., 2018, p. 2). While efforts to recruit minority teachers have been successful, minority teachers have higher attrition rates than teachers at affluent White schools due to poor working conditions (Ingersoll et al., 2019; Sutcher et al., 2016, 2019; Wiggan et al., 2020). Furthermore, recruiting and retaining minority teachers is vital for equitable access to quality STEM education for all students and to meet our nation's increasing demand for STEM knowledge and skills (Ingersoll et al., 2019). Ultimately, the low number of STEM professionals in secondary STEM classrooms impacts the quality of teaching and student learning as well as workforce development (Mentzer et al., 2019). A diverse, high-quality, STEM-certified teacher workforce will improve STEM learning outcomes and career pathways for underrepresented minority students (Mentzer et al., 2019).

Alternative STEM certification pathways offer a viable solution to close the gap of qualified teachers across the nation. Increased recruitment and retention of minority teachers through alternative STEM certification pathways will dissipate the shortage of qualified STEM teachers. Although researchers have indicated traditional pathways produce higher student outcomes (Anderson, 2006; Marder et al., 2020; Ruiz de Castilla, 2018; Schmidt et al., 2020), most states have responded to teacher shortages in various subject areas with alternative certification programs (Humphrey et al., 2008; Mentzer et al., 2019). Such programs more effectively alleviate the shortage of qualified minority STEM teachers and improve STEM learning outcomes for all students (Chamberlin-Kim et al., 2019). Since higher percentages of underrepresented minorities pursue alternative pathways over traditional certification (National Center for Education Statistics, 2018a), we posit alternative certification programs provide the most viable pathway to address the minority teacher shortage.

The purpose of this study was to examine trends in STEM teacher certification by Race or Ethnicity and gender to address minority teacher shortages in Texas. Our research question guiding the study was: What can certification data tells us about the efficacy of different pathways toward preparing teachers for STEM certification? The insights to this question gained from the current study has implications for teachers, teacher preparation programs, schools and districts, and ultimately students and our nation. This article continues with a review of the literature on teacher certification in Texas, followed by methods used in the study, results of the study, and discussion of results with implications.

Teacher Certification in Texas

Texas offers three common pathways for initial teacher certification: traditional certification, alternative certification program (ACP), and post-baccalaureate (post-bac) certification (Texas Education Agency, 2020b). For subsequent certifications in STEM and other subject areas, teachers pursue an additional certification-by-examination (cert-by-exam) pathway (Texas Education Agency, 2020a).

In the US and Texas, traditional or standard certification (also referred to as *traditional undergraduate program*) is the most prevalent pathway to teacher certification (Texas Education Agency, 2020b). Along this pathway, teacher certification is intertwined with requirements for obtaining a bachelor's degree at a four-year college or university. Most states refer to the initial teacher certification as *traditiona*l or *standard* or certification in the literature. More states have also implemented pre-collegiate urban teaching academies (e.g., Fletcher Jr. & Ashford, 2016) as magnet high school programs to encourage high school students to pursue traditional teaching pathways.

While most public school teachers are the product of traditional four-year degree programs, alternative certification (ACP) is a second pathway accounting for a growing percentage of educators. The ACP is a byproduct of Ronald Reagan's National Commission on Excellence in Education report, *A Nation at Risk*, which identified serious deficits in teacher preparation programs. According to LoCascio et al. (2016), the purpose of ACPs is "to certify candidates who have strong academic area content knowledge but have limited or no background in formal teacher preparation" (p. 105). The benefits of alternative certification include fewer course requirements and decreased program completion time for the prospective teacher. Required to hold a bachelor's degree for program admission, students are promoted to field-based teaching experiences sooner than undergraduate students enrolled in a traditional certification program. Developed in response to teacher shortages in various subject areas, ACPs are now common in most states (Bowling & Ball, 2018; Humphrey et al., 2008; LoCascio et al., 2016).

The post-baccalaureate (post-bac) certification, a third pathway, is considered an advanced degree in obtaining a teaching certificate in most states. The definition of post-bac varies across state lines and among institutions of higher education. In the state of Texas, post-bac is defined as a certification pathway best suited for those seeking an advanced (master's or doctorate) degree with certification (Texas Education Agency, 2020b). According to Ruiz de Castilla (2018), the post-bac pathway may not require completing an advanced degree while completing the certification requirements. The University of North Texas (UNT) College of Education (2020) outlines the post-bac certification program "for graduates who did not pursue education studies but want to become a teacher, principal, or superintendent. . . . Courses from each of these programs may be applied to the master's degree" (para. 1). However, "courses taken for the post-baccalaureate program do not lead to a degree. Instead, completion of the program qualifies students to take the required state certification examinations." While the terms *post-baccalaureate* and *alternative certification* are interchangeably used, a key differentiator is cost. Programs offering a certificate in conjunction with an advanced degree typically have higher cost and longer duration.

In Texas, teachers with a valid Texas classroom teaching certificate and a baccalaureate degree may obtain additional certifications through additional certification-by-exam (cert-by-exam) (Texas Education Agency, 2020a) also referred to as *endorsement by exam* (Hollo et al., 2019) and *additional exam* (Ruiz de Castilla, 2018). While advantageous for teachers desiring to expand expertise, cert-by-exam is not an option for initial certification, teachers who teach students with visual impairments, or non-classroom teaching purposes (e.g., superintendent, principal, school counselor). Cert-by-exam offers a common pathway for teachers to obtain multiple certifications. Other states have varying processes for teachers to obtain additional certification. An online search revealed the terms *additional certification by examination* or *cert-by-exam* unique to Texas. As one example, the California Commission on Teacher Credentialing defines the initial teacher certification as a *preliminary credential*, and the subsequent credential as a *clear credential* (State of California Commission on Teacher Credentialing, 2019).

Texas, the second-largest education authority in the US, offers traditional certification, ACP, and post-bac pathways for initial teacher certification and cert-by-exam pathway for subsequent certifications in STEM and other subject areas (Texas Education Agency, 2020a, 2020b). Since higher percentages of underrepresented minorities pursue alternative pathways over traditional certification pathways (National Center for Education Statistics, 2018a), we posit alternative certification programs provide the most viable pathway to address the minority teacher shortage. In the next section, we discuss the methods used to examine trends in STEM teacher certification by race or ethnicity, and gender to address minority teacher shortages in Texas.

Methods

Population

Based on teacher data for the 2018-2019 school year, this study's population consisted of 67,629 STEM teachers, representing 16.9% of the 399,670 educators teaching in Texas public schools. Within this statewide teacher population, 59.4% (n=237,467) were White, followed by 27.2% (n=108,569) Hispanic, and 10.4% (n=41,470) Black. The remaining 3.0% (n=12,164) consisted of the Other race or ethnicity category. Among White teachers, 19.0% (n=45,148) were STEM certified compared to Hispanic and Black teachers, respectively 13.0% (n=14,088) and 12.9% (n=5,328). The Other race or ethnicity group had the highest percentage of STEM teachers, 25.2% (n=3,065). By gender, there were four times more female teachers (n=265,292) than male teachers (n=66,749)2018-2019. Similarly, there were two times more female STEM teachers (n=45,001) than male STEM teachers (n=22,628). However, among only 14.5% of female teachers, were STEM certified compared to 25.3% of male teachers

Data

This study's primary data came from the teacher certification records maintained by the State Board for Educator Certification (SBEC) at the Texas Education Agency (TEA). In 1995, the Texas Legislature created SBEC to oversee all aspects of the preparation, certification, and standards of conduct of public school educators (Texas Education Code, 2020). The data contained multiple records per teacher, where each record represented the teacher's certification to teach a subject (e.g., biology, physics, chemistry) within a broader field (e.g., science). In this study, a STEM certification included any subject in computer science, mathematics, or science, or the subjects of technology applications or technology education in the field of technology in vocational education programs.

Each certification was associated with a *standard* or *provisional* type. Standard certifications must be renewed every five years if issued on or after September 1, 1999, while provisional certifications are for life if issued before September 1, 1999 (Texas Education Code, 2020). Both certification types had a start date and an expiration date for standard certifications. Moreover, all records had one of four certification pathways: (standard) traditional program, alternative certification program (ACP), post-baccalaureate (post-bac), or additional certification-by-examination (cert-by-exam). Teacher demographic data included race or ethnicity (White, Hispanic, Black, or Other) and gender (F or M).

Measures

The longitudinal nature of the data allowed us to build a certification profile for each teacher consisting of the initial and subsequent certifications in STEM fields. In so doing, we created numerical measures for the study as follows. Using the certification's start issuance date, for each teacher, we chronologically sorted records to identify the initial and subsequent certifications. We then kept the initial and the next encountered certification if in a STEM field. For each teacher, we calculated the number of years between a certification's start date and 2019, the referenced school year for data collection, and the number of years between the start dates of the initial and next encountered certification in a STEM field. We then compared these measures against our independent calculations from the raw data for teachers randomly selected so that, we could ensure our SAS code was reliable and valid in producing the measures for the current study. For analytical procedures, we relied on descriptive statistics to illuminate patterns and discoveries in the data in a meaningful way within the context of our research question.

Theoretical Framework

Since intersectionality (Crenshaw, 1991) is an appropriate theoretical framework to describe multifaceted racial and social identities of diverse individuals, we referenced its use in the context of education related to teacher professionals (Macias & Stephens, 2017). During this study, we utilized intersectionality theory to describe race and gender factors related the race or ethnicity, and gender status of diverse STEM teachers in the state of Texas.

Results

Race or Ethnicity of STEM Teachers

Table 1 shows the race or ethnicity, and gender distribution of STEM teachers in Texas. The results indicated a disparity in race or ethnicity among STEM teachers. This finding's relevance can be best realized relative to student demographics, as shown in the table. Of the 5,431,910 students in Texas public schools in the 2018-2019 school year (TEA, 2019), race or ethnicity distribution consisted of 27.4% (n=1,490,299) White, 52.6% (n=2,854,590) Hispanic, and 12.6% (n=685,775) Black students. In comparison, STEM teacher demographics for the 67,629 educators consisted of 66.8% (n=45,148) White, 20.8% (n=14,088) Hispanic, and 7.0% (n=5,328) Black teachers.

ISSN: 2149-8504 (online)

Table 1.
STEM Teachers in Texas: Race or Ethnicity, and Gender

	Students (S)		STEM Teachers (T)		S:T Ratio	Parity	
Group	**n**	**Percent**	**n**	**Percent**	**(S_n÷T_n)**	**Required**	**Shortage**
Race or Ethnicity							
White	1,490,299	27.4	45,148	66.8	33.0	45,148	0
Hispanic	2,854,590	52.6	14,088	20.8	202.6	86,479	72,391
Black	685,775	12.6	5,328	7.9	128.7	20,775	15,447
Other	401,246	7.4	3,065	4.5	130.9	12,156	9,091
Total	**5,431,910**	**100**	**67,629**	**100**	**80.3**	**164,557**	**96,928**
Gender							
Female	2,647,524	48.7	45,001	66.5	58.8	45,001	0
Male	2,784,386	51.3	22,628	33.5	123.1	47,327	24,699
Total	**5,431,910**	**100.0**	**67,629**	**100.0**	**80.3**	**92,328**	**24,699**

The students-to-teacher ratio by race or ethnicity indicated 33.0 White students to every 1.0 White STEM teacher. In comparison, there were 202.6 Hispanic students for every Hispanic teacher and 128.7 Black students for every Black teacher. To achieve parity with White STEM teachers would require 86,479 Hispanic STEM teachers, 72,391 more than in the 2018-2019 school year. Similarly, parity with White STEM teachers would require 20,775 Black STEM teachers, 15,447 more than in the 2018-2019 school year. If parity were achieved among the three major race or ethnicity groups, the result would add 96,928 STEM teachers to the 67,629 in the 2018-2019 school year, for a total of 164,557 STEM teachers available to students in public schools.

Disparities also existed by STEM teacher gender. Of the 5,431,910 students in Texas public schools in the 2018-2019 school year (TEA, 2019), the gender demographics consisted of 48.7% (n=2,647,524) female and 51.3% (n=2,784,386) male. In comparison, STEM teacher demographics for the 67,629 educators consisted of 66.5% (n=45,001) female and 33.5% (22,628) male. The ratio of students-to-teacher by gender indicated 58.8 female students for every female STEM teacher. In comparison, there were 123.1 male students for every male STEM teacher. To achieve parity with female STEM teachers would require 47,327 male STEM teachers, 24,699 more than in the 2018-2019 school year, for a total of 92,328 STEM teachers available to students in public schools.

STEM Certified Teachers by Initial and Second Certification Field

Table 2 shows the distribution of STEM certified teachers by initial and next encountered certification in a STEM field and descriptive statistics for the number of years between a certification's start date and 2019. The data include two groups of teachers: teachers with initial certification in a non-STEM field who later earned a certificate in a STEM field, and those with initial certification in a STEM field who went on to earn a second certificate in another STEM field. The 25,357 teachers with initial certification in non-STEM fields were certified, on average, 19.0 years (SD=10.4) before the 2018-2019 school year. For the first certification in a STEM field, 54.0% (n=13,692) of teachers did so in mathematics, followed by 27.7% (n=7,018) in science. The average years from their STEM certification to the 2018-2019 school year was 15.8 years (SD=10.7), which implies their STEM certifications were earned, on average, within four years after receiving their initial non-STEM certification.

In comparison, the 42,272 teachers with initial certification in STEM fields were certified on average, 13.4 years (SD=9.7) before the 2018-2019 school year. Of these STEM teachers, however, only 18.4% (n=7,787) went on to earn a second certification in another STEM field, primarily in the fields of science (n=4,999) followed by mathematics (n=1,463). For the second-STEM certified teachers, the average years from initial STEM certification and second STEM certification to the 2018-2019 school year were 18.1 years (SD=10.8) and 14.7 years (SD=11.2), respectively. Therefore, second-STEM certified teachers earned their second STEM certifications, on average, within four years after receiving their initial STEM certification.

Table 2.
STEM Certified Teachers in Texas: Initial and Second Certification Field

			Years to 2018-2019 School Year				
Field	**n**	**Percent**	**Mean**	**SD**	**Min**	**Median**	**Max**
Teachers with Initial Certification in a non-STEM Field							
Bilingual Education	1,920	7.6	9.3	6.3	0	8	32
English Language Arts	2,304	9.1	21.3	9.5	1	22	54
Fine Arts	743	2.9	17.4	10.7	1	15	47
Foreign Language	205	0.8	17.8	11.1	1	15	47
General Elementary	11,590	45.7	19.7	9.8	0	20	54
Health & PE	3,962	15.6	20.0	11.0	1	18	57
Social Studies	1,261	5.0	21.4	11.3	1	21	55
Special Education	1,072	4.2	15.5	7.1	0	15	46
Vocational Education	2,300	9.1	20.4	11.3	1	19	58
Total	**25,357**	**100**	**19.0**	**10.4**	**0**	**18**	**58**
First Certification in a STEM Field							
Computer Science	649	2.6	17.4	8.5	1	18	44
Mathematics	13,692	54.0	18.0	11.0	0	18	57
Science	7,018	27.7	15.7	10.6	0	14	52
Technology	3,998	15.8	7.9	4.9	0	7	18

Total	**25,357**	**100.0**	**15.8**	**10.7**	**0**	**14**	**57**
Teachers with Initial Certification in a STEM Field							
Computer Science	635	1.5	20.5	9.4	1	20	45
Mathematics	21,164	50.1	12.3	8.4	0	11	58
Science	19,113	45.2	14.8	10.7	0	13	60
Technology	1,360	3.2	6.6	4.6	0	5	17
Total	**42,272**	**100.0**	**13.4**	**9.7**	**0**	**12**	**60**
Second Certification in a STEM Field							
Computer Science	509	1.2	9.5	8.9	1	5	34
Mathematics	1,463	3.5	16.4	12.1	0	13	57
Science	4,999	11.8	16.0	11.2	0	13	60
Technology	816	1.9	6.9	4.6	0	6	18
Total	**7,787**	**18.4**	**14.7**	**11.2**	**0**	**12**	**60**
(Initial Certification)			18.1	10.8	0	16	60
No Second Certification	34,485	81.6	12.3	9.1	0	11	58

STEM Certified Teachers' Initial and Second Certification Pathway by Race or Ethnicity

Table 3 shows the distribution of STEM teachers' initial and second certification pathways by race or ethnicity. While 54.1% (n=36,567) of STEM teachers earned their initial certification through a Traditional program, this was not always the case by race or ethnicity. Among White and Hispanic STEM teachers, 59.4% (n=26,824) and 49.3% (n=6,947) respectively completed a Traditional program to earn the initial certification. In comparison, 55.6% (n=2,961) of Black STEM teachers completed an Alternative Certification Program (ACP) to earn their initial certification.

Table 3.
STEM Certified Teachers in Texas: Initial and Second Certification Pathway by Race or Ethnicity

	Race or Ethnicity				
Certification Pathways	**White**	**Hispanic**	**Black**	**Other**	**Total**
Initial Certification					
Traditional	26,824	6,947	1,662	1,134	36,567
	59.4	49.3	31.2	37.0	54.1
	73.4	19.0	4.6	3.1	100
ACP	13,782	6,050	2,961	1,513	24,306
	30.5	42.9	55.6	49.4	35.9
	56.7	24.9	12.2	6.2	100.0
Post-Bac	4,542	1,091	705	418	6,756
	10.1	7.7	13.2	13.6	10.0

	67.2	16.2	10.4	6.2	100.0
Second Certification					
Traditional	9,634	1,574	525	277	12,010
	21.3	11.2	9.9	9.0	17.8
	80.2	13.1	4.4	2.3	100.0
ACP	1,495	449	183	139	2,266
	3.3	3.2	3.4	4.5	3.4
	66.0	19.8	8.1	6.1	100.0
Cert-by-Exam	13,343	3,206	958	794	18,301
	29.6	22.8	18.0	25.9	27.1
	72.9	17.5	5.2	4.3	100.0
Post-Bac	424	63	62	18	567
	0.9	0.5	1.2	0.6	0.8
	74.8	11.1	10.9	3.2	100.0
Total Second Certification	24,896	5,292	1,728	1,228	33,144
	55.1	37.6	32.4	40.1	49.0
	75.1	16.0	5.2	3.7	100
No Second Certification	20,252	8,796	3,600	1,837	34,485
	44.9	62.4	67.6	59.9	51.0
	58.7	25.5	10.4	5.3	100.0
Total	45,148	14,088	5,328	3,065	67,629
	100.0	100.0	100.0	100.0	100.0
	66.8	20.8	7.9	4.5	100.0

Note. Entries: n, column%, row%.

Among the 67,629 STEM teachers, slightly less than 50% (n=33,144) completed a second certification in a STEM field. Of the four certification pathways to the second certification, the Cert-by-Exam was most prevalent for each race or ethnicity group. The least prevalent pathway was post-bac, with only 567 teachers consisting of about 75% (n=424) White, followed by 11.1% (n=63) Hispanic, and 10.9% (n=62) Black.

Of the remaining 34,485 STEM teachers not earning a second certification in another STEM field, Whites represented 58.7% (n=20,252), followed by Hispanics and Blacks with 25.5% (n=8,796) and 10.4% (n=3,600), respectively. However, the data paints a different picture when viewed by the race or ethnicity group. Among Blacks, 67.6% (n=3,600) did not earn a second certification in another STEM field, followed closely next by Hispanics with 62.4% (n=8,796). In comparison, 44.9% (n=20,252) of the White STEM teachers did not seek a second certification in another STEM field.

Initial and Second Certification Pathways for STEM Teachers

Table 4 shows the distribution of STEM teachers by initial and second certification pathways and descriptive statistics for the number of years between the start dates of the initial and next encountered certification in a STEM field, if applicable. The findings illustrated the varied pathway combinations from the initial to the second certification. The Cert-by-Exam was the most prevalent pathway to the second certification in another STEM field. This was true, regardless of whether the teachers' initial certification was in a non-STEM or STEM field.

Note the certification pathways where the initial and second certification were both post-bac. The average time interval from the initial to the second certification for teachers was zero years (SD=0). This implies the post-bac pathway provides teachers with the shortest time interval to prepare for initial certification and then add a second certification in a STEM field.

Table 4.
STEM Certified Teachers in Texas: Initial and Second Certification Pathways for Non-STEM and STEM Initial Certifications

Certification and Pathway		**Distribution**		**Years to Second Certification**				
Initial	**Second**	**n**	**Percent**	**Mean**	**SD**	**Min**	**Median**	**Max**
Teachers with Initial Certification in non-STEM field								
Traditional	Traditional	9,874	98.1	0.7	3.1	0	0.0	42
ACP		118	1.2	6.2	4.1	0	6.0	21
Post-Bac		73	0.7	3.3	3.7	0	2.0	15
Traditional	ACP	220	12.0	11.3	7.1	0	10.0	32
ACP		1,556	84.8	0.1	0.8	0	0.0	11
Post-Bac		59	3.2	4.2	3.0	0	4.0	11
Traditional	Cert-by-Exam	6,542	50.2	6.8	7.6	0	4.0	46
ACP		5,098	39.1	3.8	4.0	0	3.0	32
Post-Bac		1,385	10.6	4.6	4.9	0	3.0	23
Traditional	Post-Bac	19	4.4	0.8	3.2	0	0.0	14
Post-Bac		413	95.6	0.0	0.0	0	0.0	0
Teachers with Initial Certification in STEM Field								
Traditional	Traditional	1,846	94.9	1.0	3.2	0	0.0	40
ACP		54	2.8	5.4	3.5	0	5.0	15
Post-Bac		45	2.3	4.6	3.9	0	4.0	15

ISSN: 2149-8504 (online)

Traditional	ACP	32	7.4	10.0	7.6	0	9.5	34
ACP		386	89.6	0.1	0.9	0	0.0	11
Post-Bac		13	3.0	3.2	2.0	1	4.0	7
Traditional	Cert-by-Exam	2,045	38.8	6.1	6.8	0	4.0	47
ACP		2,250	42.7	3.3	3.7	0	2.0	24
Post-Bac		981	18.6	4.1	4.6	0	2.0	24
Traditional	Post-Bac	1	0.7	2.0		2	2.0	2
Post-Bac		134	99.3	0.0	0.0	0	0.0	0
Traditional	(No Second	15,988	46.4					
ACP	Certification)	14,844	43.0					
Post-Bac		3,653	10.6					

Note. Empty cells denote not applicable

Second Certification Pathway for STEM Teachers

Table 5 shows the distributions for the second certification pathways and, for each pathway, descriptive statistics for the number of years between the start dates of the initial and next encountered certification in a STEM field, if applicable. The table separately presents the results for teachers with initial certifications in non-STEM and STEM fields, followed by combined STEM teachers regardless of when they first certified in a STEM field. As shown earlier, the Cert-by-Exam (n=18,301) was the most prevalent second certification pathway, however, these results clearly indicated Cert-by-Exam also had the most considerable time interval of, on average, 5.1 years (SD=6.1) between initial and second certification in a STEM field. The least taken pathway to the second certification was Post Bac (n=567) but with the shortest time interval of, on average, zero years (SD=0.8) between initial and a second certification in a STEM field. Again, this implies the post-bac provides teachers with the shortest time interval to prepare for initial certification and then add a second certification in a STEM field.

Table 5.

STEM Certified Teachers in Texas: Second Certification Pathway for Non-STEM and STEM Initial Certifications

	Distribution		**Years to Second Certification**				
Second Certification Pathway	**n**	**Percent**	**Mean**	**SD**	**Min**	**Median**	**Max**
Teachers with Initial Certification in the non-STEM field							
Traditional	10,065	39.7	0.8	3.1	0	0	42
ACP	1,835	7.2	1.6	4.5	0	0	32
Cert-by-Exam	13,025	51.4	5.4	6.3	0	3	46
Post-Bac	432	1.7	0.0	0.7	0	0	14
Total	**25,357**	**100.0**	**3.2**	**5.6**	**0**	**0**	**46**

ISSN: 2149-8504 (online)

Teachers with Initial Certification in STEM Field							
Traditional	1,945	4.6	1.2	3.3	0	0	40
ACP	431	1.0	0.9	3.4	0	0	34
Cert-by-Exam	5,276	12.5	4.5	5.4	0	3	47
Post-Bac	135	0.3	0.0	0.2	0	0	2
Total	**7,787**	**18.4**	**3.4**	**5.1**	**0**	**1**	**47**
No Second Certification	34,485	81.6	n/a	n/a	n/a	n/a	n/a
All Teachers Certified in a STEM Field, Initial and/or Second Certification							
Traditional	12,010	17.8	0.9	3.2	0	0	42
ACP	2,266	3.4	1.5	4.3	0	0	34
Cert-by-Exam	18,301	27.1	5.1	6.1	0	3	47
Post-Bac	567	0.8	0.0	0.6	0	0	14
Total	**33,144**	**49.0**	**3.3**	**5.4**	**0**	**1**	**47**
No Second Certification	34,485	51.0	n/a	n/a	n/a	n/a	n/a

Discussion

Texas is the second-largest education authority in the US, educating 5.4 million students, nearly 10% of all students in the US, and employing 67,629 STEM teachers, nearly 17% of the 399,670 educators teaching in Texas public schools. Therefore, understanding most viable certification pathways to address the minority STEM teacher shortage in Texas also informs the broader minority STEM teacher shortage in the US (Schmidt et al., 2020).

Research has revealed a minority teacher shortage represented by the longstanding gap between the percentage of minority students in relation to minority teachers (Ingersoll et al., 2019; Sutcher et al., 2019). Based on this study's results, we believe Texas could dispel the minority teacher shortage gap if parity were achieved among the three major race or ethnicity groups: White, Black, and Hispanic. To achieve a comparable to the 33.0 students-to-teacher ratios for Hispanics, Texas would require 86,479 Hispanic STEM teachers to achieve parity with White STEM teachers in Texas. Similarly, Texas would need an additional 20,775 Black STEM teachers to achieve parity of Black STEM teachers in Texas. In total, Texas would need to add 96,928 STEM teachers to the current baseline of 67,629 to achieve parity among the three major ethnic groups in the 2018-2019 school year.

In the US and Texas, traditional or standard certification (also referred to as *traditional undergraduate program*) is the most prevalent pathway to teacher certification (Texas Education Agency, 2020b). Likewise, this study found 54.1% (n=36,567) of all STEM teachers in Texas earned initial certification through a Traditional program. By Race or Ethnicity, we discovered fewer Hispanic STEM teachers 49.3% (n=6,947) than White STEM teachers 59.4% (n=26,824) earned their initial certification through a traditional pathway. This is important with regards to teacher quality in the classroom. Studies situated in Texas illuminate traditional certification pathway and results. Schmidt et al. (2020) conducted a randomized survey with 2,134 newly certified teachers from traditional and ACPs and concluded teachers who receive traditional certifications have higher content preparation than alternatively certified teachers. However, certified teachers reported higher mathematics scores on college entrance examinations in grades 4-8. Anderson (2020) used a sample data set of 2,599 novice teachers (i.e., five years or less experience) matched to student achievement and found students perform better on high school math assessments when taught by a traditionally certified teacher. Marder et al. (2020) utilized multilevel models to examine student test score changes from the 2010-2011 to 2017-2018 school years, nested within classrooms, teachers, and school campuses. They found students enrolled in Algebra I classes taught by experienced teachers certified through traditional pathways achieved .03 to .05 gains (in standard deviation units) compared to students taught by alternatively certified teachers.

However, the study also found that 55.6% (n=2,961) of Black STEM teachers completed an ACP to earn their initial certification. Furthermore, more Blacks (67.6%, n=3,600) and Hispanics (62.4%, n=8,796) did not earn a second certification in another STEM field in comparison to 44.9% (n=20,252) of White STEM teachers. Therefore, findings support the notion that higher percentages of Black (13% vs. 5%), Hispanic (15% vs. 8%) and Multi-Race (32% vs. 22%) teachers pursue alternative pathways over traditional pathways (National Center for Education Statistics, 2018a).

This finding was not so surprising. Research has shown that many minority teachers who participate in ACP programs also grew up in large population urban areas compared to their White peers (Zeichner & Schulte, 2001). The benefits of alternative certification include fewer course requirements and decreased program completion time for the prospective teacher. Required to hold a bachelor's degree for program admission, ACP participating teachers are promoted to field-based teaching experiences sooner than undergraduate students enrolled in a traditional certification program.

When viewed from a macro-level, alternative licensure programs compensate for teachers' shortages in certain subjects by qualifying more people to teach in high-needs areas that have difficulty attracting and retaining teachers. Historically, alternative programs have attracted a more diverse teacher candidate population than traditional programs (US Department of Education, Office of Planning, Evaluation and Policy Development, 2016). According to Kelly and Northrup (2015), the shortage of educators in STEM fields has fueled the popularity and growth of ACPs. LoCascio et al. (2016) conducted a study in low income urban areas in northeastern New Jersey and found 45% of novice teachers participating in an ACP identified as minority. Teachers who use an ACP, however, may feel less prepared. Kee (2012) found first-year teachers with limited education coursework and field experiences felt less prepared than teachers with more complete pedagogical preparation.

ISSN: 2149-8504 (online)

Implications and Future Research

This study offers implication for theory, practice, and policy. For theory, the study further illuminates the need for high-quality, diverse, STEM-certified teachers to improve STEM learning outcomes and career pathways for all students, including significantly underrepresented minority students. The results expand the current body of knowledge about certification pathways for minority STEM teachers and identify a viable option to contribute to the broader picture of STEM teacher certification and solve STEM teacher shortages.

For practice, this study informs future teachers, current teachers, teacher preparation programs, and schools and districts. Future teachers and current teachers, as applicable, can discern and find encouragement in multiple pathways to initial and secondary certifications, including the transition from non-STEM initial certification to STEM secondary certification. Furthermore, they might choose to purse ACP as the most viable option. Teacher preparation programs can educate students about pathways, and schools and districts can promote such pathways. Policy implications at the state or national levels include identifying ACP as a viable option and supporting the expansion of this solution.

The study informs future research to address limitations and further examine STEM minority teacher. This study solely focused on Texas, did not examine student outcomes, and did not account for socioeconomic settings. We recommend replicating this study for a variety of states to determine national relevance and application in solving STEM teacher shortages. Future studies could examine the relationship between performance of minority STEM teachers who pursued ACP pathways with performance of minority students. In addition to teachers preparation and pathways, such studies might examine high-poverty, middle-class, and affluent settings to isolate structural barriers impacting student performance.

This study examined trends in STEM teacher certification by Race or Ethnicity and gender to address minority teacher shortages in Texas, the second largest education authority in the US. Results revealed disparities in Race or Ethnicity among STEM teachers that could dispel the teacher shortage gap if parity were achieved among White, Hispanic, and Black STEM teachers and identified alternative certification programs as the most viable pathway. Increased recruitment and retention of minority STEM teachers through this pathway will dissipate the shortage of STEM teachers. High-quality, diverse, STEM-certified teachers will improve student learning outcomes and career pathways for all students, including significantly underrepresented minority students, leading to a highly skilled and diverse STEM workforce and global competitiveness as a nation.

ISSN: 2149-8504 (online)

Availability of Data and Materials

Teacher certification data available from the Public Education Information Management System (PEIMS) through a public information request via the Texas Education Agency available at https://tea.texas.gov/about-tea/contact-us/public-information-requests.

References

Anderson, J. C., II. (2006). Insights for recruiting underrepresented individuals into careers in agriculture, food, and natural resources. *Agricultural Education Magazine, 78*(5), 11–13. https://www.naae.org/profdevelopment/magazine/archive_issues/Volume78/v78i5.pdf

Anderson, K. A. (2020). A national study of the differential impact of novice teacher certification on teacher traits and race-based mathematics achievement. *Journal of Teacher Education, 71*(2), 247–260. https://doi.org/10.1177/0022487119849564

Bowling, A. M., & Ball, A. L. (2018). Alternative certification: A solution or an alternative problem? *Journal of Agricultural Education, 59*(2), 109–122. http://dx.doi.org/10.5032/jae.2018.02109

Carnevale, A. P., Smith, N., & Melton, M. (2011). STEM. Georgetown University Center on Education and the Workforce. https://1gyhoq479ufd3yna29x7ubjn-wpengine.netdna-ssl.com/wp-content/uploads/2014/11/stem-complete.pdf

Castro, A., Quinn, D. J., Fuller, E., & Barnes, M. (2018). *Addressing the importance and scale of the US teacher shortage*. University Council for Educational Administration. https://files.eric.ed.gov/fulltext/ED579971.pdf

Chamberlin-Kim, J., Tarnay, J., & Wells, J. C. (2019). Alternative teacher preparation programs: examination through a marketing lens. *Rural Special Education Quarterly, 38*(3), 137–150. https://doi.org/10.1177/8756870519860737

Cherrstrom, C. A, López, O. S., & Ashford-Hanserd, S. (2022). STEM knowledge in non-STEM U.S. occupations. *Community College Journal of Research and Practice, 46*(7), 457-471.

Cowan, J., Goldhaber, D., Hayes, K., & Theobald, R. (2016). Missing elements in the discussion of teacher shortages. *Educational Researcher, 45*(8), 460–462. https://doi.org/10.3102/0013189X16679145

Crenshaw, K. (1991). Mapping the margins: Identity politics, intersectionality, and violence against women. *Stanford Law Review, 43*(6), 1241-1299.

Dee, T. S., & Goldhaber, D. (2017). *Understanding and addressing teacher shortages in the United States* (Policy Proposal 2017-05). The Hamilton Project. https://www.hamiltonproject.org/assets/files/understanding_and_addressing_teacher_shortages_in_us_pp.pdf

Fletcher Jr., E. C., & Ashford, S. (2016). Unintended consequences of a state accountability policy in a pre-collegiate urban teaching academy program. *The Journal of Research in Business Education, 57*(2), 26–41. https://www.questia.com/read/1P3-4311456961/unintended-consequences-of-a-state-accountability

Hollo, A., Floyd, K. K., & Brigandi, C. B. (2019). Endorsement by exam: Policy and practice in rural special education. *Rural Special Education Quarterly, 38*(3), 177–184. https://doi.org/10.1177/8756870519843500

Humphrey, D. C., Wechsler, M. E., & Hough, H. J. (2008). Characteristics of effective alternative teacher certification programs. *Teachers College Record, 110*(1), 1–63. https://www.tcrecord.org/content.asp?contentid=12613

Ingersoll, R., May, H., & Collins, G. (2019). Recruitment, employment, retention and the minority teacher shortage. *Education Policy Analysis Archives, 27*(37), 2-38. http://dx.doi.org/10.14507/epaa.27.3714

Kee, A. N. (2012). Feelings of preparedness among alternatively certified teachers: What is the role of program features? *Journal of Teacher Education, 63*(1), 23–38. https://doi.org/10.1177/0022487111421933

Kelly, S., & Northrop, L. (2015). Early career outcomes for the "best and the brightest": Selectivity, satisfaction, and attrition in the beginning teacher longitudinal survey. *American Educational Research Journal, 52*(4), 624-656. https://dx.doi.org/10.3102/0002831215587352

LoCascio, S. J., Smeaton, P. S., & Waters, F. H. (2016). How induction programs affect the decision of alternate route urban teachers to remain teaching. *Education and Urban Society, 48*(2), 103–125. https://doi.org/10.1177/0013124513514772

Macias, A., & Stephens, S. (2019). Intersectionality in the field of education: A critical look at race, gender, treatment, pay, and leadership. *Journal of Latinos and Education*, *18*(2), 164-170. https://doi.org/10.1080/15348431.2017.1383912

Marder, M., David, B., & Hamrock, C. (2020). Math and science outcomes for students of teachers from standard and alternative pathways in Texas. *Education Policy Analysis Archives, 28*(27), 1-39. https://doi.org/10.14507/epaa.28.4863

Mentzer, G. A., Czerniak, C. M., & Duckett, T. R. (2019). Comparison of two alternative approaches to quality STEM teacher preparation: Fast-track licensure and embedded residency programs. *School Science & Mathematics, 119*(1), 35–48. https://doi.org/10.1111/ssm.12314

National Center for Education Statistics. (2018a). *Characteristics of public school teachers who completed alternative route to certification programs*. https://nces.ed.gov/programs/coe/pdf/coe_tlc.pdf

National Center for Education Statistics. (2018b). *Fast facts: Back to school statistics*. https://nces.ed.gov/fastfacts/display.asp?id=372

Ruiz de Castilla, V. (2018). *Teacher certification and academic growth among English learner students in the Houston Independent School District*. (REL 2018-284). https://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2018284.pdf

Schmidt, W. H., Burroughs, N. A., Houang, R. T., & Cogan, L. S. (2020). The role of content knowledge in mathematics teacher preparation: A study of traditional and alternative teacher preparation in Texas. *Journal of Teacher Education, 71*(2), 233–246. https://doi.org/10.1177/0022487118805989

ISSN: 2149-8504 (online)

State of California Commission on Teacher Credentialing. (2019). *Multiple subject teaching credential. Requirements for teachers prepared in California*. Author. https://www.ctc.ca.gov/docs/default-source/leaflets/cl561c.pdf

Sutcher, L., Darling-Hammond, L., & Carver-Thomas. (2016). A coming crisis in teaching? Teacher supply, demand, and shortages in the US. Learning Policy Institute. https://learningpolicyinstitute.org/sites/default/files/product-files/A_Coming_Crisis_in_Teaching_REPORT.pdf

Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2019). Understanding teacher shortages: An analysis of teacher supply and demand in the united states. *Education Policy Analysis Archives, 27*(35), 1-36. https://doi.org/10.14507/epaa.27.3696

Texas Education Agency. (2020a). *Additional certification by examination*. https://tea.texas.gov/texas-educators/certification/additional-certifications/additional-certification-by-exam-information

Texas Education Agency. (2020b). *Which educator preparation program (EPP) should I choose?* https://helpdesk.tea.texas.gov/hc/en-us/articles/360035205713-Which-Educator-Preparation-Program-EPP-should-I-choose-

Texas Education Code. (2020). 19 Tex. Admin Code §235.61. http://ritter.tea.state.tx.us/sbecrules/tac/chapter235/ch235d.html

US Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. (2016). *The state of racial diversity in the educator workforce*. http://www2.ed.gov/rschstat/eval/highered/racial-diversity/state-racial-diversity-workforce.pdf

University of North Texas College of Education. (2020). *Post-baccalaureate teacher certification*. https://coe.unt.edu/teacher-education-and-administration/certification

Vincent, G. K., & Velkoff, V. A. (2010). *The next four decades: The older population in the United States: 2010 to 2050* (No. 1138; Issue 1138). US Department of Commerce, Economics and Statistics Administration. https://www.census.gov/prod/2010pubs/p25-1138.pdf

Wiggan, G., Smith, D., & Watson-Vandiver, M. J. (2020). The national teacher shortage, urban education and the cognitive sociology of labor. Advance online publication. Urban Review, 1. https://doi.org/10.1007/s11256-020-00565-z

Zeichner, K. M., & Schulte, A. K. (2001). What we know and don't know from peer-reviewed research about alternative teacher certification programs. *Journal of Teacher Education, 52*(4), 266–282. https://doi.org/10.1177/0022487101052004002

ISSN: 2149-8504 (online)

Journal of Research in STEM Education
Vol 8. No 2. December 2022. 79-110
https://doi.org/10.51355/jstem.2022.125

RESEARCH REPORT

Categorizing Classroom-based Argumentation in Elementary STEM Lessons: Applying Walton's Types of Argument Dialogue

Jonathan K. Foster[a1], Joanna G. Schneider[b], Lorraine Franco[b], Yuling Zhuang[c], Barbara A. Crawford[b], AnnaMarie Conner[b]

[a]*University of Virginia, USA;* [b]*University of Georgia, USA;* [c]*Emporia State University, USA*

Abstract:
Argumentation is a practice that spans STEM disciplines and is an explicit goal for K12 students in reform-based standards documents. The purpose of this study was to investigate the applicability of Douglas Walton's theoretical model for describing the types of argument dialogue encountered in elementary classrooms focused on learning concepts in science, mathematics, and computer coding. We examined two elementary teachers' STEM classrooms to explore the types of argument dialogue that were evident. We found evidence of six types of dialogues: persuasion, negotiation, information-seeking, deliberation, inquiry, and discovery based on Walton's model. Our findings demonstrate the applicability of Walton's types of argument dialogue to argumentation in elementary STEM contexts. Even though our work takes place in the United States with teachers of children in grades 3-5 (ages 8-10 years), we believe our approach is applicable to other dialogues found in K12 STEM education. We postulate that students having opportunities to engage in arguments with a diverse range of goals (e.g., to prove a hypothesis, to persuade, or to exchange information) is important for their development in learning how to argue in STEM.

Keywords: *Collective argumentation, Walton's dialogue types, STEM, elementary education, Toulmin diagrams*

1 Corresponding Author: 405 Emmet St S, Charlottesville, VA, USA, Email: fosterjk@virginia.edu

To cite this article: Foster, J. K., Schneider, J. G., Franco, L., Zhuang, Y., Crawford, B. A., & Conner, A. (2022). Categorizing classroom-based argumentation in elementary STEM lessons: Applying Walton's types of argument dialogue. *Journal of Research in STEM Education, 8*(2), 79-110. https://doi.org/10.51355/jstem.2022.125

ISSN: 2149-8504 (online)

Argumentation is an explicit goal for K12 students in curricular recommendations in STEM education (K–12 Computer Science Framework [CSF] Steering Committee, 2016; National Governors Association [NGA] Center for Best Practices & Council of Chief State School [CCSS] Officers, 2010; Next Generation Science Standards [NGSS] Lead States, 2013). For example, in science, students may argue by using evidence to explain various phenomena. In engineering, students may argue to determine the best solution to a design challenge. In mathematics, students may argue to critique the reasoning of others or to establish a result; and in computer science, students may argue to describe and justify their computational solutions. These examples illustrate different goals for argumentation within curriculum recommendations for STEM education.

Argumentation is also a professional practice that cuts across the STEM disciplines. STEM professionals will need to be able to proficiently craft multiple arguments with distinct goals. We highlight how some STEM professions may engage in arguments with different goals. For example, scientists may need to persuade others in the scientific community that a new methodology is more reliable and valid in comparison to an accepted standard. Other times scientists may need to provide argument-driven informative arguments to the public as in the case of climate change and the COVID-19 pandemic. Mathematicians work together to prove or disprove mathematical conjectures, such as was the case when Andrew Wiles and other mathematicians jointly constructed mathematical proofs that ultimately lead to confirming Fermat's Last Theorem. These brief examples illustrate just some of the various goals for argument dialogues: to resolve methodological issues within a scientific community, to exchange information with the public, or to prove a mathematical conjecture.

Douglas Walton (1998, 2010) theorized how different goals for argument dialogues shape the nature of argumentative discourse. We reason that Walton's types of argument dialogue are informative for STEM education, especially for understanding students' opportunities to learn how to argue across the STEM disciplines. Our purpose in this study was to operationalize Walton's theoretical model for analyzing the types of argument dialogue in which elementary teachers and students engage when learning STEM content such as mathematics, science, and coding with robotics.

Our study differs from most previous studies using Walton's dialogue theory because we considered both the content of the argument components and their relationship (i.e., the structure of the arguments). Previous studies privileged the content of the argumentative discourse in relation to the types of argument dialogue. For example, Rapanta and Christodoulou (2022) applied Walton's dialogue theory by examining the content of whole-class discussions led by the teacher in secondary science and social science classrooms. We built on their study by applying Walton's dialogue theory to data gathered in elementary classrooms that considered the content and structure of the arguments. Furthermore, we were more inclusive in our data analysis parameters by including whole-class and small-group discussions with and without the teacher's participation. The significance of our study includes the development of a method for investigating the types of argument dialogue found in natural and authentic settings of elementary STEM classrooms. This study may be useful to other STEM education researchers interested in supporting arguments of different kinds as students engage in learning STEM concepts through argumentation and learning how to argue in STEM.

Theoretical Framework and Related Literature

Collective Argumentation

Argumentation is the process by which individuals construct and critique arguments; an argument is the product of argumentation (Nussbaum, 2011). In this study, we use the term collective argumentation when teachers or students make claims and provide evidence and reasoning to support them in a social setting (Conner et al., 2014; see also Forman et al., 1998; Krummheuer, 1995; Whitenack & Knipping, 2002; Yackel, 2002). Other researchers have similarly described our sense of collective argumentation as collaborative argumentation (Nussbaum, 2008), critical discussions (Keefer et al, 2000), accountable talk (Michaels et al., 2008), and exploratory talk (Mercer, 2000). Collective argumentation and these other similar classroom-based discussions have been found to promote students' conceptual understanding of content (Kim & Hand, 2015; Nussbaum, 2008; Walshaw & Anthony, 2008; Webb et al., 2019).

Toulmin's Model for Argumentation

Toulmin's (1958/2003) model for argumentation has been a prominent theoretical framework for education researchers studying the content and structure of argumentation (Nussbaum, 2011). Many mathematics and science education scholars have applied Toulmin's model to analyze argumentation practices in classrooms (e.g., Cross, 2009; Erduran et al., 2004; Krummheuer, 1995; Osborne et al., 2004; Yackel, 2002; Zohar & Nemet, 2002). We follow Krummheuer's (1995) adaptation of Toulmin's model and the work of other mathematics education researchers building on Krummheuer's work (e.g., Forman et al., 1998; Knipping, 2008; Rasmussen et al., 2015; Yackel, 2002). The core structure of Toulmin's model includes a statement that is being established (i.e., claim) with evidence to support the statement (i.e., data), and reasoning justifying the relation of the evidence for supporting the statement (i.e., warrant). Other argument components in Toulmin's model include statements describing the circumstances or conditions under which the warrant may not be valid or applicable (i.e., rebuttals), statements about the authority of the warrant (i.e., backings), and statements about the certainty of the claim (i.e., modal qualifiers). These argument components or statements are determined from the interactions of the collective and are not predefined by logic or the contents of the statement (Krummheuer, 1995; Yackel, 2001).

ISSN: 2149-8504 (online)

In this study, extended Toulmin diagrams (See Figure 1; Conner, 2008) framed our understanding of the content and structure of arguments and the participation of the teacher and students in the process of collective argumentation. We used extended Toulmin diagrams for the following reasons. First, extended Toulmin diagrams allow us to signify who contributed which argument components (claim, data, warrant, etc.) through line style and color. Red solid lines denote teacher contributions, blue dotted lines denote student contributions, and purple dot-dash lines denote joint teacher-student contributions. If a warrant is not explicitly stated by a teacher or her students, then it is inferred from context and stated within a cloud. In fact, Toulmin (1958/2003) asserted that warrants and backings are often left implicit unless there is an explicit need for clarity. For example, a mathematics classroom community over time may develop normative ways of reasoning and members may not provide explicit reasoning (i.e., warrants) for well-established claims or backings for the permissibility of their warrants (Rasmussen et al., 2015). Second, extended Toulmin diagrams allow us to signify teacher or student actions that either directly prompt or respond to argument components. For example, if a teacher restates a student's claim, a solid red line oval with the teacher's restatement would be placed on the student's dashed-line blue claim box. Furthermore, the extended Toulmin model allows us to capture an argument component that performs more than one function, such as a component that functions as a claim in the beginning of an argument and then as data for a subsequent claim (Conner, 2008). To attend to these components with more than one function, an extended Toulmin diagram places the component within one box and labels all functions (e.g., Data/Claim; see Figure 1). The extended Toulmin model follows Whitenack and Knipping's (2002) distinction between Toulmin's warrants and backings[2].

[2] We note that backings are almost always implicit within our conceptualization of collective argumentation because they indicate the relevance of the warrant in the field in which the argument is situated (see also Toulmin, 1956/2003, p. 95-98).

ISSN: 2149-8504 (online)

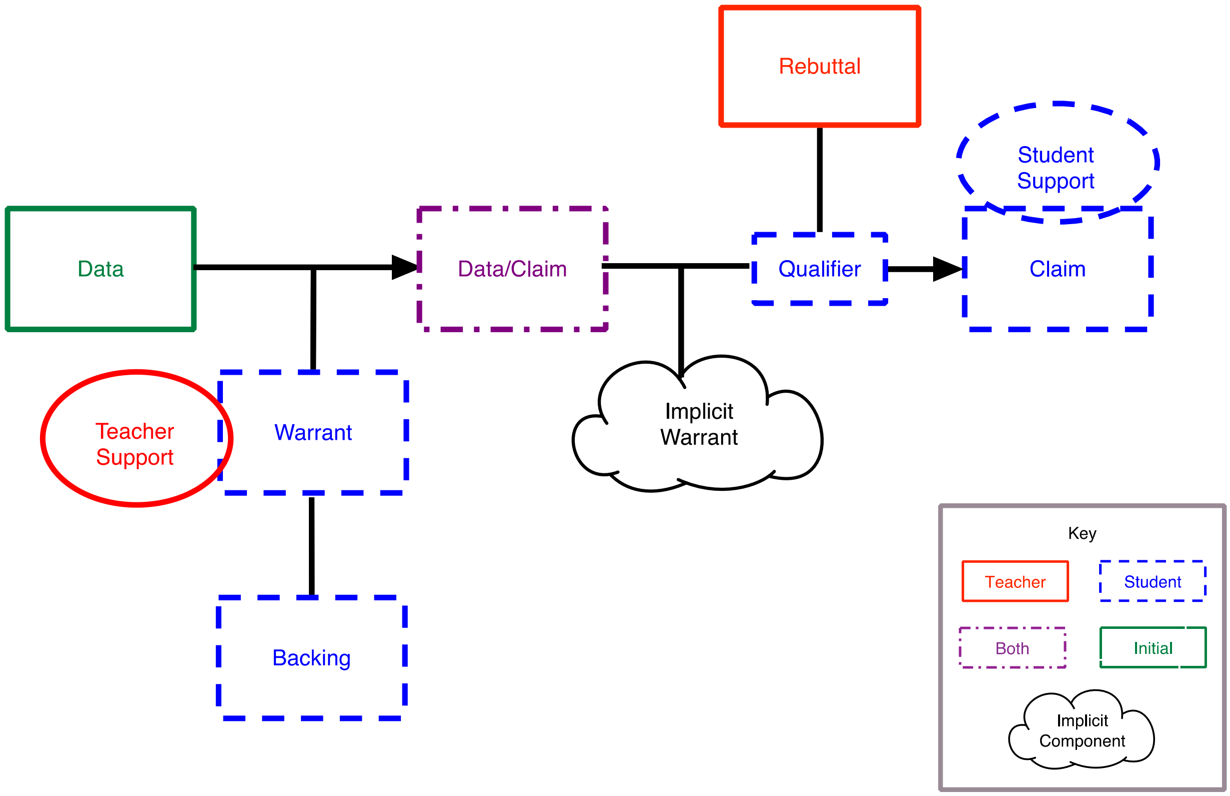

Figure 1. Extended Toulmin (1958/2003) Diagram

Note. Adapted from "Expanded Toulmin diagrams: A tool for investigating complex activity in classrooms" by Conner, 2008, *Proceedings of the Joint Meeting of the International Group for the Psychology of Mathematics Education 32 and the North American Chapter of the International Group for the Psychology of Mathematics Education XXX*, p. 361-368.

Walton's Dialogue Theory Model: Types of Argument Dialogue

Walton (1998) defined dialogue as a "framework in which two (or more) parties reason together with each other by verbal exchanges in order to fulfil a conventionalized goal" (p.6), a definition we find consistent with our conception of collective argumentation. Walton (2022; 1998) theorized seven types of argument dialogue, which differ in terms of the initial situation from which the argument arose, the goals of each participant, and the overall goal of the argument dialogue. Table 1 summarizes these seven types in terms of these criteria. In addition to these types, Walton and Krabbe (1995) posited the existence of complex dialogue, a single sequence of dialogue that contained a shift from one type of dialogue to another. These dialogical shifts occur when the context or topic of the argument changes within the dialogue. Walton's types of argument dialogue framed our understanding of students' and teachers' goals and the goal of the argument dialogue.

ISSN: 2149-8504 (online)

Table 1.
Walton's Types of Argument Dialogue

Type	Initial situation	Participant's goal	Goal of dialogue
Persuasion	Conflict of opinions	Persuade other party	Resolve or clarify issue
Inquiry	Need to have proof	Find and verify evidence	Prove (disprove) hypothesis
Discovery	Find an explanation	Find suitable hypothesis	Discover best hypothesis
Negotiation	Conflict of interests	Get what you want most	Reasonable settlement
Information-seeking	Need Information	Acquire or give information	Exchange information
Deliberation	Practical Choice	Coordinate goals/actions	Decide best action
Eristic	Personal conflict	Hit out at opponent	Reveal deeper conflict

Note. From "Formal Dialogue Models for Argumentation in Education and Linguistics" by D. Walton, 2022, *Learning, Culture and Social Interaction,* Advance online publication.

Toulmin and Walton: Complementary Argumentation Models

Although Toulmin's model has a different focus from Walton's model for describing argumentation, these models are complementary with each other when analyzing classroom-based collective argumentation (Nussbaum, 2011). At their core, Toulmin's and Walton's models of argumentation both position argumentation as dialectical, meaning one party may put forth an argument with the other party providing "counterarguments, refutations, elaborations, questions, and other argument related speech acts to achieve a common purpose" (Nussbaum, 2011, p.87). However, these models seek to describe the argumentation with different grain sizes. Toulmin's model considers the microstructure of arguments; by which we mean the model seeks to describe the content of argument components (e.g., claims or warrants) and distinguish the relationship (i.e., structure) among argument components. In comparison, Walton's types of argument dialogue consider the macrostructure of arguments; by which we mean the model seeks to describe the normative ways in which individuals participate in argumentation and their collective goals.

In order to investigate argumentation at the macrostructure level, we faced the challenge of assessing the teacher's and students' participation goals and the goal of the argumentation in

ISSN: 2149-8504 (online)

the moment. Guided by Walton's (1998; 2010) theory that the content and structure of arguments are shaped by the participants' goals, we used Toulmin's model for argumentation to first model the content and structure of the argument and thus consider the arguments-as-products. Then, using Walton's model, we inferred participants' goals and classified the dialogue type and thus considered the arguments-as-process. Toulmin's model has been used to assess the content and structure of arguments in classrooms (e.g., Conner et al., 2014; Erduran et al., 2004; Rasmussen et al., 2015), but, to our knowledge, no one has used Toulmin's model to assist with the classification of arguments into Walton's types of argument dialogue.

Background and Methods

The Professional Development (PD) Project

The larger project from which this study comes focused on increasing the ability and willingness of elementary teachers to include coding using argumentation into their general curriculum. It consisted of two cohorts of elementary (grades 3-5, ages 8-10 years) school teachers (30 teachers total) from suburban and rural schools in the southeastern United States. These teachers participated in a one-semester PD course, which focused on enhancing teachers' knowledge of collective argumentation and its application within the context of STEM learning, increasing teachers' ability to code robots, and developing teachers' capacity to use collective argumentation in coding activities integrated with content learning.

The Collective Argumentation Learning and Coding (CALC) framework provided the structure for the PD course activities and content for teachers (Conner et al., 2021). The CALC framework includes three elements: teacher support for collective argumentation, choice of tasks, and coding content. Teacher support for collective argumentation is based on a framework in mathematics education and conceptualizes support as providing a direct contribution (e.g., a claim), asking questions to prompt a contribution (e.g., requesting elaboration), or supporting the contribution in some other way (e.g., restating the claim) (Conner et al., 2014). Choice of task element includes the content learning goals, whether the intellectual demand of the task is likely to engage students in higher order reasoning, and the extent to which the task is likely to be motivating and engaging for students. The coding content element includes common coding control structures that are likely to be accessible for elementary students. Sample PD content and activities included having teachers create a set of directions (i.e., pseudocode) to instruct a person how to complete a task, introduction to various coding control structures and their applications, and describing the characteristics of argumentation from videos of mathematics and science instruction.

After the PD course, we followed 10 teachers into their classrooms to support their design and implementation of lessons using the CALC approach and observed how they engaged

ISSN: 2149-8504 (online)

students in argumentation. The research team observed up to three lessons in each participating teacher's classroom.

Study Design and Case Selection

Our qualitative study adapted case study approaches (Stake, 1995; Merriam, 1998) with an instrumental focus as it tests the application of Walton's (1998, 2010) dialogue theory to the phenomenon of argumentation in elementary STEM lessons. We bound the case to episodes of argumentation in two teachers' STEM lessons. Our research question was: What types of argument dialogue are evident in elementary STEM lessons?

Participants and Their Lessons

We chose Sarah and Erica (pseudonyms) for this study because they were the only two teachers observed teaching lessons that included all three disciplines of mathematics, science, and coding with robotics. Sarah was a Gifted Resource Specialist who worked with second through fifth grade students. She had been teaching for over 20 years, her certification was in elementary education, and she taught all levels from Pre-K to fifth grade over the course of her career. Starting in the fall of 2018, her school moved to a push-in co-teaching gifted model, wherein Sarah went into advanced content classes to co-teach with the general classroom teacher. Sarah's school classified this model as Advanced STEAM (Science, Technology, Engineering, Arts, and Mathematics) Inquiry Project-Based Learning; the content Sarah taught was dependent on where the students were in their inquiry-based units.

Sarah's first lesson was a coding-focused lesson with a basis in the engineering design process. Her second was a mathematics lesson on equivalent representations of decimals and fractions. Her third was a science lesson about surface features on Mars. All three lessons involved a coding component, with the first and second lessons involving a greater emphasis on coding than the third.

Erica was an early career teacher with 6 years of teaching experience; 2 years as an elementary STEM teacher for kindergarten through fifth grade students and 4 years as a fourth-grade teacher. Her certification was in elementary education. In fall of 2019, Erica taught fourth-grade students. She described several of her students as advanced in STEM content areas, meaning they were above grade level in content knowledge.

Erica's first lesson was a science lesson focused on simple machines. Her second lesson was a mathematics lesson about scale factors and polygons. Her third lesson was a mathematics lesson about the relationships between distance, time, and speed. All three lessons involved a coding component, with the second and third lessons involving a greater emphasis on coding than the first.

ISSN: 2149-8504 (online)

Table 2.
Overview of Teachers' Lessons

Teacher	Lesson	Goal(s)	Instructional Days	Video Recordings [hh:mm]
Sarah	1	Students will use proportional reasoning to determine the time delay for a robot traveling 6, 12, and 24 inches.	2	01:42
	2	Students will move flexibly among equivalent representations of fractions and decimals and identify ways to shorten coding sequences using mathematical structure.	1	01:35
	3	Students will apply their understanding of surface features on Earth to make predictions about what caused the surface features on Mars.	2[a]	01:46
Erica	1	Students will design and create Rube Goldberg machines and identify at least three different simple machines in their design.	1	09:30
	2	Students will create a coding sequence for a robot to travel the perimeter of a polygon and similar polygon.	1	06:30
	3	Students will understand the relationship between speed, time, and distance by holding one parameter constant, varying another parameter, and then measuring the outcome of the third parameter.	2	07:16

Note. [a] Research team was unable to observe the second instructional day for this lesson.

Data Collection

As part of the PD project, we video recorded three STEM lessons in each of Sarah's and Erica's classrooms. At least two members of the research team videotaped each of the lessons. One camera was used to record the teacher's actions and at least one other camera was used to record small group interactions. This resulted in approximately 5 hours of video recordings from Sarah's lessons and 23 hours of video recordings Erica's lessons to be used for data analysis. There were considerably more hours of video recording from Erica's lessons because multiple cameras focused on small groups of students during Erica's lessons, which also extended over several hours. A research team member collected the tasks and handouts used during the lessons.

Data Analysis

Reduction of the Data for Analysis

At least two members of the research team, one of whom observed and video recorded the lesson, identified and transcribed potential episodes of collective argumentation focused on mathematics, science, or coding content. An episode of argumentation at minimum included a student or teacher making a claim with data and warrant accompanying it, with recognition that sometimes the warrant could be implicit (Toulmin, 1958/2003). We extended the episode if a teacher or student continued the argument by building from a previous argument component. We ended the episode if the collective's data, claim, and warrant did not build on a previous argument component. We were inclusive in our analysis of arguments by including instances when a teacher or student attempted to make a mathematical, scientific, or coding-related claim. We did not limit our analysis to arguments that were deemed mathematically or scientifically correct by our expert opinion. We excluded arguments that were quarrels or "a kind of angry or adversarial verbal exchange based on a conflict between two parties" (Walton, 1998, p. 178). This kind of dialogue is characterized by Walton as eristic. We excluded eristic dialogues because this type of verbal exchange falls outside our definition of collective argumentation and is not included in reform-oriented standards for argumentation in STEM education (CCSS, 2010; K-12 CSF Steering Committee, 2016; NGSS, 2013).

Next, the team met together to reach consensus regarding if the identified episodes included collective argumentation, excluding episodes that did not meet our criteria. We identified 57 (approximately 74 minutes) and 37 (approximately 50 minutes) episodes of argumentation from Sarah's and Erica's lessons, respectively. As part of the larger research study, we randomly selected episodes to obtain at least 5 minutes of small-group arguments and 10 minutes of whole-class arguments for each teacher observed. This random selection was done iteratively. We kept randomly selecting episodes until each threshold was met. We decided to limit the data by random selection to provide a balanced and representative selection of episodes of collective argumentation across all the teachers observed as part of the larger research study.

Creation of Extended Toulmin Models and Transcripts

To represent the content and structure of the collective argumentation in the lessons, we created extended Toulmin's (1958/2003) diagrams for each episode of collective argumentation (as described in Conner, 2008; See Figure 1). A subgroup of at least two research team members watched the episodes of argumentation and examined tasks used in the lesson, enriched the transcripts with teacher's and students' gestures, and developed extended Toulmin diagrams. If the subgroup could not reach consensus on the Toulmin model for an episode, then the subgroup would ask members of the research team that observed and video recorded the lesson for their

ISSN: 2149-8504 (online)

input on the context of the episode and their interpretation of the argument's structure. The research team met until consensus was reached on the extended Toulmin models for each episode. These extended Toulmin diagrams and annotated transcripts were our primary data source for this study.

Interpretation of the Argument Dialogue from the Extended Toulmin Model and Transcripts

To analyze the types of argument dialogue in the lessons, we began examining the content and structure of the extended Toulmin models and transcripts to categorize each episode of collective argumentation into one of Walton's seven dialogue types by using an adapted version of a decision tree to determine the type of argument dialogue (Walton & Krabbe, 1995, p. 81). In the original decision tree by Walton and Krabbe, discovery dialogue was not included. Discovery dialogue was proposed by McBurney and Parsons (2001) as an additional type of argument dialogue. Walton (2019) accepted discovery dialogue as a new type of argument dialogue and agreed with McBurney and Parsons's distinction between inquiry and discovery dialogues. In an inquiry dialogue, the statement to be proved true is set at the beginning of the dialogue, whereas in discovery dialogue, the truth of a statement only emerges during the dialogue. Therefore, there is no statement set early on to be proven or disproven in a discovery dialogue. Other educational researchers (Macagno, 2022; Rapanta & Christodoulou, 2022) have also taken up this distinction between inquiry and discovery argument dialogues. We added the question "Is the aim to prove or disprove?" to our adapted tree to distinguish between inquiry and discovery dialogues (see Figure 2).

As an example of applying our adapted decision tree, when determining if there was conflict during the argumentation episode, we looked at structural features within the extended Toulmin models that may suggest conflict, such as rebuttals or competing claims. In the Findings section, we explicate how we used the decision tree in Figure 2 to interpret the extended Toulmin models to identify the type of argument dialogue. We also describe general trends between the extended Toulmin models and argument dialogue types.

To confirm our interpretation of the argument dialogue type based on the extended Toulmin models, we triangulated our interpretation by going back to the annotated transcripts or video recording to identify confirming or disconfirming evidence of the initial situation, teacher's and students' goals, and the goal of the dialogue for our interpretation of the dialogue type as recommended by Walton and Krabbe (1995). At least two research team members met to discuss their classification with evidence. If consensus could not be reached among the subgroup, the episode of argumentation was brought to the entire research group and was discussed until consensus was reached.

Figure 2. Decision Tree for Classifying the Type of Argument Dialogue

Note. Adapted from *Commitment in dialogue: Basic concepts of interpersonal reasoning* by D. Walton and E. C. Krabbe, 1995, p. 81, SUNY Press.

Findings

We begin by providing an overview of the types of argument dialogue across the three lessons in each of Sarah's and Erica's classrooms. We then present illustrative argumentation episodes for each dialogue type that we found in these elementary STEM lessons to answer the question of what types of arguments were evident in elementary STEM lessons. Our purpose in presenting these episodes is to demonstrate how formal models of argumentation theory (Toulmin and Walton) can be applied to model the argumentation in elementary classrooms. We also use these illustrative episodes to highlight the content, structure, and dialogue goals within a specific argument dialogue and summarize argument diagram trends across dialogue types.

ISSN: 2149-8504 (online)

Sarah's and Erica's Enactment of Argumentation and Types of Argument Dialogue

We randomly selected 26 and 21 episodes of argumentation from Sarah's and Erica's three lessons, respectively, for a total of 47 episodes of argumentation. Two of the episodes included a dialogical shift between argument dialogue types and thus we analyzed a total of 49 argument dialogues (See Table 3). A total of 5 episodes were chosen where the primary focus of the dialogue was coding, 21 episodes were chosen with a primary focus of mathematics, and 25 episodes were chosen with a primary focus of science practices.

The majority of the 49 argument dialogues were either information-seeking (17, 35%) or deliberation (14, 29%). These types of argument dialogue appeared in arguments when the contents focused on mathematics, science, or coding. Also appearing in the data with some frequencies were persuasion (7, 14%) and discovery (7, 14%) argument dialogues. Persuasion and discovery argument dialogues were only evident when the content of the argument centered on mathematics and science. To an even lesser extent, negotiation and inquiry argument dialogues were evident in the data. There were two negotiation argument dialogues in Erica's science lesson. One inquiry argument dialogue was found in each of Sarah's and Erica's lessons with the content of the argument focused on mathematics. There were no eristic argument dialogues because we excluded these types of dialogues from analysis.

Table 3.
Argument Classifications by Teacher and Disciplinary Focus

Walton's Types of Argument Dialogue	Sarah's argument dialogues			Erica's argument dialogues			Total
	Coding focused	Math focused	Science focused	Coding focused	Math focused	Science focused	
Persuasion	0	2	1	0	1	3	7
Negotiation	0	0	0	0	0	2	2
Eristic	0	0	0	0	0	0	0
Information-seeking	2	3	8	0	3	1	17
Deliberation	1	5	1	2	1	4	14
Inquiry	0	1	0	0	1	0	2
Discovery	0	0	2	0	4	1	7
Total	**3**	**11**	**12**	**2**	**10**	**11**	**49**

Note. The information in this table is intended to give a descriptive summary of the nature of argument dialogues in the data. A comparison of the number of argument dialogues across teachers or disciplines is not appropriate because these numbers do not provide a complete representation of the arguments across Erica's and Sarah's lessons. The order of the types of dialogues in the table mirrors the order in the findings.

ISSN: 2149-8504 (online)

A Persuasion Argument

This first illustrative episode of argumentation (diagrammed in Figure 3) is a persuasion type of argument dialogue. It comes from Erica's first lesson during a small-group interaction with Erica present. The students were working on a portion of their Rube Goldberg machine, which was supposed to move a ball into a box (Data 3.2). Prior to the episode, students unsuccessfully tested their machine (Data 3.1). Erica came over after the testing and asked the students how the test went (Support 3.3), prompting this argument. One student claimed they need to shorten the string tying the ball to their robot (Data/Claim 3.1), and another student thought they needed to make the string longer (Data/Claim 3.2).

We used our adapted decision tree (
) to aid our inference for the dialogue type classification. First, we considered whether there were any potential conflicts in the given argument, referring specifically to conflicting points of view (Walton & Krabbe, 1995). In our extended Toulmin models, we observed a conflict is often represented with competing claims, counterclaims, or rebuttals. Figure 3 included a set of counterclaims (Data/Claims 3.3 and 3.4) and a rebuttal (Rebuttal 3.1). Therefore, our analysis concluded that there were potential conflicts in this episode of argumentation. One conflict arising early in the argument was the disagreement about whether a shorter or longer string will solve the students' design issues (Data/Claim 3.3 and Data/Claim 3.5). We determined that resolution was the overall goal because each student was trying to convince Erica and the other student that their claims were correct (Warrants 3.1 and 3.2). However, resolution was not reached because neither student successfully convinced the other of their point of view, as evidenced by the two parallel claims at the end of the argument (Claim 3.6 and Claim 3.8).

ISSN: 2149-8504 (online)

Journal of Research in STEM Education
Vol 8, No 2, December 2022, 79-110

Figure 3. Extended Toulmin Diagram of a Persuasion Argument Dialogue

Note. Numbers in the diagram correspond to references in text. This argument continues in Figure 4.

To check the validity of this finding, we compared this argument with Table 1, which gives the criteria for each type of dialogue according to Walton's classification. In persuasion dialogue, the initial situation should be a conflict of opinions, which fits with this argument because the initial situation was a disagreement about whether a shorter or longer string would work better. The participant's goals should be to persuade the other party, which is true here: The students were seeking to convince Erica and each other. The overall goal of the dialogue should be to resolve or clarify the issue. As stated previously, while resolution was not reached in this

ISSN: 2149-8504 (online)

episode, we determined that resolution was a goal. The episode of argumentation depicted in **Error! Reference source not found.**, therefore, met the all the criteria for persuasion dialogue.

A Negotiation Argument

Our second illustrative example is a direct continuation of the previous persuasion argument dialogue example, with two of the final claims in **Error! Reference source not found.** (Claims 3.9 and 3.11) corresponding to the initial data in **Error! Reference source not found.** (Data 4.1 and 4.2). At the beginning of this episode, the small group had reached an impasse, with neither side able to convince the other of their perspective as evidenced by the two parallel pieces of data (Data 4.1 and 4.2). Erica prompted the students to think about how they could proceed despite this impasse (Support 4.1). Eventually, the students decided they could test both shorter and longer string on their machine to see which one would work better (Claim 4.1).

This episode still involved conflict because no resolution was reached previously. However, resolution is no longer the goal of the dialogue because neither side is attempting to persuade anymore. Walton and Krabbe (1995) elaborated that a settlement is a way of "finding a compromise that will be attractive to both parties" (p. 72). Using this elaboration, we concluded that the goal of this episode was settlement because Erica prompted students to think about what they could do to move forward with their machine design and students came to decide they could try both options.

Erica's support for argumentation instigated this dialogical shift. She questioned how they could proceed despite their disagreement (Support 4.1), prompting the students to look for a compromise instead of remaining fixated on convincing one another that their idea was correct (Claim 4.1). She affirmed their decision to try both suggestions (Supports 4.2 and 4.3). When one student tried to suggest they would only try a longer string if a shorter one didn't work (Support 4.4), she reminded them of their compromise (Support 4.5) and helped students articulate why trying both was a good idea (Warrant 4.1).

We again compared our understanding of the episode of argumentation to the criteria in Table 1 to validate our finding. A negotiation dialogue should have a conflict of interest as the initial situation, and in this episode the conflict is still the disagreement about what length string to use. The goal of the dialogue should be reasonable settlement, and participants' individual goals should be getting what they want most. In this case, the students wanted to try their idea, so the reasonable settlement that gives them each what they want most is the decision to try both lengths. This episode therefore met all the criteria for negotiation dialogue.

ISSN: 2149-8504 (online)

Journal of Research in STEM Education
Vol 8, No 2, December 2022, 79-110

Figure 4. Extended Toulmin Diagram of a Negotiation Argument Dialogue

Note. Numbers in the diagram correspond to references in text. This argument dialogue is a continuation of the dialogue diagramed in Figure 3.

An Information-seeking Argument

The third episode of argumentation (diagrammed in Figure 5) is an example of an information-seeking type of argument dialogue. It comes from Sarah's second lesson during a small-group interaction with Sarah present. The students were tasked with developing a coding sequence that would program a robot to travel around a meter square such that the area enclosed by the robot's path of travel would be six-tenths of the meter square. Students were given a meter square partitioned into 10 equal sized pieces (See Data 5.1 in Figure 5).

ISSN: 2149-8504 (online)

Figure 5. Extended Toulmin Diagram of an Information-Seeking Argument Dialogue

Note. Numbers in the diagram correspond to references in text.

In our analysis of this episode, there does not appear to be any conflicts as evidenced by a lack of competing claims, counterclaims, or rebuttals in the extended Toulmin's model (Figure 5). As in the previous episodes of persuasion and negotiation types of argument dialogue, a conflict can be represented with competing claims, counterclaims, or rebuttals in an extended Toulmin model. Within this episode, there is a lack of a common problem or task to be solved because students and Sarah were still exchanging relevant information to support students in solving the ultimate task of programming the robot. If they were solving the task, then there would have been evidence in students' or Sarah's claims about the time delay for 10 centimeter and the coding structure for the robot. During this episode, the students and Sarah were focused on the lengths of sections from meter square, which were relevant to solving the task, but they had not yet focused on the time delay necessary for completing the task as evident in Sarah's initial prompt, "Now you need to figure out a code then – a delay – that will get you 10 centimeters" (Data 5.2 in Figure 5).

Again, we compared our understanding of the episode of argumentation to the criteria in Table 1 to validate our finding. There was an unstated need of information at the beginning of the episode. Prior to the episode, students determined that one-tenth of a meter was approximately 4 inches. Upon hearing the students were measuring in inches, Sarah asked students to measure in centimeters. Sarah's request for centimeters prompted a need for new measurements for students. The goal of the argument dialogue was to exchange information. Sarah provided information not known to the group; the students did not know that one-tenth of a meter was 10 centimeters. When Sarah stated to students that they should be working in centimeters, students began trying to determine the length of one-tenth of a meter by using a ruler. Upon seeing and hearing students discussing their measurements, Sarah provided new information (Data 5.2 in Figure 5). Students also provided information to Sarah that six tens would be 60 centimeters (Claim 5.1 in Figure 5) at Sarah's request (Support 5.1 in Figure 5). Therefore, Sarah sought to give students information that one-tenth of a meter was 10 centimeters and acquired information from students that six-tenths of a meter was 60 centimeters.

A Deliberation Argument

Our next episode of argumentation is an illustration of deliberation type of argument dialogue. This episode comes from Erica's second lesson. In the episode, a small group of students were attempting to program their robot to travel forward a certain distance and then turn around and come back to the starting distance (Data 6.1). Erica was not present during this interaction. A student claimed the robot needed to turn 180 degrees (Data/Claim 6.1) with reasoning to support the claim by simulating two turns of 90 degrees (Warrant 6.1). A second student questioned whether the turn would be 80 degrees (Support 6.1). A third student misspoke that the turn would be 80 degrees but corrected to confirm the turn is 180 degrees (Support 6.2). As the first student attempted to change the code, they realized that program does not support input values not in the range of -128 to 127 (Data/Claim 6.2 and Warrant 6.2). The students decided to use two lines of code for turning (Claim 6.3 and Support 6.3) because two turns of 90 degrees is a turn of 180 degrees (Warrant 6.3) and meets the constraints of the programming language (Data/Claim 6.2 and Warrant 6.2)

Deliberation dialogue is not adversarial like persuasion dialogues; it is a collaborative dialogue that seeks to solve a practical problem or issue (Walton, 1998). As shown in the diagram (Figure 6), there is no evidence of conflicts (that is, there are no competing claims, counterclaims, or rebuttals). However, diagrams classified as deliberation dialogues may still have these structural elements. As Walton (1998) stated, "In many, but not all, cases of deliberation, there is a conflict between two possible courses of action, and a choice needs to be made between them" (p. 151). Deliberation dialogues are distinguished from information-seeking dialogues in that deliberation dialogues seek to solve a common problem. As evident in the content of Data/Claim

6.2 and Warrant 6.2, the students encountered a problem when the programming language for the robot would not accept numeric values greater than 127. This problem was not a theoretical problem, but it was a practical limitation of the programming language that the students did not expect. A student offered a potential solution by modifying their code to have the robot turn 90 degrees twice to complete the 180-degree turn needed to return to the starting point.

Figure 6. Extended Toulmin Diagram of a Deliberation Argument Dialogue

Note. Numbers in the diagram correspond to references in text.

In summary, the students encountered an initial dilemma when the programming language would not accept their original input of 180-degree turn (Data/Claim 6.2). Together, the students had a set goal of making the robot turn around to come back, which they agreed was a 180-degree turn (Data/Claim 6.1; Supports 6.1 and 6.2). When facing the dilemma of how to turn 180 degrees given the limitations on the numerical value inputs in the programming language, a student offered a potential solution with their reasoning (Claim 6.3 and Warrant 6.3). Collectively, the group agreed this was the best course of action to take given these limitations (Support 6.3). Therefore, the argument met the validation for our finding of this deliberation dialogue with Table 1.

An Inquiry Argument

Our illustrative example of inquiry dialogue comes from a whole class discussion during Erica's third lesson. At this point in the lesson, students had collected data individually about how far their robot could travel in 5 seconds at different speeds, and they had graphed each group's data on the board. Prior to the start of this episode, Erica asked students to describe the relationship between speed and distance. One student suggested that the relationship between speed and distance at a constant time was similar to the relationship between time and distance at a constant speed (Claim 7.1), which they had discussed as a class the previous day (Data 7.2). As a group, the students used their previous discussion and data (Data 7.1 & 7.3) to elaborate that speed and time both affect distance, and specifically, an increase in the robot's speed will cause it to go further (Data/Warrant/Claim 7.2). Erica then asked the rest of the class to consider whether their data supported that claim (Support 7.7), directing their attention to the graph on the board (Support 7.8). The class agreed (Claim 7.3), and with additional prompting from Erica (Supports 7.9, 7.10, & 7.11), they decided that the fact that their distance got farther every time their speed got faster (Warrant 7.3) meant that their data supported the original student claim.

Figure 7. Extended Toulmin Diagram of an Inquiry Argument Dialogue

Note. Numbers in the diagram correspond to references in text.

To categorize this argument dialogue, we first concluded that there was no conflict in this episode in the sense that Walton intended, which eliminated persuasion, negotiation, and eristic dialogues. Next, we had to determine whether there was a common problem to be solved. The common problem was the relationship between speed and distance that the students were exploring. This common problem meant that this could not be an information-seeking dialogue. In this episode, students were not searching for a course of action to pursue, so we decided that this problem was a theoretical one, which meant this dialogue could be either inquiry or

ISSN: 2149-8504 (online)

discovery. The distinction between these two types is particularly subtle, with discovery dialogue seeking to develop a hypothesis or proposition, while inquiry dialogue seeks prove or disprove a proposition. In this case, a student put forward a hypothesis (Data/Warrant/Claim 7.2) and the rest of the class evaluated the validity of this claim (Claim 7.3 and Warrant 7.3). We therefore decided the goal of the dialogue was to prove or disprove, which made this episode of argumentation an inquiry argument dialogue.

As a validity check of our decision of inquiry, we checked the initial situation, participant's goal, goal of dialogue for discovery in Table 1. We concluded that the initial situation for this episode was that students needed to prove the relationship between speed and distance using the evidence they collected with their robots., which fit with the initial situation given for inquiry dialogue. The goal for each participant and goal of the dialogue was to find and verify evidence to prove their hypothesis, which can be seen with Erica's support (Support 7.7) and the final claim (Claim 7.3). These goals matched those from Table 1 for inquiry.

A Discovery Argument

The final episode is the argument dialogue type of discovery, which is diagrammed in

. In this episode, a small group of students, without Sarah's presence, were analyzing a given photograph of Mars. They were tasked with making and justifying claims about the potential causes of the surface features on their photo of Mars. One student put forth the idea that the surface feature could have been caused by water (Claim 8.1). Another student suggested a potential justification for this idea (Warrant 8.1), saying that the pattern on the surface looked like bubbles, and the first student agreed with this reasoning and elaborated on it (Warrant 8.1 and Support 8.1).

Figure 8. Extended Toulmin Diagram of a Discovery Argument Dialogue

Note. Numbers in the diagram correspond to references in text.

In this episode, multiple students contributed to the argument without resulting in any counterclaims or rebuttals, which suggests that there was no conflict of opinions. There was a common problem, as evidenced by the question in the initial data (Data 8.1) and the collaborative creation of the claim and warrant. Since the common problem is not one with an actionable solution, it fit the definition for a theoretical problem. This brought us to the last two types of dialogue, discovery and inquiry. From

, in inquiry dialogue the goal is to prove or disprove a particular proposition, such as the student claim (Data/Warrant/Claim 7.2) in the previous example. In this argument, there was no expectation to prove or disprove anything. The students were instead attempting to make sense of information given to them (i.e., photographs of Mars' surface). This led us to classify this argument as discovery dialogue.

As a final validity check of our decision of discovery, we checked the initial situation, participant's goal, goal of dialogue for discovery in Table 1. We concluded that the initial situation for this episode was that students needed to find an explanation for what caused the surface

ISSN: 2149-8504 (online)

features on Mars, which fit with the initial situation given for discovery dialogue, to find an explanation. The goal for each participant was to find and defend their ideas for what caused the surface features, and the overall goal of the dialogue was to choose the best idea on what caused the surface feature. These goals matched those from the table for discover, which were to find a suitable hypothesis and discover the best hypothesis, respectively.

Discussion

We examined the applicability of Walton's (1998, 2010) types of argument dialogue for describing episodes of argumentation in elementary STEM lessons. Across six lessons with different STEM foci, we transcribed and diagrammed 47 episodes of argumentation using the extended Toulmin's (1958/2003; Conner, 2008) model. Then, we classified the episodes using Walton's types of argument dialogue. We determined that two episodes of argumentation contained a dialogic shift, and so our final analysis consisted of 49 argument dialogues. We found that the two US elementary teachers and their students engaged in six of the seven dialogue types: deliberation, discovery, information-seeking, inquiry, negotiation, and persuasion argument dialogues. We did not find evidence of eristic dialogues because we excluded episodes of argumentation that were purely student quarrels. The results extend previous findings that only identified a subset of these types of argument dialogue (information-seeking, discovery, inquiry, and persuasion) in secondary science and social science classrooms (Rapanta & Christodoulou, 2022). Our results also extend previous findings by revealing elementary students can engage in some of these argument dialogues without the teacher directly facilitating the argumentation (see argument diagrams in Figures 6 and 8).

Types of Argument Dialogue in the STEM Lessons

Our study provides initial insights into the types of argument dialogue present in elementary STEM lessons. The most common type of argument dialogue we identified was information-seeking, followed closely by deliberation. Excluding eristic dialogue, the two least common types were negotiation and inquiry. There could be various reasons for these different frequencies in argument dialogue outcomes: students' knowledge and beliefs about STEM disciplinary practices (Baytelman et al., 2020; Nussbaum et al., 2008), teacher's role in the argument dialogue or their beliefs about argumentation discourse (Conner & Singletary, 2021; Walshaw & Anthony, 2008), or the nature of the instructional task (Felton et al., 2009, 2015; Gilabert et al., 2013).

Overall, few of the arguments in our data were initially adversarial (i.e., persuasion, negotiation, and eristics; see left branch of Figure 2). Most of them were collaborative in nature (i.e., information-seeking, deliberation, inquiry, and discovery; see right branch of Figure 2). Some researchers contend that argument dialogues with an initial adversarial situation are not

ISSN: 2149-8504 (online)

supportive of building knowledge or learning STEM concepts (e.g., Felton et al., 2015). They recommend teachers promote argument dialogues that are more collaborative and cooperative. Other scholars are more inclusive in their assessment of argument dialogue types for building knowledge (e.g., Aberdein, 2020; Rapanta, 2018). For example, Rapanta (2018) posited that persuasion dialogue types, along with information-seeking, discovery, and inquiry dialogues, have pedagogical potential for students to critically examine each other's ideas. Comparably, Aberdein (2020) was even more inclusive of the argument dialogue types that mathematicians engage in to build knowledge: inquiry, persuasion, information-seeking, deliberation, and negotiation. We recognize that argumentation is important for students to develop their understanding of STEM concepts, but our study is unable to clarify if certain types of argument dialogue are more productive for learning STEM concepts.

Learning to Argue in STEM

Both arguing to learn STEM concepts and learning to argue are important in STEM education (Staples & Newton, 2016; von Aufschnaiter et al., 2008). While the nature of our data does not allow for claims about which types of argument dialogue are productive for building STEM knowledge in elementary classrooms, our illustrations and descriptions of the argument dialogues may provide insights about learning to argue in STEM for researchers and teachers.

First, we contend these opportunities to learn to argue within various argument dialogues are vital for preparing a future STEM workforce. If students are limited in their opportunities to learn to argue within certain dialogue types, then they could be ill equipped for their STEM profession. For example, Gainsburg et al. (2016) noted that persuasion is not often a goal of engineer's arguments. They noted that engineers often make use of their knowledge of scientific and mathematical models to deliberate between potential design solutions to a problem. Future engineers with limited opportunities to develop their understandings of how to engage in deliberation argument dialogues but many opportunities to develop their understanding of persuasion argument dialogues may not be prepared to engage in the kind of argumentation needed for their role.

Second, the study provides evidence that elementary students can engage in argumentation for multiple purposes. The students in Sarah's and Erica's classrooms attempted to persuade classmates about design solutions for a simple machine, negotiated a plan to test designs for simple machines, used information about the relationship between centimeters and meters to develop reasoned based claims, deliberated how to modify their coding sequence given input restrictions in the programming language, used data to prove their conjectures about the relationship between speed and distance, and developed reasonable hypothesizes about what caused surface features on Mars. They were able to engage in multiple types of argument dialogue. It is, however, unknown if this is comparable for other students in the US and around

ISSN: 2149-8504 (online)

the world. Future research may extend our findings to other contexts to determine if other teachers provide similar experiences to engage in different types of argument dialogue during STEM instruction.

Contributions to Argumentation Theory and Research

Toulmin and Walton did not develop their models for argumentation for educational research purposes. Still, Toulmin's model has been productive for educational researchers wanting to describe the structure and content of arguments in classrooms (Nussbaum, 2011). In comparison, Walton's model for argumentation has not been taken up to the same extent as Toulmin's; even though there are researchers who have advocated for Walton's model or have applied parts of his theory for educational research (e.g., Felton et al., 2015; Mextaxas et al., 2016; Nussbaum, 2011; Ozdem et al., 2013; Rapanta, 2018; Rapanta & Christodoulou, 2022). Walton (2022) acknowledged that authentic linguistic data from education contexts are important sites for testing and building upon formal argumentation models. In our study of the argumentative discourse in two elementary teacher's STEM lessons, we did not find a need to expand or modify Walton's theory of argument dialogues to be able to accurately describe the nature of the argumentation. When viewed alongside studies such as that of Rapanta and Christodoulou (2022), which applied a modified version of Walton's theory to a secondary science context, our findings support the idea that Walton's types of argument dialogue is applicable for modeling argumentation in educational contexts. Our study specifically identified examples, not yet described by others, of the negotiation argument dialogue type in a STEM educational context.

Our method of analysis contributes to classroom-based argumentation research by coordinating two argumentation models. Argumentation models can serve various purposes. Toulmin's model describes the structure and content of arguments. Walton's types of argument dialogue describe the ways in which people tend to argue. Our method of analysis brought these two purposes together. In the moment of argumentation, educational researchers are not able to probe a teacher's or students' goals. However, researchers are often able to capture the contents of the dialogue with audio and video recordings. From these recordings, researchers have used Toulmin models to describe the content and structure of arguments. Building on Walton's claim that the content and structure of arguments are shaped by participants' goals, we drew on the content and structure in Toulmin's models to make inferences about participants' goals and thus describe a type of argument dialogue. To our knowledge, no one has coordinated Toulmin's model for argumentation and Walton's types of argument dialogue for research in this manner, although Kolstø and Ratcliffe (2007) suggested such a coordination as promising for research and teacher education. Our combining of Walton's types of argument dialogue with Toulmin's model expands the analytical power of Toulmin's models for STEM education researchers to not only be able to describe the content and structure of the argument but also be able to describe the

ISSN: 2149-8504 (online)

nature of the argumentation (i.e., the initial situation, participants' goals, and the goal of the argument dialogue).

We believe coordinating argumentation models can enrich our understanding of the nature of argumentation in STEM classrooms. Other scholars, Macagno et al., (2015), Metaxas et al. (2016), and Ozdem et al. (2013), coordinated Toulmin's model for argumentation and Walton's argumentation schemes, which are classifications for the patterns of reasoning found in warrants and backings (e.g., justifying a claim based on an expert's opinion). Note that Walton's argumentation schemes are different from Walton's types of argument dialogue. Like these scholars, we found the combined perspective of Toulmin and Walton offers a systematic way for researchers to examine the nature of argumentation in often-messy classroom interactions.

Contribution to Practice

The purpose of argumentation is often cast as to persuade others (Meiland, 1989). Walton's types of argument dialogue provide a more holistic description of the varying nature of goals for argumentation and is more aligned to what have been described in the field of STEM education. Providing teachers with Walton's types of argument dialogue may be productive in supporting them to think about the kinds of argumentation in which they want to engage their students. The examples in this paper provide teachers and teacher educators illustrations for each of Walton's types of argument dialogue and exemplify features we found typical of these types of argument dialogue. While we do not claim these arguments are ideal or should be replicated in classrooms, these examples do illustrate the distinguishing features of the types of argument dialogue within a context with which elementary teachers may identify.

Furthermore, combining the extended Toulmin models and Walton's model for argument dialogues may reveal ways in which teachers support students' engagement in certain types of argument dialogue during dialogical shifts. Recall, Erica's supportive actions (Support 4.1-4.3; 4.5) were significant in shifting the persuasion argument dialogue (Figure 3) to one of negotiation (Figure 4). Analyzing teachers' supportive actions during dialogical shifts may illuminate ways teachers support students' engagement with specific dialogue types. Cataloguing and describing these supportive actions may assist teachers in planning and guiding students in learning how to argue and arguing to learn in STEM.

Conclusion

We believe all students need opportunities to engage in a variety of argument dialogues across the STEM disciplines. Whether proving a hypothesis (inquiry), deciding on the best available course of action (deliberation), or resolving an issue to persuade another party (persuasion), students' engagement in these types of argument dialogue (and others) create different opportunities to learn through disciplinary-based argumentative discourse. If students

ISSN: 2149-8504 (online)

are given opportunities to develop their argumentative practice with only a subset of argument dialogues, then they could be ill-prepared for the complexities of argumentation in the STEM disciplines. Our approach of applying Walton's dialogue theory to extended Toulmin models facilitates a systematic way to investigate the content, structure, and opportunities for various goals of argumentation across STEM disciplines.

Acknowledgement

This paper is based on work supported by the National Science Foundation under Grant No. DRL-1741910. Opinions, findings, and conclusions in this paper are those of the authors and do not necessarily reflect the views of the funding agency.

References

Aberdein, A. (2020). Dialogue Types, Argumentation Schemes, and Mathematical Practice: Douglas Walton and Mathematics. *Journal of Applied Logics, 8*(1), 159-182.

Baytelman, A., Iordanou, K., & Constantinou, C. P. (2020). Epistemic beliefs and prior knowledge as predictors of the construction of different types of arguments on socioscientific issues. *Journal of Research in Science Teaching*, *57*(8), 1199–1227. https://doi.org/10.1002/tea.21627

Conner, A. (2008). Expanded Toulmin diagrams: A tool for investigating complex activity in classrooms. *Proceedings of the Joint Meeting of PME,* 32, 361-368.

Conner, A., Crawford, B. A., Foutz, T., Hill, R. B., Jackson, D. F., Kim, C., & Thompson, S. A. (2021). Argumentation in primary grades STEM instruction: Examining teachers' beliefs and practices in the USA. In J. Anderson & Y. Li (Eds.), *Integrated Approaches to STEM Education: An International Perspective* (pp. 427–446). Springer International Publishing. https://doi.org/10.1007/978-3-030-52229-2_23

Conner, A., & Singletary, L. M. (2021). Teacher support for argumentation: An examination of beliefs and practice. *Journal for Research in Mathematics Education*, *52*(2), 213–247. https://doi.org/10.5951/jresematheduc-2020-0250

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

Cross, D. I. (2009). Creating optimal mathematics learning environments: Combining argumentation and writing to enhance achievement. *International Journal of Science and Mathematics Education*, *7*(5), 905–930. https://doi.org/10.1007/s10763-008-9144-9

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's Argument Pattern for studying science discourse. *Science Education*, *88*(6), 915-933. https://doi.org/10.1002/sce.20012

Felton, M., Garcia-Mila, M., & Gilabert, S. (2009). Deliberation versus dispute: The impact of argumentative discourse goals on learning and reasoning in the science classroom. *Informal Logic*, *29*(4), 417. https://doi.org/10.22329/il.v29i4.2907

Felton, M., Garcia-Mila, M., Villarroel, C., & Gilabert, S. (2015). Arguing collaboratively: Argumentative discourse types and their potential for knowledge building. *British Journal of Educational Psychology*, *85*(3), 372–386. https://doi.org/10.1111/bjep.12078

Forman, E. A., Larreamendy-Joerns, J., Stein, M. K., & Brown, C. A. (1998). "You're going to want to find out which and prove it": Collective argumentation in a mathematics classroom. *Learning and Instruction*, *8*(6), 527–548. https://doi.org/10.1016/S0959-4752(98)00033-4

Gainsburg, J., Fox, J., & Solan, L. M. (2016). Argumentation and decision making in professional practice. *Theory Into Practice*, *55*(4), 332–341. https://doi.org/10.1080/00405841.2016.1208072

Gilabert, S., Garcia-Mila, M., & Felton, M. K. (2013). The effect of task instructions on students' use of repetition in argumentative discourse. *International Journal of Science Education*, *35*(17), 2857–2878. https://doi.org/10.1080/09500693.2012.663191

ISSN: 2149-8504 (online)

K–12 Computer Science Framework Steering Committee. (2016). *K–12 Computer Science Framework*. http://www.k12cs.org

Keefer, M. W., Zeitz, C. M., & Resnick, L. B. (2000). Judging the quality of peer-led student dialogues. *Cognition and Instruction*, *18*(1), 53–81. https://doi.org/10.1207/S1532690XCI1801_03

Kim, S., & Hand, B. (2015). An analysis of argumentation discourse patterns in elementary teachers' science classroom discussions. *Journal of Science Teacher Education*, *26*(3), 221–236. https://doi.org/10.1007/s10972-014-9416-x

Kolstø, S. D., & Ratcliffe, M. (2007). Social aspects of argumentation. In S. Erduran & M. P. Jiménez-Aleixandre (Eds.), *Argumentation in Science Education: Perspectives from Classroom-Based Research* (pp. 117–136). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6670-2_6

Krummheuer, G. (1995). The ethnography of argumentation. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 229-269). Erlbaum.

Macagno, F. (2022). Coding relevance. *Learning, Culture and Social Interaction*, *36*, 1-15. Advance online publication. https://doi.org/10.1016/j.lcsi.2019.100349

Macagno, F., Mayweg-Paus, E., & Kuhn, D. (2015). Argumentation theory in education studies: Coding and improving students' argumentative strategies. *Topoi*, *34*(2), 523–537. https://doi.org/10.1007/s11245-014-9271-6

McBurney, P., & Parsons, S. (2001). Chance discovery using dialectical argumentation. In T. Terano & Jinkō Chinō Gakkai (Eds.), *New Frontiers in Artificial Intelligence: Joint JSAI 2001 Workshop Post-Proceedings*. Springer.

Meiland, J. R. (1989). *Argument as inquiry and argument as persuasion. Argumentation, 3*(2), 185–196.

Mercer, N. (2000). *Words and minds: How we use language to think together*. Routledge.

Merriam, S. B. (1998). *Qualitative research and case study applications in education.* Jossey-Bass.

Metaxas, N., Potari, D., & Zachariades, T. (2016). Analysis of a teacher's pedagogical arguments using Toulmin's model and argumentation schemes. *Educational Studies in Mathematics*, *93*(3), 383-397. https://doi.org/10.1007/s10649-016-9701-z

Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education*, *27*(4), 283–297. https://doi.org/10.1007/s11217-007-9071-1

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

Next Generation Science Standards Lead States. (2013). APPENDIX F: Science and Engineering Practices in the Next Generation Science Standards. In *Next Generation Science Standards: For States, By States* (pp. 382–412). The National Academies Press. https://doi.org/10.17226/18290

Nussbaum, E. M. (2008). Collaborative discourse, argumentation, and learning: Preface and literature review. *Contemporary Educational Psychology,* *33*(3), 345-359. https://doi.org/10.1016/j.cedpsych.2008.06.001

ISSN: 2149-8504 (online)

Nussbaum, E. M. (2011). Argumentation, dialogue theory, and probability modeling: Alternative frameworks for argumentation research in education. *Educational Psychologist*, *46*(2), 84–106. https://doi.org/10.1080/00461520.2011.558816

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, *41*(10), 994-1020. https://doi.org/https://doi.org/10.1002/tea.20035

Ozdem, Y., Ertepinar, H., Cakiroglu, J., & Erduran, S. (2013). The nature of pre-service science teachers' argumentation in inquiry-oriented laboratory context. *International Journal of Science Education*, *35*(15), 2559–2586. https://doi.org/10.1080/09500693.2011.611835

Rapanta, C. (2018). Potentially argumentative teaching strategies—And how to empower them. *Journal of Philosophy of Education*, *52*(3), 451–464. https://doi.org/10.1111/1467-9752.12304

Rapanta, C., & Christodoulou, A. (2022). Walton's types of argumentation dialogues as classroom discourse sequences. *Learning, Culture and Social Interaction*, *36*, 1-15. Advance online publication. https://doi.org/10.1016/j.lcsi.2019.100352

Rasmussen, C., Wawro, M., & Zandieh, M. (2015). Examining individual and collective level mathematical progress. *Educational Studies in Mathematics, 88*(2), 259-281. https://doi.org/10.1007/s10649-014-9583-x

Stake, R. E. (1995). *The Art of Case Study Research*. SAGE Publications.

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

Toulmin, S. E. (2003). *The Uses of Argument*. Cambridge University Press. https://doi.org/10.1017/CBO9780511840005 (Original work published 1958)

von Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching*, *45*(1), 101–131. https://doi.org/10.1002/tea.20213

Walshaw, M., & Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Review of Educational Research*, *78*(3), 516–551. https://doi.org/10.3102/0034654308320292

Walton, D. N. (1998). *The New Dialectic: Conversational Contexts of Argument*. University of Toronto Press.

Walton, D. (2010). Types of dialogue and burdens of proof. *Frontiers in Artificial Intelligence and Applications*, *216*, 13–24. https://doi.org/10.3233/978-1-60750-619-5-13

Walton, D. (2019). How the context of dialogue of an argument influences its evaluation. In F. Puppo (Ed.), *Informal Logic: A "Canadian" Approach to Argument* (pp. 196–233). Windsor Studies in Argumentation.

Walton, D. (2022). Formal dialogue models for argumentation in education and linguistics. *Learning, Culture and Social Interaction*, *36*, 1-4. Advance publication online. https://doi.org/10.1016/j.lcsi.2020.100388

Walton, D., & Krabbe, E. C. (1995). *Commitment in dialogue: Basic concepts of interpersonal reasoning.* SUNY Press.

ISSN: 2149-8504 (online)

Webb, N. M., Franke, M. L., Ing, M., Turrou, A. C., Johnson, N. C., & Zimmerman, J. (2019). Teacher practices that promote productive dialogue and learning in mathematics classrooms. *International Journal of Educational Research*, *97*, 176–186. https://doi.org/10.1016/j.ijer.2017.07.009

Whitenack, J. W., & Knipping, N. (2002). Argumentation, instructional design theory and students' mathematical learning: A case for coordinating interpretive lenses. *The Journal of Mathematical Behavior*, *21*(4), 441–457.

Yackel, E. (2001). Explanation, justification and argumentation in mathematics classrooms. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th International Conference on the Psychology of Mathematics Education* (Vol. 1, pp. 9–23). IGPME.

Yackel, E. (2002). What we can learn from analyzing the teacher's role in collective argumentation. *The Journal of Mathematical Behavior*, 21(4), 423-440. https://doi.org/10.1016/S0732-3123(02)00143-8

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, *39*(1), 35-62. https://doi.org/https://doi.org/10.1002/tea.10008

ISSN: 2149-8504 (online)

J-STEM

Journal of Research in STEM Education

j-stem.net